
Gestaltungsorientierte Evaluation und der Return on Investment von Weiterbildungsprogrammen

1 Einführung

Problemstellung und Zielsetzung

Evaluation von Weiterbildung diente in den 1970er Jahren dem Auf- und Ausbau der Weiterbildungspraxis. In den 1980er Jahren sank aufgrund enttäuschter politischer Hoffnungen zunächst das Interesse an Evaluation. Seit den 1990er Jahren gewinnt das Thema wieder an Bedeutung. „Antreiber“ ist diesmal allerdings keine gesellschaftliche Aufbruchsstimmung wie in den 1970er Jahren. Evaluation übernimmt heute zunehmend Funktionen des (Bildungs-) Controllings. Das Ziel von Evaluation scheint weniger die Qualitätsverbesserung von Weiterbildungsprogrammen oder die Bereitstellung von Daten, um eine Entscheidungshilfe für Investitionen zu leisten, als vielmehr die Qualitätskontrolle und die Legitimation bzw. Abwehr von Budgetkürzungen.¹ Zwei Interessen treffen hier aufeinander: Qualitätsentwicklung (pädagogisch motiviert) und Controlling (ökonomisch orientiert). „Rhetorisch ließen sich solche Gegensätze schnell glätten – etwa durch den Hinweis, dass auch das Controlling auf pädagogische Qualitätskriterien verpflichtet werden könne bzw. die Pädagogik Lernerfolge kontrollieren würde. Dennoch bleiben potentielle Konfliktpunkte, die daraus resultieren, dass jeweils eigene Handlungsgewohnheiten durch Prämissen, Denkweisen und Werthaltungen herausgefordert werden.“ (EULER 2005, 19)

Dieses Spannungsverhältnis ist charakteristisch für die betriebliche Weiterbildungs-evaluation. Sinnvoll ist weder eine einseitige Parteinahme noch eine Nivellierung der Interessen. Betriebliche Weiterbildung und betriebliche Weiterbildungs-evaluation ist vielmehr die Einheit der Differenz. Planung, Durchführung und Evaluation von Weiterbildung erfordern eine Bewertung der Ausgangssituation, der Prozesse sowie der Ergebnisse. Bewertung sollte wiederum darauf ausgerichtet sein, Gestaltungsbereiche zu identifizieren und Daten für Entscheidungen zur Verfügung zu stellen. Modelle, die dieses Spannungsverhältnis aufgreifen, liegen bislang nur in Ansätzen vor. Andere Autoren kritisieren grundsätzlich die fehlende Modellbildung in der Evaluation betrieblicher Weiterbildung.²

¹ KUHLENKAMP fasst die aktuelle Finanzsituation in der Weiterbildung wie folgt zusammen: „Die öffentliche Hand vermindert ihre Anteile an der Weiterbildungs-Finanzierung, die Bundesagentur für Arbeit zieht sich massiv aus der Finanzierung der beruflichen Weiterbildung zurück, die Betriebe vermindern ihre finanziellen Aufwendungen für die Weiterbildung, gesellschaftliche Großorganisationen vermindern ihre finanziellen Aufwendungen für die Weiterbildung, die Finanzierungsanteile der Teilnehmenden an der Weiterbildung steigen und die Europäischen Sozialfonds übernehmen mit den Kofinanzierungen der Länder zunehmend in der Weiterbildung sozialpolitische Funktionen.“ (KUHLENKAMP 2005, Vortrag auf der Sektions-tagung Erwachsenenbildung in Potsdam am 22.09.05).

² „Außer regelmäßig wiederkehrenden Bekräftigungen über die Bedeutung der Evaluation wurden im Kontext der betrieblichen Weiterbildung keine adäquaten praxistauglichen Modelle entwickelt“ (GÜLPEN 1996, 2).

Nachfolgend wird die Entwicklung eines Gestaltungsorientierten Modells zur Evaluation betrieblicher Weiterbildung vorgestellt, das einem wertpriorisierenden Ansatz folgt (vgl. BEYWL et al. 2004, 88). Qualitätsentwicklung und Qualitätskontrolle sollen gleichermaßen Zielsetzungen der Evaluation sein. Die Leitfragen der Modellierung lauten: (1) Wie sollten Weiterbildungsprogramme gestaltet werden? (2) Wie sollten Weiterbildungsprogramme bewertet werden? Evaluation soll hierbei der Verbesserung des Weiterbildungsprogramms dienen: *„the most important purpose of evaluation is not to prove, but to improve“* (STUFFLEBEAM 2002, 283).

Es bestehen weitere Spannungsverhältnisse: Die Gestaltung von Weiterbildung und die Ermittlung von Prozess- und Ergebnisdaten (u.a. Lernerfolg, Return on Investment) erfordern es, Wirkungszusammenhänge, Bedingungen und Ursachen zu identifizieren. Zwischen den Pfaden „Begründungsdiskurs“, der die Interessen des Lernsubjekts in den Blick nimmt und dem „Bedingtheitsdiskurs“, der die Gestaltungsparameter der Lernumgebung und das didaktische Handeln der Lehrenden fokussiert, scheint es wenig Spielraum zugeben.³ M.E. ergänzen sich diese „Diskurse“: Mit dem Gestaltungsansatz bewegen wir uns in der Planung vom Lernenden ausgehend („Begründungsdiskurs“) in Richtung Gestaltung und Entscheidung des Lehr-/Lerndesigns („Bedingtheitsdiskurs“), während wir im Evaluationsansatz vom Lehr-/Lerndesign ausgehend („Bedingtheitsdiskurs“) u.a. die Erfahrungen der Teilnehmenden erfragen („Begründungsdiskurs“). Für die Modellbildung bedeutet dies, dass beide Betrachtungsweisen zu berücksichtigen sind.

Training „on-the-job“ bzw. „near-the-job“ und Training „off-the-job“ werden oftmals als gegensätzliche Formen diskutiert und mit Attributen wie „traditionell“ und „modern“ bzw. „wissensorientiert“ und „erfahrungsorientiert“ belegt. Eine Gestaltungsorientierte Evaluation orientiert sich an Problemlagen im Betrieb und deren Auflösung. Weiterbildungsprogramme, die Elemente verschiedener Weiterbildungsformen miteinander verbinden und in einem System integrieren, können auf vielfältige Problemstellungen reagieren. Reflexion und neue Deutungen sind dann ebenso möglich wie Erfahrung und veränderte Handlungsmuster. Für die Modellbildung bedeutet dies, dass eine Anschlussfähigkeit für verschiedene Trainingsformen und weitere Formen der betrieblichen Weiterbildung gewährleistet sein muss.

Der gestaltungsorientierte Evaluationsansatz integriert in einem Modell verschiedene Aktivitäten (vgl. Kapitel 3). Werden die „Stakeholder“ (engl. to have a stake in = Interesse haben an; Stakeholder sind alle an einem Weiterbildungsprogramm interessierte, potentiell interessierte, beteiligte und/oder betroffene Personen bzw. Personengruppen) nicht frühzeitig an dem Prozess beteiligt, bestehen unausbalancierte Interessen und das Modell wird ein Instrument zur Durchsetzung der betrieblichen Disziplin. Entwicklung erfordert Dialog, Auseinandersetzung und Herausforderung bestehender Deutungs- und Handlungsmuster. In diesem

³ „Wir haben aufzuweisen versucht, dass menschliches Handeln generell im Medium eines „Begründungsdiskurses“ geschieht und dass dementsprechend die Vorstellungen der traditionellen Psychologie, sie analysiere „Bedingungen“ oder „Ursachen“ von Handlungen ein folgenschweres Selbstmißverständnis darstellt.“ (HOLZKAMP 1996, S. 30).

Sinne könnte durch die frühzeitige Einbindung der Stakeholder das System Organisation in einen expansiven Lernprozess eintreten. (vgl. ENGESTRÖM 1999)

Aufbau des Artikels

Im zweiten Kapitel werden drei einschlägige Evaluationsmodelle vorgestellt: KIRKPATRICK (Kapitel 2.1), PHILLIPS (Kapitel 2.2) sowie STUFFLEBEAM (Kapitel 2.3). Im dritten Kapitel werden die drei Modelle im Gestaltungsorientierten Evaluationsmodell verbunden (Kapitel 3.1) und an einem Anwendungsbeispiel illustriert (Kapitel 3.2). Das Kapitel und der Artikel schließen mit einem Ausblick (Kapitel 3.3) und einer kurzen Schlussbemerkung (Kapitel 4). Die Ermittlung des Return on Investment (ROI) von Weiterbildungsprogrammen wird besprochen in den Kapiteln 2.2 & 3.1 (Modellbildung) sowie im Kapitel 3.3 (Anwendungsbeispiel).

2 Modellierung von Weiterbildungsevaluation

Auf der Basis des Vier-Ebenen-Modells von DONALD L. KIRKPATRICK (Kapitel 2.1), das JACK J. PHILLIPS um eine fünfte Ebene erweiterte, werden auf Seite der Ergebnisse verschiedene Parameter erfasst, die u.a. für die Ermittlung des Return on Investment von Weiterbildungsprogrammen erforderlich sind (Kapitel 2.2). Diese Output-Orientierung ist für einen Gestaltungsansatz in ein umfassenderes Verständnis von Evaluation einzubetten. DANIEL L. STUFFLEBEAM hat bereits vor 40 Jahren (STUFFLEBEAM 1966) seinen Ansatz veröffentlicht und seitdem fortlaufend weiterentwickelt. Im CIPP-Modell wird der Kontext (Context), der Input, der Prozess und das Ergebnis (Product) bewertet (Kapitel 2.3). Die drei Modelle bilden den Referenzrahmen des Gestaltungsorientierten Evaluationsmodells (Kapitel 3).

2.1 Das Vier-Ebenen-Modell von Donald Kirkpatrick

KIRKPATRICK entwickelte 1954 die Grundlagen des Vier-Ebenen-Modells, das er 1959 in vier Artikeln publizierte.⁴ Die Evaluation von Weiterbildung auf Basis des „Kirkpatrick Modells“ birgt Unwägbarkeiten: Was soll mit welchen Methoden auf welcher Ebene von wem, mit welcher Güte und wann gemessen werden? Kann das Ergebnis (Output) bzw. die Wirkung (Outcome) auf die Weiterbildung zurückgeführt werden (Ebene 1 = Reaction/Zufriedenheit der Teilnehmer, Ebene 2 = Learning/Lernerfolg, Ebene 3 = Behavior/ Lerntransfer und Ebene 4 = Results/Unternehmenserfolg)? Von Ebene 1 bis 4 steigt der Schwierigkeitsgrad, Belege für den Zusammenhang von Training (Prozess) und Zufriedenheit/Lernerfolg/ Lerntransfer/Unternehmenserfolg (Produkt) zu finden.

⁴ „In 1954, at the University of Wisconsin in Madison, I (Don) completed my PhD dissertation entitled 'Evaluation a Human Relations Training Program for Supervisors'. Based on the dissertation, I wrote a series of four articles for the T&D on the American Society for Training and Development (ASTD): 'Evaluating Reactions', 'Evaluation Learning', 'Evaluating Behavior,' and 'Evaluating Results.'” (KIRKPATRICK/ KIRKPATRICK 2005, 4).

Die Daten der Ebenen 1 und 2 werden häufig erhoben und bilden den Normalfall der Weiterbildungs-evaluation. Dies gilt insbesondere, wenn (1) die Weiterbildungseinrichtung oder -einheit den Forderungen der DIN EN ISO 9001 folgt und die Kundenzufriedenheit misst (Ebene 1) und wenn (2) die Teilnehmer das Training mit einer Zertifizierung abschließen (Ebene 2). Die Zufriedenheit der Teilnehmer (Reaction), die im Anschluss an einen Lehrgang erhoben wird (Ebene 1), hat vermutlich keine signifikante Auswirkung auf die Ergebnisse der anderen Ebenen. Zumindest konnten ALLINGER und JANAK einen entsprechenden Zusammenhang nicht nachweisen (ALLINGER/ JANAK 1989, 334). Auch wenn Zufriedenheit den Erfolg nicht wahrscheinlicher macht, so könnte Unzufriedenheit jedoch den Erfolg auf den nächsten Ebenen verhindern. Die Reaktion der Teilnehmenden ist aus weiteren Gründen von Bedeutung: „In a situation where they pay, their reaction determines whether they attend again or recommend the training to others from their institution. [...] Reactions of participants should be measured on all programs for two reasons: to let the participants know value their reaction, and to measure their reactions and obtain suggestions for improvement.” (KIRKPATRICK/ KIRKPATRICK 2005, 5).⁵

Auf der zweiten Ebene betrachtet KIRKPATRICK den Lernerfolg der Teilnehmer (learning): „Three things can be accomplished in a training program: (1) Understand the concepts, principles, and techniques being taught. (2) Develop and/or improve skills. (3) Change attitudes.“ (ebd.). KIRKPATRICK definiert fünf „Guidelines“ für diese Ebene: (1) Messe Fähigkeiten, Fertigkeiten, Kenntnisse und Einstellungen vor und nach dem Training. (2) Verwende einen Papier-und-Bleistift-Test für die Messung von Kenntnissen und Einstellung. (3) Verwende einen Performanz-Test für die Messung von Fähigkeiten und Fertigkeiten. (4) Strebe eine 100 % Antwortquote an. (5) Wenn möglich, verwende eine Kontrollgruppe (vgl. ebd., 6).

Auf der dritten Ebene wird untersucht, in welcher Weise sich das Verhalten am Arbeitsplatz aufgrund des Trainings verändert (behavior). Falls das Gelernte am Arbeitsplatz nicht angewendet wird, obwohl auf der zweiten Ebene ein guter Lernerfolg nachgewiesen werden konnte, gilt das Trainingsprogramm als gescheitert. „Therefore, measuring behavior change is necessary, not only to see if behavior has changed, but also to determine the reasons why change has not occurred“ (vgl. ebd., 6). Auch für diese Ebene bestehen „Guidelines“: (1) Wenn möglich, dann evaluiere vor und nach dem Training. Da dies in der Praxis meist unmöglich ist, so bewerte nach dem Training und stelle fest, welches Verhalten sich in welcher Weise verändert hat. (2) Verhaltensänderung und die Entwicklung neuer Verhaltensweisen benötigt Zeit. (3) Befrage einen oder mehrere der folgenden Personenkreise: Teilnehmende, Vorgesetzte, Mitarbeiter von Teilnehmenden und andere, die das Verhalten der Teilnehmer bewerten können. (4) Wiederhole die Bewertung in einer angemessenen Zeit. (5) Wäge die Kosten der Evaluation mit dem möglichen Nutzen ab. (vgl. ebd., 6 f.)

⁵ Die Zitate entstammen dem Buch KIRKPATRICK/KIRKPATRICK (2005) von Vater (Donald) und Sohn (James). Die zitierten Textstellen verfasste Donald Kirkpatrick.

Auf der vierten Ebene werden die Geschäftsergebnisse erhoben (results). Die Fragestellung lautet: In welcher Weise haben sich die Geschäftsergebnisse aufgrund des Trainings verändert? KIRKPATRICK nennt verschiedene Beispiele, die als Parameter den Erfolg des Trainings belegen können: „Results could be determined by many factors including less turn over, improved quantity of work, improved quality, reduction of waste, reduction in wasted time, increased sales, reduction in costs, increase in profits, and return on investment (ROI).“ (vgl. ebd., 7) Diese Ergebnismessung ist für soziale Themen besonders schwierig: “it is difficult if not impossible to measure final results for programs on such topics as leadership, communication, motivation, time management, empowerment” (KIRKPATRICK 1998, 23). Hinzu kommt die Wirkung weiterer Variablen (Störvariablen) wie z.B. andere Programme im Unternehmen, Performanzentwicklung der Konkurrenz oder konjunkturelle Entwicklungen, die das Ergebnis beeinflussen. Um den Effekt des Trainings bestimmen zu können, müssten diese Effekte isoliert werden.

Das Vier-Ebenen-Modell kann mit seiner Perspektive auf die Ergebnisse des Trainings zudem für die Bedarfsanalyse verwendet werden und einen Beitrag zur Kalibrierung von Weiterbildungsprogrammen leisten: „Start with Level 4 (results) and determine with the line-of business managers what needs to happen. Then, work backward with the question, ‘What behaviors (Level 3) need to be put into practice to achieve the desired results?’ Then, ask, ‘What knowledge, skills and attitudes (Level 2) will the targeted employees need to have to behave appropriately?’ Follow by, ‘How will we get them to come to training and be receptive to these changes (Level 1)?’” (KIRKPATRICK/ KIRKPATRICK 2005, 9). In Tabelle 1 wurden die Merkmale der vier Ebenen zusammengefasst.

Tabelle 1: Charakteristika des Vier-Ebenen-Modells von KIRKPATRICK

	Reaction	Learning	Behavior	Results
Ziel	- Um die Selbstbewertung der Teilnehmenden und deren Reaktion zu erfassen - Um Informationen über ggf. negative Einstellung zu erhalten - Um Informationen zur Verbesserung des Trainings zu erhalten - Um die Teilnehmenden als Experten zu würdigen	- Um den Lernerfolg der Teilnehmenden zu messen - Um den Output des Trainings zu messen - Um eine Bewertungsgrundlage zur Entwicklung des Trainings zu erhalten - Um die Erfolgsaussicht des Lerntransfers einschätzen zu können	- Um Veränderungen des Verhaltens am Arbeitsplatz zu erfassen und den Transfer des Gelernten zu bewerten - Um Zusammenhänge zwischen Reaction, Learning und Behavior zu identifizieren	Um die Auswirkungen des Trainings auf den Geschäftserfolg zu erfassen und zu bewerten
Fragestellung	- Wie zufrieden sind die Teilnehmenden mit dem absolvierten Training? Erfragt wird z.B. Kompetenz und Verhalten des Trainers, Relevanz der Inhalte, Lernformen, Medieneinsatz, Organisation und Zeitstruktur, Räumlichkeiten - Welche Pläne haben die Teilnehmenden hinsichtlich der Umsetzung des Gelernten? Erfragt wird z.B. Was wird sich wann wie ändern? Welche Schwierigkeiten nehmen die Teilnehmenden wahr? Welche Unterstützung bräuchten sie?	- Haben die Teilnehmenden die Ziele des Workshops erreicht? Untersucht wird, inwiefern sich Einstellungen verändert haben, Wissen erworben und Fähigkeiten und Fertigkeiten entwickelt wurden. - Was haben die Teilnehmenden wie gut gelernt? - Was haben die Teilnehmenden nicht gelernt?	Wenden die Teilnehmenden das Gelernte am Arbeitsplatz an?	Inwiefern hat das Training Auswirkung auf z.B. die Kundenzufriedenheit, Kosten, Qualität der Arbeit, Qualität der Ergebnisse, Lieferzeiten, Anzahl an Beschwerden, Krankheitsstand
Methoden	- Selbsteinschätzung der Teilnehmenden - Fragebogen oder Interview	Pre-/Post-Test, Beobachtung, Interview, Rollenspiele oder Simulationen	Interview, Beobachtung, 360-Grad-Feedback (von internen und externen Kunden, Mitarbeitern & Vorgesetzten)	Datenanalyse, Kennzahlen
Zeitpunkt	Direkt zum Abschluss des Trainings	Während des Trainings, zum Abschluss des Trainings oder zeitnah im Anschluss an das Training	3 bis 6 Monate nach dem Training	6 bis 12 Monate nach dem Training

2.2 Die Ebene des Return on Investment (ROI) von Jack PHILLIPS

PHILLIPS erweitert das Modell von KIRKPATRICK um eine fünfte Ebene (vgl. PHILLIPS 1994). Er isoliert den Return on Investment von den anderen Ergebnissen der vierten Ebene. Der Return on Investment (ROI) bildet das Verhältnis ab zwischen dem betrieblichen Aufwand von Weiterbildung (Programmkosten) und dessen Ertrag (Programmleistung bzw. Programmnutzen) entsprechend folgender Formel:

$$ROI(\%) = \frac{\text{Programmnutzen} - \text{Programmkosten}}{\text{Programmkosten}} \times 100$$

Ein ROI von z.B. 130% bedeutet, dass für jeden investierten Euro ein Return in Höhe von 1,30 € (netto) erzielt wurde. Entsprechend gibt ein ROI < 100 an, dass die Investition höher als der Return ist. (vgl. PHILLIPS/ SCHIRMER 2005, 31)

Die Definition der Aufwandsgröße (Programmkosten) beeinflusst erheblich den ROI. Dieser kann, je nachdem, welcher Aufwand als Basis verwendet wird, positiv oder negativ ausfallen. Werden nur die tatsächlich anfallenden Kosten des Weiterbildungsprogramms eingerechnet (wie z.B. Raumkosten, Materialkosten, Trainerhonorar) oder auch Aufwände, die z.B. durch die Planung, den Arbeitsausfall oder die Nachbereitung entstehen? Die Bewertung des Ertrags (Programmnutzen) ist gleichermaßen sensibel: Wird nur der finanzielle Ertrag gemessen, der z.B. durch einen erhöhten Umsatz möglich wurde oder auch die Veränderung immaterieller Werte wie z.B. Arbeitszufriedenheit, Arbeitsplatzsicherheit oder Mündigkeit?

Aufwand und Ertrag kumulieren in der Frage, ob ein z.B. positiver ROI (Ertrag > Aufwand) tatsächlich auf das Weiterbildungsprogramm zurück zu führen ist oder ob nicht vielmehr andere Faktoren, wie z.B. eine veränderte Preisgestaltung, eine gestiegene Nachfrage oder sonstige Veränderungen im Umfeld des Unternehmens für den wirtschaftlichen Erfolg verantwortlich sind? Hier kommt der Faktor Zeit zum Tragen:

- (1) Innerhalb welchen Zeitraums ist der ROI zu betrachten? Je größer der Zeitabstand zwischen Weiterbildung und Ergebnismessung ist, desto spekulativer ist die Bewertung des Zusammenhangs aufgrund der Einwirkung von Störvariablen.
- (2) MARTENS hat darauf hingewiesen, dass sich die Verlaufskurven von „Verhaltenstrainings“ mit dem Ziel der Einstellungsänderung von „Lehrtrainings“ mit dem Ziel des Wissenszuwachses unterscheiden. Während kognitive Lernziele schneller zu erreichen sind und die Vergessenrate anschließend höher ist, stellen sich Einstellungsänderungen langsam ein und nehmen dafür im Zeitverlauf nur langsam ab (MARTENS 1987, 101).

Der Aufwand zur Ermittlung des ROI sollte aufgrund der zusätzlichen Kosten nicht für alle Programme betrieben werden, sondern nur für 5 % bis 10 % der Programmaktivitäten. PHILLIPS und SCHIRMER schlagen folgende Schwerpunktsetzung vor:

Tabelle 2: Evaluationsziel je Stufe (vgl. PHILLIPS/ SCHIRMER 2005, 32)

Evaluationsstufe		Evaluationsziel Prozent aller Programme
Stufe 5	ROI	5 - 10 %
Stufe 4	Auswirkungen auf das Geschäft	10 - 20 %
Stufe 3	Anwendungen am Arbeitsplatz	ca. 30 %
Stufe 2	Lernen	40 - 60 %
Stufe 1	Reaktion, Zufriedenheit, geplante Aktionen der Teilnehmer	90 - 100 %

Die Isolierung von Effekten ist für die Aussagekraft des ROI ebenso entscheidend, wie das Bewusstsein, dass es keinen Sinn macht, alle immateriellen Werte für die Berechnung des ROI in monetäre Größen zu konvertieren. Dies bedeutet nicht, dass immaterielle Werte nicht erfasst, nicht konvertierbar und deren Entwicklung und Veränderung nicht bewertet wird. Ihre Qualität bleibt oftmals allerdings nur erhalten, wenn sie monetär nicht quantifiziert werden.

Die Schritte zu Ermittlung des ROI stellen PHILLIPS und SCHIRMER in einem Prozessdiagramm dar (vgl. Abbildung 1).

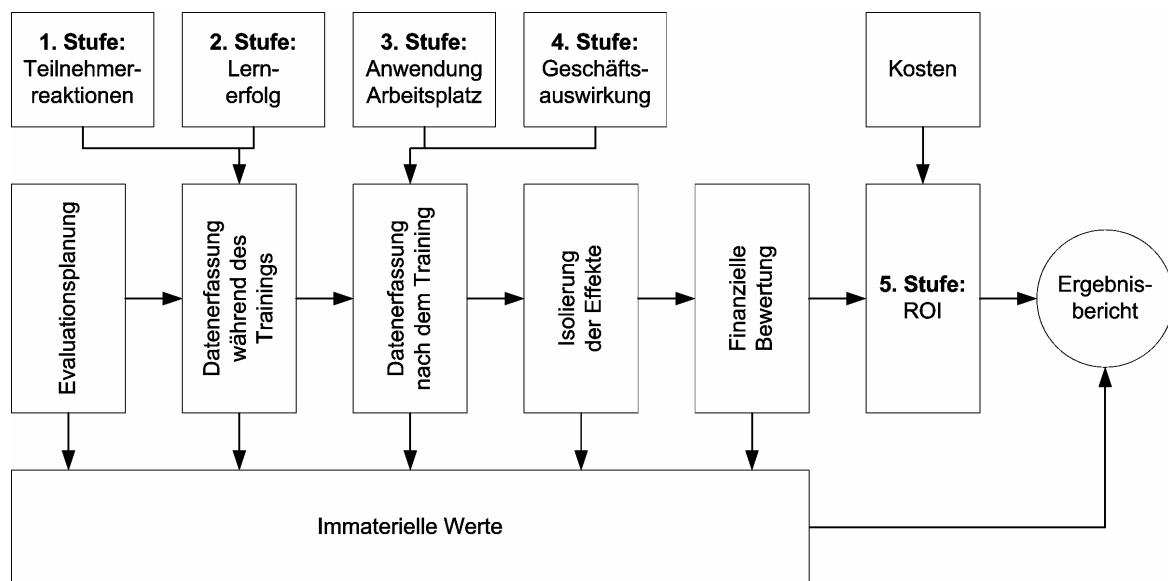


Abb. 1: Der ROI-Prozess nach PHILLIPS (in Anlehnung an PHILLIPS/ SCHIRMER 2005, 23 sowie 125)

PHILLIPS knüpft explizit an das KIRKPATRICK-Modell an (vgl. Abb. 1: 1. bis 4. Stufe). Nachfolgend werden deshalb nur die weiterführenden Prozessschritte „Evaluationsplanung“, „Isolierung der Effekte“, „Finanzielle Bewertung und immaterielle Werte“ sowie „Kosten“ erläutert. Auf den Gesichtspunkt Evaluationsplanung wird zum Abschluss des Teilkapitels eingegangen, da dieser das Verständnis der anderen Prozessschritte voraussetzt.

Isolierung der Effekte

Zur Isolierung der Effekte stellen PHILLIPS und SCHIRMER neun Methoden vor: (1) Kontrollgruppen, (2) Trendlinien-Analyse, (3) Prognose-Methoden, (4) Bewertung/ Einschätzungen durch Teilnehmer, (5) Bewertung/Einschätzungen durch Vorgesetzte, (6) Bewertung/ Einschätzungen durch Management, (7) Bewertung/Einschätzungen durch Experten, (8) Bewertung/Einschätzungen anderer Faktoren, (9) Kunden bewerten die Auswirkungen (vgl. ebd., 79 ff.).

Von Interesse sind insbesondere die zwei zuerst genannten Methoden, da diese nicht auf subjektiven Bewertungen basieren, sondern auf (relativ) objektiven Daten. Überraschend ist allerdings der Vorschlag, dass „Kontroll- und Experimentalgruppe nicht notwendigerweise vor Programmbeginn gemessen [werden]. Die Messungen erfolgen nach Programmende. Der Leistungsunterschied der beiden Gruppen belegt die Auswirkungen des Trainings.“ (PHILLIPS/ SCHIRMER 2005, 85). Dieser Vorschlag wird nicht weiter begründet.

Interessant ist hingegen der Hinweis, dass Kontrollgruppen in der praktischen Durchführung relativ einfach zu bilden sind, indem verschiedene Teilnehmergruppen zu unterschiedlichen Zeitpunkten das Programm beginnen. Durch diese Ungleichzeitigkeit entstehen Vergleichsmöglichkeiten.

Die zweite Methode ist die Trendlinien-Analyse: „Diese Methode verlängert auf Basis früherer Daten eine Trendlinie in die Zukunft. Nach Abschluss des Trainings wird die tatsächliche Leistung mit der projizierten Trendlinie verglichen. Jede Verbesserung der Leistung über die Trendlinie hinaus wird dem Training zugerechnet.“ (PHILLIPS/ SCHIRMER 2005, 89)

Voraussetzung hierfür ist, dass die Daten und eine realistische Trendaussage bereits vor dem Training vorliegen und dass keine neuen (!) Variablen außer dem Training in den Prozess eindringen.

Finanzielle Bewertung und immaterielle Werte

Die Autoren unterscheiden „harte“ und „weiche“ Daten. Die harten Daten gliedern sie in vier Gruppen (P:Produktivität, Z:Zeit, K:Kosten und Q:Qualität) und nennen insgesamt 46 Parameter (z.B. P: Verkaufte Artikel, Z: Zeiten für Meetings, K: Fixkosten, Q: Abweichungen vom Standard).

Die weichen Daten untergliedern die Autoren in sechs Gruppen (A:Arbeitsgewohnheiten, K: Kundenzufriedenheit, AK: Arbeitsklima, EF: Entwicklung/Förderung, E: Einstellung zur Arbeit und I:Initiative) mit insgesamt 32 Parametern (z.B. A: Arbeitstempo, K: Kundenloyalität, AK: Konflikte, EF: Leistungsbeurteilungen, E: Zuversicht in die Zukunft, I: Umsetzung neuer Ideen). (vgl. PHILLIPS/ SCHIRMER 2005, 106 f.)

Für die Konvertierung schlagen die Autoren einen Fünf-Stufen-Prozess vor (vgl. Tab. 3):

Tabelle 3: **Datenkonvertierung (vgl. PHILLIPS/ SCHIRMER 2005, 107 f.)**

Schritt	Beschreibung	Beispiel
1	Einen Parameter auswählen	Anzahl an Beschwerden
2	Basiswert bestimmen	Interne Experten schätzen, dass eine Beschwerde Kosten in Höhe von A € verursachen
3	Veränderung des Parameters ermitteln	6 Monate nach dem Training wird die durchschnittliche Zahl der Beschwerden / Monat ermittelt. Diese liegt 70 % unter der Vergleichszahl des Vorjahres (VV). Nach Isolierung weiterer Effekte (Störvariablen) werden 50 % des Nutzens dem Training zugerechnet ($\frac{1}{2}$ VV).
4	Veränderung auf das vorausgehende Jahr beziehen	Die Veränderung wird als gleich bleibend für das Jahr angenommen ($\frac{1}{2}$ VV x 12).
5	Gesamtwert der Veränderung berechnen	Wert der Veränderung = $A \times \frac{1}{2} \text{ VV} \times 12$

Im Zweifelsfall werden immaterielle Werte nicht konvertiert. Dies gilt insbesondere, wenn keine Methode zur Konvertierung vorliegt. Liegt eine Methode vor, die allerdings aufwendig ist, wird auf eine Konvertierung verzichtet. „Immaterielle Daten sind Daten, die absichtlich nicht in materielle Werte umgewandelt werden.“ (ebd., 24)

Kosten

Auf Seiten der Kosten gilt es zu unterscheiden zwischen Investitionen, die sich zu 100 % „aufbrauchen“ (z.B. Kosten für Räumlichkeiten) und Investitionen, die nach Ablauf des Programms einen Restwert behalten ($0 < x < 100 \%$), da z.B. Daten, Erkenntnisse, implementierte Maßnahmen im nächsten Programmzyklus verwendbar sind. Beispielsweise sind Investitionen in die Bedarfsanalyse, in die Entwicklung eines Programms und in die Evaluation bereits Eingangswerte für den nächsten Programmzyklus (vgl. Tab. 4).

Tabelle 4: **Kostenarten einer ROI-Vollkostenrechnung**

Kostenart	Aufgeteilt	Direkte Kosten
Bedarfsanalyse	✓	
Programmentwicklung	✓	
Programmdurchführung		✓
• Gehälter, Bonus der Trainer		✓
• Programmunterlagen und Lizenzen		✓
• Reisen, Übernachtung und Verpflegung		✓
• Räumlichkeiten		✓
• Gehälter, Boni der Teilnehmer		✓
o Vorbereitungs- und Durchführungszeit		✓
o Reisezeit		✓
Evaluation	✓	
Overhead für Training/Entwicklung	✓	

Die Kosten werden grundsätzlich auf einer Vollkostenbasis erfasst. Hierzu gehören variable und anteilige Kosten für die (1) Bedarfsanalyse, (2) Programmentwicklung, (3) Programmimplementierung und (4) Evaluation/Ergebnisbericht sowie anteilig (5) die Fixkosten (z.B. Kosten für die Verwaltung) (vgl. PHILLIPS/ SCHIRMER 2005, 143).

Evaluationsplanung

Die Evaluationsplanung umfasst (1) die Bedarfsanalyse, (2) den Datenerfassungsplan, (3) den ROI-Analyseplan sowie (4) den Projektplan. Die Bedarfsanalyse wurde im Teilkapitel 2.1 „Das Vier-Ebenen-Modell“ bereits angesprochen. PHILLIPS und SCHIRMER greifen den Vorschlag von KIRKPATRICK auf: „Start with Level 4 (results) and determine with the line-of business managers what needs to happen. Then, work backward“ (KIRKPATRICK/ KIRKPATRICK 2005, 9). Auf Basis der Daten der Bedarfsanalyse werden Programmziele und Evaluationsziele abgeleitet. Das von den Autoren vorgeschlagene Verfahren startet mit der Bestimmung der „Potenziellen Erlöse“. Die weiteren Schritte werden hiervon abgeleitet (vgl. Abb. 2). Unklar ist, wie immaterielle Werte bei diesem Ansatz Ausgangs- und Zielpunkte der Weiterbildung sein können.

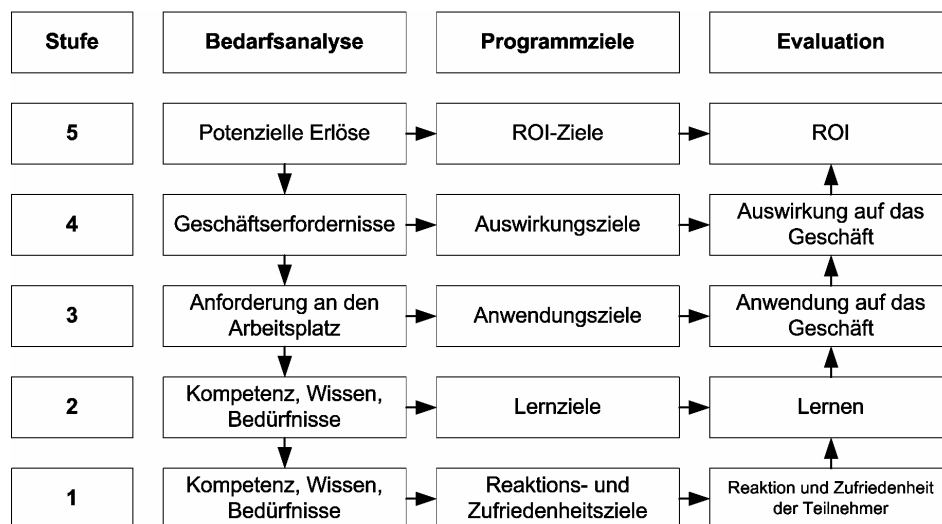


Abb. 2: Bedarfsanalyse, Programmziele und Evaluation (vgl. PHILLIPS/ SCHIRMER 2005, 37)

Im Datenerfassungsplan sind (1) die zu erhebenden Messdaten („harte“ und „weiche“ Daten) jeder Evaluationsebene zu bestimmen und auf die Ziele des Weiterbildungsprogramms zu beziehen. Des Weiteren sind (2) die Ergebnisziele je Ebene zu definieren, (3) die Datenquellen sowie die (4) Methoden zur Datenerfassung je Ebene festzulegen und schließlich (5) der Zeitpunkt bzw. die Zeitpunkte der Datenerhebung zu bestimmen. Im ROI-Analyseplan wird pro (1) Messdatum festgelegt, (2) mit welcher Methode die Effekte isoliert werden und (3) mit welcher Methode die Daten in monetäre Werte konvertiert werden bzw. (4) welches die immateriellen Werte sind, die nicht konvertiert werden. Zudem werden (5) die zu erfassenden Kostenkategorien bestimmt und (6) mögliche Einflussfaktoren (Störvariablen) bzw. notwendige Rahmenbedingungen für die Durchführung identifiziert.

2.3 Das CIPP-Modell von Daniel STUFFLEBEAM

STUFFLEBEAM (2002) unterscheidet vier Formen von Evaluation: Context Evaluation, Input Evaluation, Process Evaluation und Product Evaluation. Die vier Formen ordnet er sachlogisch in einem Flussdiagramm (vgl. Abbildung 3).

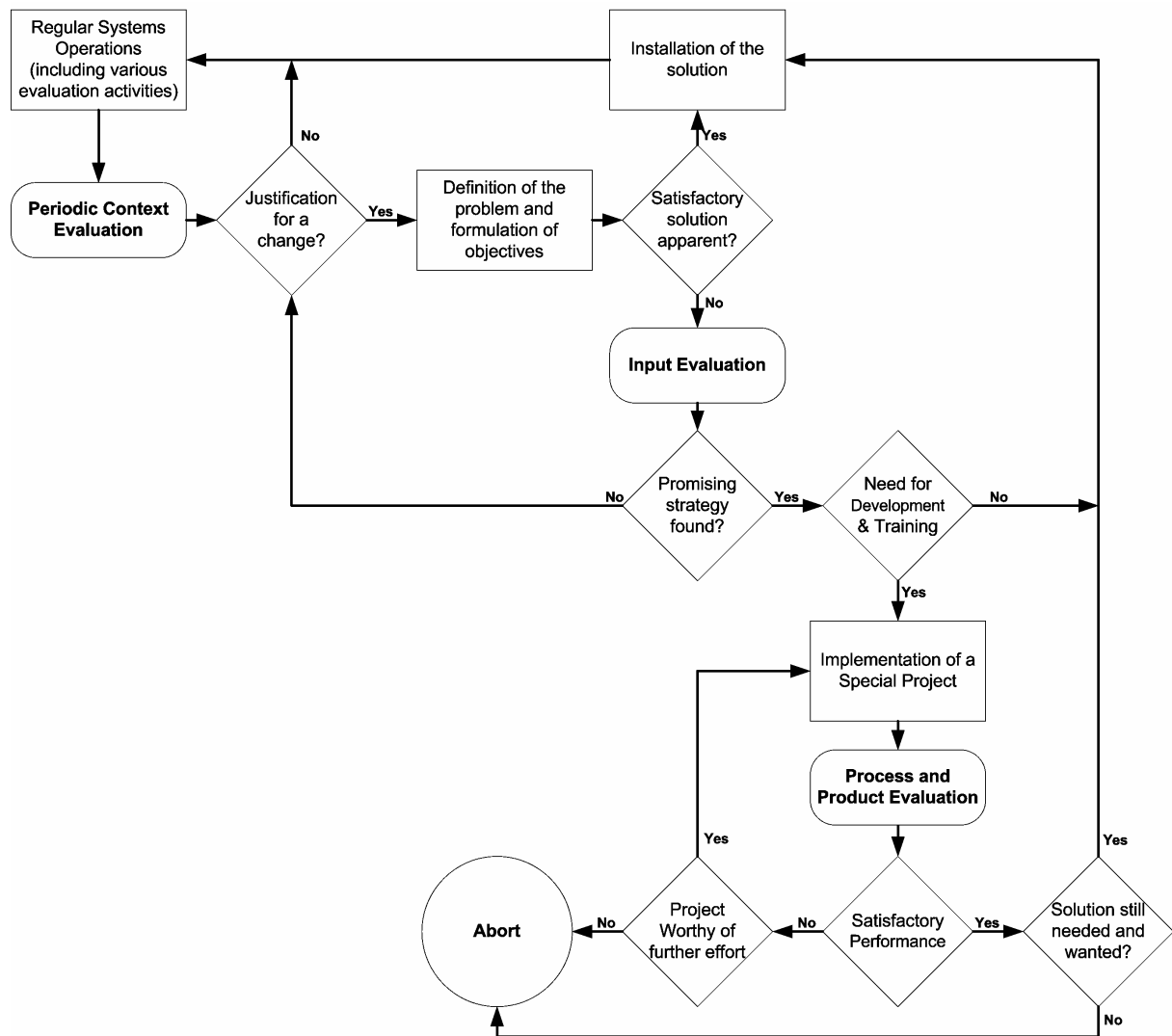


Abb. 3: CIPP-Modell von STUFFLEBEAM als Flussdiagramm
(vgl. STUFFLEBEAM 2002, 284)

Das Flussdiagramm stellt die CIPP-Evaluationen in einer logisch-sachlichen Anordnungsbeziehung dar. Eine Institution (oben links) operiert zunächst auf Basis ihrer regulären Prozesse, wozu auch verschiedene Evaluationen gehören können. In periodischen Abständen findet eine Kontext Evaluation statt. Es werden Bedürfnisse der Stakeholder (Interessengruppen) identifiziert sowie Möglichkeiten definiert, wie auf diese Bedürfnisse zu reagieren ist. „A key concept used in the model is that of the stakeholders. They are those persons who are intended to use the findings, persons who may otherwise be affected by the evaluation, and those expected to contribute to the evaluation.” (STUFFLEBEAM 2002, 280). Die den Bedürfnis-

sen zugrunde liegenden Probleme werden diagnostiziert und die Justierung der bestehenden Ziele überprüft. Eine solche Kontext-Evaluation dient der Erhebung des Ist-Zustandes der Institution mittels einer Selbst- oder Fremdbewertung, um Stärken, Schwächen, Chancen und Risiken und Stakeholder zu identifizieren. Die Bewertung führt zu einer Entscheidung, ob eine Veränderung notwendig ist. Insofern eine Änderung notwendig ist, werden das zu lösende Problem und das Ziel bestimmt. Liegt eine gangbare Lösung vor, wird die Umsetzung angegangen, andernfalls werden in einer Input Evaluation auf Basis der Ziele und Bedürfnisse alternative Lösungsansätze hinsichtlich deren Machbarkeit und Kostenwirksamkeit geprüft. Die Input-Evaluation dient der Identifikation und Bewertung der Kapazitäten der Einrichtung und der Untersuchung der geplanten Prozeduren, Budgets und Zeitpläne. Ziel der Input-Evaluation ist es, eine „Landkarte“ zur Steuerung der Projektimplementation⁶ zu erhalten. Insofern eine geeignete Strategie gefunden ist und ein Entwicklungsbedarf besteht, wird ein Veränderungsprojekt installiert. Die Einrichtung eines solchen Projektes führt direkt zum nächsten Schritt - der Prozess- und Produkt-Evaluation.

Aufgabe der Prozess-Evaluation ist es, Fehler im Evaluationsdesign zu identifizieren, Informationen für Entscheidungen im Prozess bereit zu stellen und Daten des laufenden Prozesses für die spätere Aus- und Bewertung zu sammeln. Es wird das Ziel verfolgt, den Prozess der Implementation zu steuern und zu kontrollieren. Ein weiteres Ziel der Prozess-Evaluation ist die zyklische Messung der Akzeptanz der eingeleiteten Maßnahme bei den Akteuren und anderen Stakeholdern. Veränderungen bei einer Zielgruppe erfordern die zumindest teilweise Übereinstimmung der Ziele der Beteiligten mit den Zielen des Programms (Akzeptanz) und die systematische Berücksichtigung ihrer Bedürfnisse und Ressourcen in der Prozessgestaltung. Insofern handelt es sich bei Stakeholdern weniger um Kunden als um aktiv Mitgestaltende. Der Einbezug der verschiedenen Stakeholder ermöglicht zudem, dass bestehende Interessenkonflikte identifizierbar und Entscheidungen verhandelbar sind.

Die Prozess-Evaluation mündet in die Produkt-Evaluation. Deren Ziel ist die methodische Erfassung und Bewertung der erzielten Ergebnisse. Betrachtet werden beabsichtigte und unbeabsichtigte wie auch kurz- und langfristige Ergebnisse. Die Ergebnisse werden mit den definierten Zielen und den Daten der Kontext, Input und Prozess Evaluation verglichen, um über die Fortführung, Beendigung, Modifikation oder Redefinition der Maßnahme entscheiden zu können.

BEYWL unterscheidet wertdistanzierte, wertrelativistische, wertpositionierende und wertpriorisierende Modelle der Evaluation und ordnet das CIPP-Modell in die letzte Gruppe ein. „Für dieses Modell sind individuelle, organisatorische und soziale Werte/Interessen von zentraler Bedeutung. Sie verfahren in Anerkenntnis eines politischen/ökonomischen Kontexts, in dem Programme durchgeführt werden und Evaluationen stattfinden“ (BEYWL et al. 2004,

⁶ Entsprechend des Kriteriums „Umfang“ können Maßnahmen, Projekte und Programme unterschieden werden. Projekte beinhalten Maßnahmen und Programme umfassen Projekte. Abweichend zu diesem gängigen Begriffsverständnis verwenden wir durchgängig den Begriff „Weiterbildungsprogramm“, da sich der Aspekt der Integration (vgl. Kapitel 3, Rahmung 1) in diesem Begriff am deutlichsten verkörpert. Programme sind wie Projekte zeitlich befristete Vorhaben mit einem definierten Anfangs- und Endzeitpunkt sowie mitlaufenden Restriktionen als gegebene Rahmenbedingungen (z.B. Ressourcen).

88). Die bereits eingangs zitierte Aussage unterstreicht diese Bewertung: „*the most important purpose of evaluation is not to prove, but to improve*” (STUFFLEBEAM 2002, 283). In Tabelle 5 wurden die Merkmale der vier Evaluationsformen zusammengefasst.

Tabelle 5: **Charakteristika des CIPP-Modells (vgl. STUFFLEBEAM 2002, 302)**

	Context Evaluation	Input Evaluation	Process Evaluation	Product Evaluation
Objective	- To define the institutional / service context - To identify the target population and assess its needs - To identify pertinent area assets and resource opportunities for addressing the needs - To diagnose problems underlying the needs - To judge whether goals are sufficiently responsive to the assessed needs	- To identify and assess system capabilities and alternative service strategies - To closely examine planned procedures, budgets and schedules for implementing the chosen strategy	- For identifying or predict defects in the procedural design or its implementation - To provide information for the programmed decisions - To record procedural events and activities for later analysis and judgment	- To collect descriptions and judgments of outcomes - To relate outcomes to goals and to context; input and process information - To interpret the effort's merit and worth
Method	By using such methods as survey, document review, secondary data analysis, hearings, interviews, diagnostic tests, system analysis and Delphi technique	- By inventorying and analyzing available human and material resources - By using such methods as literature search, visits to exemplary programs, advocate teams and pilot trials to identify and examine promising solution strategies - By critiquing procedural designs for relevance, feasibility, cost and economy	- By monitoring the activity's potential procedural barriers and remaining alert to unanticipated ones - By obtaining specified information for programmed decisions - By interviewing beneficiaries, describing the actual process, maintaining a photographic record and continually interacting with and observing the activities of staff and beneficiaries	- By operationally defining and measuring outcomes - By collecting judgments of outcomes from stakeholders - By performing both qualitative and quantitative analyses - By comparing outcomes to assessed needs, goals, and other pertinent standards
Relation to decision making in the change process	- For deciding on the setting to be served - For defining goals and setting priorities - For surfacing and addressing potential barriers to success - For providing assessed needs as a basis for judging outcomes	- For selecting sources of support and solution strategies - For explicating a sound procedural design, including a budget, schedule and staffing plan - For providing a basis for monitoring and judging implementation	- For implementing and refining the program design and procedures, i.e., for effecting process control - For logging the actual process to provide a basis for judging implementation and interpreting outcomes	- For deciding to continue, terminate, modify or refocus a change activity - For presenting a clear record of effects (intended and unintended, positive and negative) - For judging the effort's merit and worth

3 Gestaltungsorientierte Evaluation von Weiterbildungsprogrammen

In Kapitel 3.1 wird das Gestaltungsorientierte Evaluationsmodell vorgestellt. Es basiert auf den Modellen von KIRKPATRICK, PHILLIPS und STUFFLEBEAM (vgl. Kapitel 2). Der Aspekt „Gestaltung“ bezieht sich auf die Konfiguration des Weiterbildungsprogramms und nicht auf die mikro-didaktische Gestaltung einzelner Einheiten (z.B. eines Trainings). In Kapitel 3.2 wird ein Anwendungsbeispiel aus dem Bankensektor vorgestellt. Einen Ausblick auf die Fortsetzung des Forschungsprogramms gibt Kapitel 3.3.

3.1 Gestaltungsorientiertes Evaluationsmodell

Fünf (normative) Rahmungen prägen den Gestaltungsorientierten Evaluationsansatz:

Rahmung 1: Vision und Integration

Weiterbildung kann wirksam werden, wenn Maßnahmen und Weiterbildungsformen aufeinander abgestimmt und in einem Programm integriert sind. Hierfür ist eine gemeinsame Vision als Bezugs- und Integrationspunkt erforderlich.

Rahmung 2: Ziele und Stakeholder

Von der Vision ausgehend werden Ziele abgeleitet, die von den Stakeholdern getragen werden und das Weiterbildungsprogramm steuern. Die Stakeholder sind bereits frühzeitig (möglichst bereits während der Visionsbildung) in den Prozess einzubeziehen.

Rahmung 3: Kontextualisierung und Change Management

Weiterbildung kann wirksam werden, wenn die integrierten Maßnahmen kontextualisiert sind und die Prozesse der Organisation auf die Lern- und Arbeitsprozesse der beteiligten Stakeholder abgestimmt sind. Weiterbildungsprogramme erfordern auf betrieblicher Seite oftmals ein begleitendes Change Management.

Rahmung 4: Commitment

Weiterbildung kann wirksam werden, wenn sich die verantwortlichen, beteiligten und betroffenen Parteien und Personen dem Programm verpflichtet fühlen. Hierfür ist die Selbstverpflichtung der Leitungsebene ebenso erforderlich wie die Selbstverpflichtung der Teilnehmenden.⁷

⁷ REICHERTS und KRUMKAMP benennen vergleichbare Maßnahmen zur Verbesserung des Trainingserfolgs: (1) Operationalisierte Weiterbildungsziele, die u.a. in beobachtbarem Verhalten konkretisiert sind. (2) Begleitende Organisationsentwicklung als Ergänzung zur Weiterbildung sowie die Vermittlung des „Wollens“: „Der Mensch muss den unbedingten Willen haben, die im Lernfeld erworbenen Dispositionen als Mitarbeiter im Funktionsfeld wirklich zu realisieren“ (REICHERTS/ KRUMKAMP 1987, 269).

Rahmung 5: Gestaltung und Evaluation

Gestaltung erfordert Evaluation und Evaluation erfordert Gestaltung. Die Evaluationsplanung ist Bestandteil der Gestaltungsarbeit vor Beginn der Implementation des Weiterbildungsprogramms. Bedarfsanalyse und Zielsetzung zu Beginn unterstützen die anschließende Durchführung und helfen bei der Implementation des Programms und der Bewertung. Wenn die zu erzielende Verbesserung der Performanz definierbar ist (Evaluation), dann sind auch Wege bestimmbar, diese zu erreichen (Gestaltung).

In Anlehnung an STUFFLEBEAM und WITTMANN kann eine Programmevaluation einzelne oder mehrere der folgenden Komponenten umfassen: Kontext, Input, Prozess und/oder Produkt (vgl. Abb. 3: Das CIPP-Modell von STUFFLEBEAM als Flussdiagramm) bzw. Programmträger, Handlung, Klientel, Hilfsmittel, Ziele, Annahmen bzw. Theorien und/ oder Unterstützungsfunktionen (vgl. WITTMANN 1985, 22 f.).

Programmevaluation dient im Gestaltungsorientierten Modell der methodischen Erfassung, Dokumentation und Bewertung des Kontextes, des Inputs, des Prozesses sowie der Produkte von Weiterbildungsprogrammen zur Gestaltung und Entwicklung des Programms durch fortlaufende Ergebnis- und Wirkungskontrolle sowie reflexive Steuerung auf Ebene des Kontextes, des Inputs, des Prozesses sowie der Ergebnisse.

Das Vier-Ebenen-Modell von KIRKPATRICK (vgl. Kap. 2.1) sowie das ROI-Modell von PHILLIPS (vgl. Kap. 2.2) wurden im Gestaltungsorientierten Evaluationsmodell (vgl. Abb. 4) eingebettet in das CIPP-Modell von STUFFLEBEAM (vgl. Kap. 2.3).

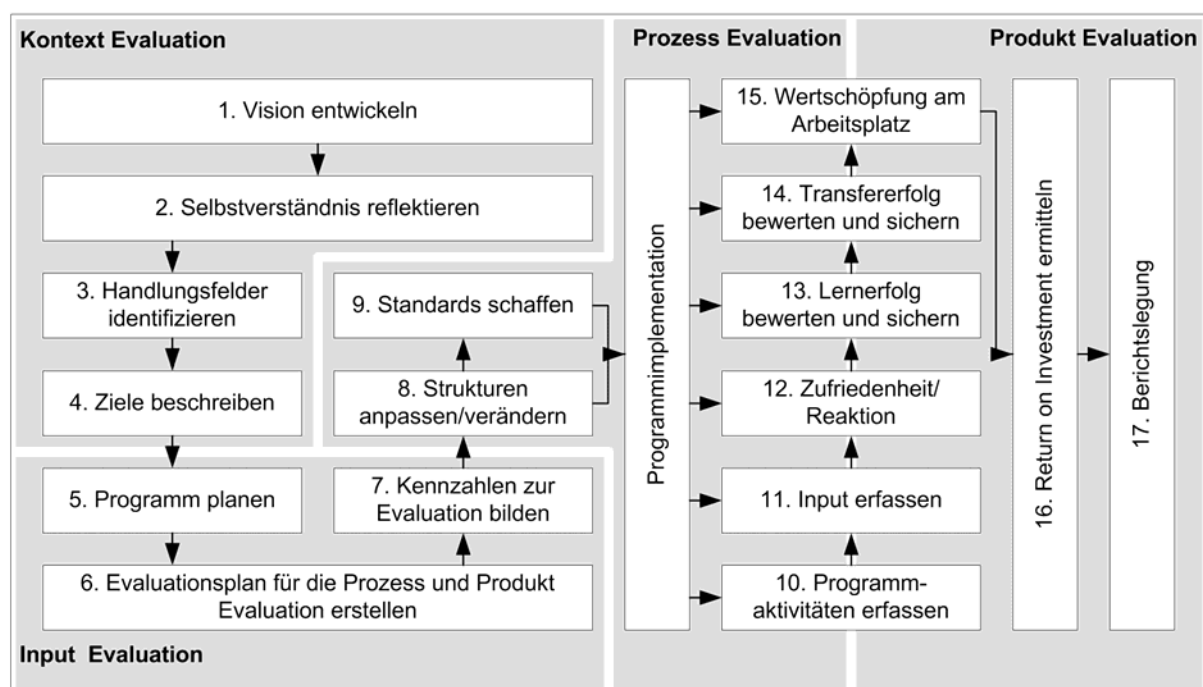


Abb. 4: Gestaltungsorientiertes Evaluationsmodell

Das Gestaltungsorientierte Evaluationsmodell gliedert sich in die vier Evaluationsformen von STUFFLEBEAM (vgl. Kapitel 2.3) sowie insgesamt 17 Prozessschritte. Die der Prozess- und Produkt-Evaluation zugehörigen Schritte 12 bis 15 entstammen dem Modell von KIRKPATRICK (vgl. Kapitel 2.1) und der Prozessschritt 16 basiert auf dem ROI-Modell von PHILLIPS (vgl. Kapitel 2.2). Die Prozessschritte werden nun erläutert.

Zu Beginn wird eine **(1) Vision** für den Geschäftsbereich entwickelt, welche die zukünftige Ausrichtung der Organisation oder Organisationseinheit zum Gegenstand hat. Entscheidend für den Erfolg des Weiterbildungsprogramms ist die Teilhabe der Stakeholder an diesem Prozessschritt. Die Vision (lat. visio = Anblick, Erscheinung, Vorstellung, Idee) ist in einem allgemeinen Verständnis zunächst einmal das Bild einer möglichen Zukunft. Die Unternehmensvision ist das Wunschbild einer zukünftigen Unternehmensidentität. Dieser erwünschten Zukunft wird eine imaginäre Anziehungskraft zugeschrieben. Die Vision vermittelt zudem eine gemeinschaftliche Orientierung, die divergente Ideen zu integrieren vermag. Der Anfang der Unternehmensvision liegt in der Entwicklung persönlicher Visionen. „Eine Organisation, die gemeinsame Visionen aufbauen will, ermutigt ihre Mitglieder dazu, ihre persönlichen Visionen zu entwickeln. Wenn Menschen keine eigene Vision haben, können sie sich nur für die Vision eines anderen ‚vertraglich verpflichten‘. Das Ergebnis ist lediglich eine Einwilligung, nicht Engagement“ (SENGE 1996, 258).

Anschließend wird das **(2) Selbstverständnis** der Mitarbeiter reflektiert, das sich auf Basis der *Konstanz der Erwartungen* gebildet hat und welches die Mitglieder der Organisation oder der Organisationseinheit auf sich selbst beziehen (Selbstbild) bzw. welches von anderen Personen (z.B. Kunden, Gesellschaft) auf die Organisation oder Organisationseinheit bezogen wird (Fremdbild). Von der Vision wird ein Selbstverständnis abgeleitet und mit dem aktuellen Selbstverständnis kontrastiert. Auf Basis der Differenzwerte werden **(3) Handlungsfelder** identifiziert.

Die Handlungsfelder werden in messbare **(4) Zieldefinitionen** „übersetzt“. Aus einer oftmals abstrakten Vision werden nun konkrete Ziele. Das Erreichen dieser Ziele führt zu einer Auflösung der Handlungsfelder in Form einer veränderten Praxis. Ausgerichtet auf die Ziele werden die **(5) Maßnahmen des Weiterbildungsprogramms** geplant. Verantwortlichkeiten, Ressourcen, Zeitpunkte oder -räume sowie Schnittstellen zu anderen Bereichen oder Personen (Stakeholdern) werden ermittelt. Auf Basis dieses Programms wird der **(6) Evaluationsplan** für die Prozess- und Produkt-Evaluation festgelegt. In einem vorerst letzten Schritt werden auf Basis der Ziele und der Programmmaßnahmen **(7) Kennzahlen** (Key Performance Indicator – KPI) gebildet, anhand derer der Erfüllungsgrad der Ziele und der Fortschrittsgrad der Maßnahmen bewertbar ist.

Im Übergang von der Kontext-/Input-Evaluation zur Prozess-/Produkt-Evaluation findet eine Entscheidungsklausur statt. Diese Klausur ist bereits Teil der **(8) Anpassung der Strukturen**, da die Akzeptanz des Programms und die Selbstverpflichtung der Stakeholder erst eine Veränderung und Anpassung der Strukturen ermöglicht. Werden die Ergebnisse der Prozessschritte 1 bis 7 akzeptiert, können **(9) Standards** für die Programmaktivitäten festgelegt wer-

den. Die für die Schritte 8 und 9 erforderliche Akzeptanz und die notwendige Selbstverpflichtung (Commitment) sind abhängig von der Beteiligung der Stakeholder an den Prozessschritten 1 bis 7. Die Beteiligung sichert sowohl die richtige Justierung des Programms als auch die spätere Identifikation mit dem Programm. Wird der Plan abgewiesen, ist eine Neuausrichtung zu verhandeln. Die Prozessschritte 8 und 9, „Anpassung der Strukturen“ sowie „Festlegung von Standards“, sind bereits Teil der **Programmimplementation**.

Die weiteren Schritte der Prozess- und Produktevaluation folgen weitgehend den beschriebenen Modellen von KIKRPATRICK und PHILLIPS.

(10) Programmaktivitäten: Es wird erhoben, wie oft eine Maßnahme angeboten wird (Auflagenhöhe), wie stark sie nachgefragt wird (Teilnehmeranzahl) und welcher Art der Abschluss ist (Zertifikat). Diese Daten haben statistischen Charakter und dienen dem Geschäftsbereich Personal zur internen Steuerung.

Auf der gleichen Stufe (Ebene 0) wird fortlaufend der **(11) Input** bewertet. Hierzu gehören u.a. Geldeinsatz, Personaleinsatz und als erhebliche Kostengröße der entstandene Arbeitsausfall während der Weiterbildung. Die Inputgrößen werden für alle Maßnahmen erhoben. Die Werte gehen in die Berechnung des ROI ein.

Die erste Ebene erfasst die **(12) Zufriedenheit der Teilnehmer (Output I)**. Die Daten werden in der Evaluation berücksichtigt, obwohl ihre Aussagekraft vermutlich gering ist (siehe Kap. 2.1).

Zwei bis vier Wochen nach dem Seminar findet ein Gespräch zwischen dem Mitarbeiter, der am Training teilnahm, und dem jeweiligen Vorgesetzten statt. In diesem Gespräch wird einerseits der **(13) Lernerfolg (Output II)** besprochen und bewertet und andererseits bietet das Gespräch dem Vorgesetzten die Möglichkeit, Transferhindernisse zu identifizieren und notwendige Unterstützungsleistungen zu identifizieren und zu vereinbaren.

Weitere drei bis vier Monate später findet ein erneutes Treffen zwischen dem Mitarbeiter und dem Vorgesetzten statt. Ziel des Treffens ist die gemeinsame Bewertung der bisherigen Umsetzung und die Sicherung und Entwicklung des **(14) Lerntransfers (Outcome I)**. Ein Beispiel hierfür sind die Doppelbesuche⁸ im Pharmabereich. Durch gezieltes Feedback unterstützt der Vorgesetzte den Mitarbeiter beim Lerntransfer. Auf der vierten Ebene wird der konkrete Beitrag zur **(15) Wertschöpfung am Arbeitsplatz (Outcome II)** gemessen. Alle Daten werden im Vergleich zu einem Vorjahres- oder Vormonatswert betrachtet. Etwa sechs bis zwölf Monate nach Abschluss des Trainings wird der **(16) Return on Investment** ermittelt. Hierfür müssen die Effekte isoliert und der finanzielle Wert der Veränderung (Nutzen) ermittelt werden. Abschließend wird der **(17) Ergebnisbericht** verfasst. Das Modell wird nun an einem Anwendungsbeispiel verdeutlicht.

⁸ Ein in der Pharmabranche übliches Vorgehen, bei dem Regionalleiter ihre Mitarbeiter beim Besuch der zu betreuenden Ärzte begleiten und nach den Verkaufs-/ Beratungsgesprächen direktes Feedback geben können, welches beim nächsten Besuch direkt umgesetzt werden kann. Ein ähnliches Vorgehen ist in einer Bank möglich, indem die direkte Führungskraft den Mitarbeiter bei Beratungsgesprächen unterstützt.

3.2 Anwendungsbeispiel

Das Anwendungsbeispiel entstammt dem Bankensektor. Im internationalen Vergleich besitzt Deutschland eine hohe Banken- und Filialdichte und die Produktpalette im Endkundengeschäft ist bei allen Voll-Banken nahezu identisch. Unterschiede bestehen hingegen in der Beratung, im Kundenkontakt bzw. in der Servicequalität. Ziel des Weiterbildungsprogramms war die Verbesserung der Servicequalität. Der Betrachtungszeitraum umfasst die Jahre 2001 bis 2004.⁹ An dem Weiterbildungsprogramm waren Mitarbeiter von über 40 Bankfilialen beteiligt. Mit der Konzeption und Durchführung des Weiterbildungsprogramms wurden externe Organisationsberater beauftragt.

In einem ersten Schritt (Dezember 2000) wurde unter Beteiligung der Regionalverantwortlichen Nord und Süd sowie der Filialverbundleiter in einem eintägigen moderierten Workshop die Vision für das Jahr 2005 formuliert. Die Teilnahme war für die Filialverbundleiter freiwillig und die Hälfte der Führungskräfte nahm an diesem Prozessschritt teil. Filialverbundleiter sind Vorgesetzte von ca. 2 bis 5 Filialleitern. In einem zweiten Schritt (Januar 2001) wurde in einem ebenfalls eintägigen und moderierten Workshop die Vision konkretisiert und ihre Erreichbarkeit wurde mit Unterstützung von Experten aus den Zentralbereichen Marketing, Personal sowie Controlling bewertet. Außer den Experten waren die Regionalverantwortlichen sowie drei Filialverbundleiter, die selbst Filialen leiten, beteiligt. Im Mai 2001 fand ein Koordinierungstreffen zwischen den Regionalverantwortlichen und dem externen Berater statt, um den weiteren Zeitplan zu vereinbaren. In der Zeit zwischen Mai und November 2001 wurden das Verhaltenstraining „Service und Beratung“ sowie das Führungskräfte-Training, beide sollten noch 2001 starten, vorbereitet.

Operationalisierung der Strategie und Implementation des Weiterbildungsprogramms:

Im November 2001 startete das erste dreitägige Training „Service und Beratung“. Dieses Training wurde jeweils für die Filialleiter und Filialmitarbeiter einer Filiale durchgeführt. Das Training fand anschließend zeitversetzt in den beteiligten Filialen statt. Gegenstand des Trainings waren die Teilschritte (1) Vision, (2) eigenes Selbstverständnis, (3) Aktionsfelder identifizieren, (4) Ziele definieren, (5) Maßnahmenplanung, (6) Entwicklung von Kennzahlen und (7) Evaluationsplan vereinbaren. Dem Workshop folgte im Dezember 2001 ein zweitägiges Führungskräfte-Training für die Filialleiter und deren Stellvertreter. Dieses Training hatte den Zweck, die notwendigen betrieblichen Rahmenbedingungen für den Transfer der Ergebnisse des Trainings „Service und Beratung“ zu schaffen. Gegenstand des Trainings waren: (8) Anpassung der Strukturen (inkl. Entscheidungsklausur) und schließlich (9) Definition von Standards für die Programmaktivitäten. Parallel wurden bereits Daten der Evaluationsstufen Ebene 0 (Programmaktivitäten und Input), Ebene 1 (reaction: zum Abschluss des Trainings), Ebene 2 (learning: Rollenspiele im Training sowie 3 Wochen nach Abschluss des Trainings im Gespräch mit dem Vorgesetzten), Ebene 3 (behavior: Mitarbeitercoachings 3 bis 4 Monate nach Abschluss des Trainings) und Ebene 4 (results: monatliche Bewertung des Umsatzes).

⁹ Die Kapitel 3.2 und 3.3 sind unter Mitwirkung von Jürgen Radel M.A. entstanden. Herr Radel promoviert zu dem Thema „Gestaltung und Evaluation von Weiterbildungsprogrammen“ an der Universität Bremen.

Der Return on Investment wurde erstmals nach einem Jahr für das Weiterbildungsprogramm ermittelt.

Die Ergebnisse des Führungskräfte-Training wurden leicht modifiziert erhoben und bewertet. Die Daten der Ebene 0 wurden fortlaufend erhoben. Die Ebenen 1 und 2 wurden während des Trainings bewertet und die Ergebnisse der Ebene 3 waren ab 2003 Gegenstand des Coachings der Filialleiter (Coach: externe Beratung). Die Ebene 4 wurde erst im Rahmen der Reifegradbestimmung der Filialen gegen Ende des Weiterbildungsprogramms bewertet (ab 2004). Hierbei wurden sowohl Marktdaten betrachtet, als auch Mitarbeiterinterviews durchgeführt. Von zentraler Bedeutung war hierbei die Kompetenzentwicklung und Performanz der Filialleiter.

Einen Überblick über den Ablauf des Programms gibt Tabelle 6.

Tabelle 6: Ablauf des Weiterbildungsprogramms

Evaluation	Zeitraum	Prozessschritt
Kontext	Dezember 2000	Entwicklung der (1) Vision für die Strategiefelder „Markt/Kunde“, „Investor“, „Produkte/ Prozesse“ und „Mitarbeiter“
	Januar 2001	Konkretisierung der Vision mit Beteiligung von Experten
	Mai 2001	Koordinierungstag zur Abstimmung des weiteren Vorgehens
	Mai bis November 2001	Vorbereitung der Trainings
	ab November 2001	Training „Beratung und Service“ (BS) für die Mitarbeiter und den Leiter einer Filiale. Gegenstand der Trainings waren folgende Prozessschritte: (1) Vision und (2) eigenes Selbstverständnis (3) Handlungsfelder identifizieren und (4) Ziele definieren (5) Maßnahmenplanung (6) Entwicklung von Kennzahlen (7) Evaluationsplan (12) Zufriedenheit (Output I) (zum Abschluss des Trainings) (13) Lernerfolg (Output II) (während und zum Abschluss des Trainings)
Input		Begleitend zu den Trainings wurden ab November 2001 (10) die Programmaktivitäten sowie (11) der Input erfasst.
	Dezember 2001	Führungskräfte-Training „Leadership 1“ (8) Anpassung der Strukturen (inkl. Entscheidungsklausur) (9) Definition von Standards für die Pgm.-Implementation (12) Zufriedenheit der Teilnehmenden (Output I) (13) Lernerfolg der Teilnehmenden (Output II)
Prozess und Produkt	ab Januar 2002	Fortsetzung der Programmaktivitäten. Teilnehmende des Trainings „Beratung und Service“ wurden von ihren Vorgesetzten begleitet und gecoacht hinsichtlich (14) des Lerntransfers (Outcome I) sowie der (15) Wertschöpfung.
	November 2002 ab Januar 2003	Führungskräfte-Training „Leadership 2“ Fortsetzung der Programmaktivitäten und zusätzlich

		Ermittlung des
		(16) Return on Investment sowie
		(17) Erstellung von Zwischenberichten
	ab Juni 2003	Beginn des Coachings für Führungskräfte (Filialleiter)
Produkt	ab Mai 2004	Ermittlung des Reifegrads der Filialen (insbesondere Kompetenz und Performanz der Filialleiter)
		(14) Lerntransfer der Mitarbeiter (Outcome I)
		(15) Wertschöpfung der Mitarbeiter am Arbeitsplatz
		(16) Return on Investment des Weiterbildungsprogramms
		(17) Abschlussbericht

3.3 Ausblick

Das Weiterbildungsprogramm hat zu einer nachweisbaren Verbesserung der Geschäftsergebnisse geführt, wobei sich der Umsatz in Filialen (Filiale 1 und 2) mit kombinierten Trainings („Service und Beratung (SB)“ sowie Führungskräfte-Training) besser entwickelte als in Filialen (Filiale 3) mit einfachen Trainings (nur Führungskräfte-Training) (vgl. Abb. 5):

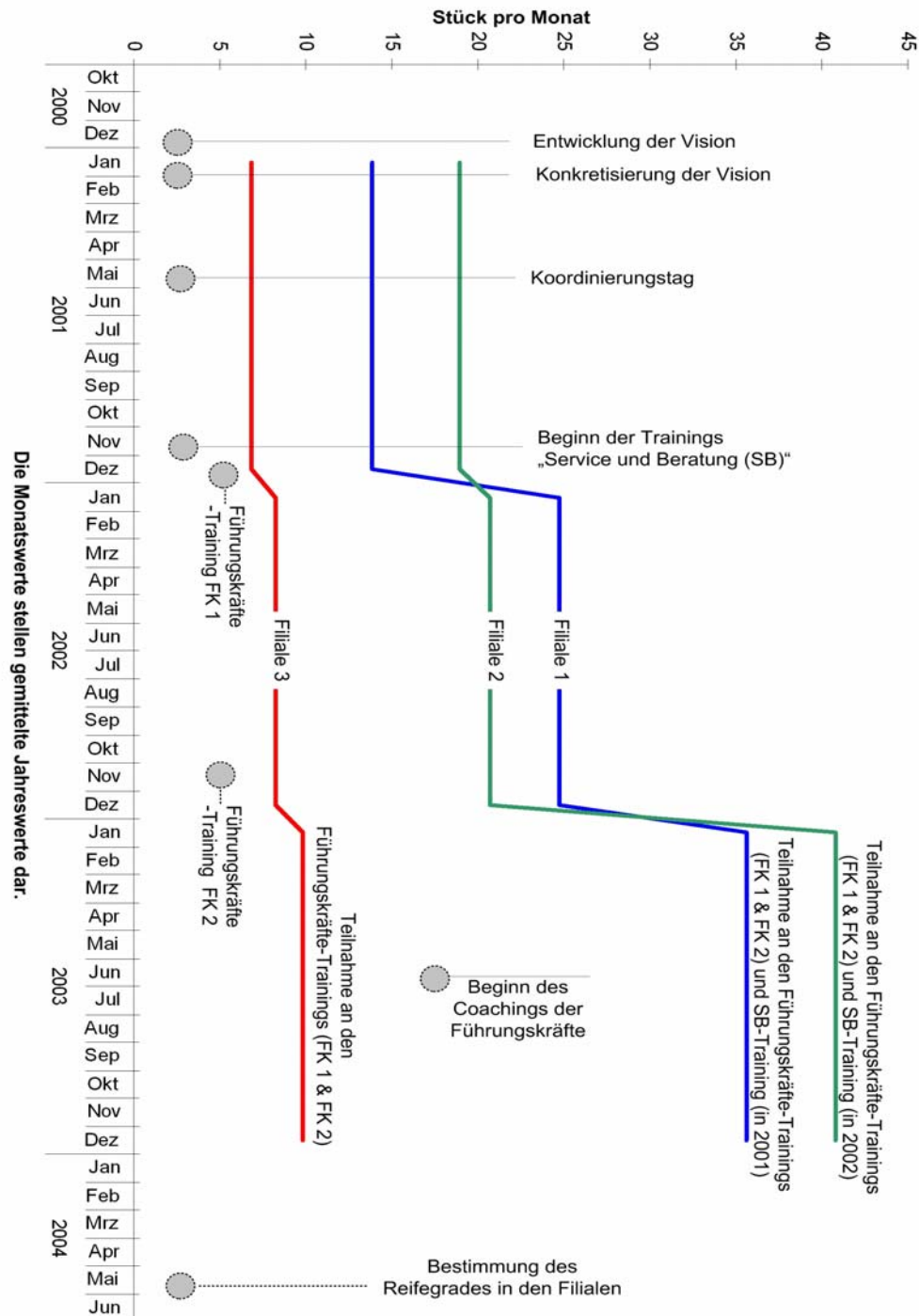


Abb. 5: Geschäftsergebnisse dreier Filialen im Vergleich

Der Return on Investment wurde auf Basis folgender Kostenarten ermittelt:

Tabelle 7: **Kostenarten der ROI-Vollkostenrechnung**

Kostenart	Aufgeteilt	Direkte Kosten
Analyse der Ausgangssituation		✓
Programmentwicklung		✓
Programmdurchführung		✓
• Gehälter der Trainer		✓
• Programmunterlagen		✓
• Reisen, Übernachtung und Tagungspauschale Trainer und Teilnehmer		✓
• Räumlichkeiten und Technik		✓
• Gehälter und Reisezeit der Teilnehmer		✓
o Regionalleiter		✓
o Filialverbundleiter		✓
o Filialleiter		✓
o Mitarbeiter in die Filialen		✓
o Experten (Marketing, Personal, Controlling)		✓
Evaluation		✓
Abzinsung (Inflation pro Jahr) mit 1,7 %		✓
Opportunitätskosten		✓

Die Kostenarten stimmen weitgehend mit dem Vorschlag von PHILLIPS überein (vgl. Tab. 4). Ergänzt wurden ein Abzinsfaktor auf das direkt eingesetzte Kapitel sowie Opportunitätskosten. Diese Kosten entstanden, da durch die Teilnahme am Training Möglichkeiten (Opportunitäten) zur Wertsteigerung am Arbeitsplatz entgangen sind. Zudem wurden einzelne Kosten (z.B. für die Evaluation) nicht aufgeteilt, sondern direkt zugerechnet. Diese Berechnung verschlechtert den ROI. Da der verbleibende Wert nicht zu quantifizieren war, wurde dieser Weg gewählt. Nicht erfasst wurden Overheadkosten, da Kosten, die hier zu berücksichtigen wären, weitgehend direkt erfasst wurden (z.B. Kosten der Experten). Auf der Nutzenseite wurde die Verbesserung des Geschäftsergebnisses über alle Filialen erfasst. Die Veränderung immaterieller Werte wurde monetär nicht bewertet.

Der ROI wurde auf Basis dieser Aufwands- und Nutzengrößen entsprechend der Formel von PHILLIPS (Kapitel 2.2) ermittelt und betrug ROI > 350 %. Der eigentliche ROI lag vermutlich höher, da z.B. Kosten nicht aufgeteilt und immaterielle Wertentwicklungen nicht berücksichtigt wurden.

Die Gestaltungsorientierte Programmevaluation wird derzeit in einer weiteren Branche (Produktion) erprobt. Für das Jahr 2007 ist eine umfassende Publikation der hier skizzierten Daten sowie der Daten der zweiten Programmevaluation aus dem Produktionssektor geplant.

4 Schlußbemerkung

Aufgrund der langen Laufzeit von Weiterbildungsprogrammen ist die Bewertung von Wirkungszusammenhängen kritisch. Dies gilt insbesondere für die Ermittlung des Return on Investment. Selbst konservative Berechnungsverfahren bergen Unsicherheiten. Andererseits ermöglichen gerade Programme in ihrer Summe synergetische Transfereffekte. Der Beitrag der einzelnen Wirkungsfaktoren (u.a. Workshop, Training, Coaching, Besprechung und Beratung) ist jedoch noch unklar. Unklar ist auch, wie immaterielle Werte angemessen berücksichtigt, erfasst und bewertet werden können.

Der vorgestellte Ansatz und die skizzierten Ergebnisse bilden nicht den Abschluss, sondern den Beginn des Forschungsprogramms „Gestaltungsorientierte Evaluation von Weiterbildungsprogrammen“.

Literatur

ALLINGER, G.M./ JANAK, E.A. (1989). Kirkpatrick's Levels of Training Criteria. Thirty years later. In: Personnel Psychology, 42, 331-342.

BEYWL, W./ SPEER, S./ KEHR, J. (2004): Wirkungsorientierte Evaluation im Rahmen der Armuts- und Reichtumsberichterstattung. Perspektivstudie im Auftrag des Bundesministeriums für Gesundheit und Soziale Sicherung (BMGS). Online im WWW: <http://www.univation.org/index.php?class=seite&id=9031> (22.12.05)

BORTZ, J (1984): Lehrbuch der empirischen Forschung für Sozialwissenschaftler. Berlin u.a.

CAMPBELL, D.T. (1983): Reforms as Experiments. In: STUENING, E.L./ BREWER, M.L. (Hrsg.): Handbook of Evaluation Research. Beverly Hills, 107-137.

ENGESTRÖM, Y. (1999): Lernen durch Expansion. Marburg.

EULER, D. (2005): Qualitätsentwicklung in der Berufsausbildung. Materialien zur Bildungsplanung und zur Forschungsförderung. Heft 127. Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung. Bonn. Online im WWW: <http://www.blk-bonn.de/papers/heft127.pdf> (22.12.05)

GÜLPEN, B. (1996): Evaluation betrieblicher Verhaltenstrainings unter besonderer Berücksichtigung des Nutzens. München und Mering.

HOLZKAMP, K. (1996): Wider den Lehr-Lern-Kurzschluß. Interview zum Thema Lernen. In: FAULSTICH, P. / Ludwig, J. (Hrsg., 2004): Expansives Lernen. Baltmannsweiler, 29-38.

KIRKPATRICK, D. L. (1998): Evaluating Training Programs. The Four Levels. San Francisco.

KIRKPATRICK, J. D./ KIRKPATRICK, D. L. (2005): Transferring Learning to Behavior. Using the Four Levels to Improve Performance. San Francisco.

KUHLENKAMP, D. (2005): Finanzielle Ressourcen zur Teilhabe an Weiterbildung. Vortrag auf der Sektionstagung Erwachsenenbildung in Potsdam am 22.09.05.

MARTENS, J.-U. (1987). Empirische Erprobung „objektiver Lehrsysteme“. In: WILL, H./ WINTELER, A./ KRAPP, A. (Hrsg.): Evaluation in der beruflichen Aus- und Weiterbildung. Heidelberg, 89-107.

PHILLIPS, J./ SCHIRMER (2005): Return on Investment in der Personalentwicklung. Der 5-Stufen-Evaluationsprozess. Berlin, Heidelberg.

PHILLIPS, J.J. (1994): In Action: Measuring Return on Investment - Volume 1. Alexandria.

REICHERTS, H.-J. & KRUMKAMP, J. (1987): Praktische Ansätze betrieblicher Weiterbildungskontrolle. In: BLOMBERG, P.v. et al. (Hrsg.): Weiterbildung im Wandel. Konzeptionelle und methodische Innovationen. Hamburg, 263-281.

SENGE, P. (1996): Die fünfte Disziplin. Kunst und Praxis der lernenden Organisation. 3. Aufl., Stuttgart.

STUFFLEBEAM, D.L. (1966): A depth study of the evaluation requirement. Theory Into Practice, 5(3), 121-133.

STUFFLEBEAM, D.L. (2002): The CIPP Model for Evaluation. In: STUFFLEBEAM, D.L./ MADAUS, G.F./ KELLAGHAN, T.: Evaluation Models. Viewpoints on Educational and Human Services Evaluation. Second Edition. eBook. New York u.a., 290-317.

WITTMANN, W. (1985): Evaluationsforschung. Aufgaben, Probleme und Anwendungen. Berlin, Heidelberg, New York u.a.