

Evaluation betrieblicher Weiterbildung aus Sicht der Verantwortlichen



Selina Ebert, Caroline Bonnes und Sabine Hochholdinger

Zusammenfassung: Kontinuierliches Lernen von Beschäftigten wird immer wichtiger, und die Ausgaben der Unternehmen für Weiterbildung steigen stetig an. Daher ist es nötig, die Wirksamkeit von Weiterbildung systematisch zu evaluieren. Auf Basis einer explorativen Interviewstudie ($N = 18$) wird im vorliegenden Beitrag die aktuelle Praxis der Weiterbildungsevaluation in deutschsprachigen Unternehmen aus Sicht der Verantwortlichen betrachtet. Die Ergebnisse der Interviews verdeutlichen, dass zwar in allen Unternehmen betriebliche Weiterbildungen evaluiert werden, dieser Prozess aber nur bei einer Minderheit der Unternehmen modellgestützt erfolgt und methodische Qualitätsstandards nur teilweise umgesetzt werden. Auch die Nutzung und die Weiterverarbeitung der Evaluationsergebnisse ist über alle Unternehmen der Studie hinweg ausbaufähig.

Stichworte: Evaluation, Controlling, Weiterbildung, Experteninterviews, 4-Ebenen-Modell, Unternehmen

Evaluation of company further training from the perspective of the responsible persons

Abstract: In view of the increasing importance of continuous learning for employees and the rising expenditure of companies on training, it is necessary to systematically evaluate the effectiveness of further training. On the basis of an explorative interview study ($N = 18$), this article examines the recent practice of further training evaluation in German-speaking companies from the perspective of those responsible for the evaluation process. The results of the interviews show that all companies evaluate further training programmes, but only a minority of them do so in a model-based and scientifically oriented way. The use and further processing of the data generated by the evaluation can be improved across all companies in the study.

Keywords: evaluation, controlling, training, expert interviews, four levels of training evaluation, company

1 Ausgangssituation und Problemstellung

Damit Unternehmen in einer sich permanent verändernden Welt erfolgreich bestehen können, müssen ihre Beschäftigten kontinuierlich dazulernen (Schäfer & Ebersbach, 2021). Effektive und hochwertige Personalentwicklungsmassnahmen beeinflussen die

Wettbewerbsfähigkeit von Unternehmen und deren Beschäftigten massgeblich (Aguinis & Kraiger, 2009). Die betrieblichen Ausgaben für Weiterbildung steigen kontinuierlich (Seyda & Placke, 2020). Umso wichtiger ist es, dass Unternehmen die Wirksamkeit und den Nutzen ihrer betrieblichen Weiterbildungsangebote kennen, um diese strategisch zu steuern sowie Entscheidungen zu treffen. Dem dient ein systematisches Bildungscontrolling, insbesondere durch die Evaluation von Weiterbildungsangeboten.

Der Fokus der Studie liegt auf der gelebten Praxis bei der Evaluation betrieblicher Weiterbildung, welche einen wesentlichen Teil des Bildungscontrollings darstellt (Hoffmann, 2018). Ein etabliertes wissenschaftliches Konzept zur Weiterbildungs- oder Trainingsevaluation ist das Vier-Ebenen-Modell von Kirkpatrick (Kirkpatrick & Kirkpatrick, 2016). Bisherige Studien konzentrieren sich auf Teile des Vier-Ebenen-Modells oder des Evaluationsprozesses (Kennedy et al., 2014; Bächmann et al., 2019), legen den Fokus auf einzelne Weiterbildungsmassnahmen (Williams & Nafukho, 2015; Yates, 2015) oder Branchen (Kucherov & Manokhina, 2017), berücksichtigen jedoch nicht die Qualität des gesamten Evaluationsprozesses und die Konsequenzen für evaluierte Weiterbildungen. Auch erleichtert es inzwischen die Digitalisierung, hierfür aussagekräftige Daten mit vergleichsweise geringem Aufwand zu erheben und aufzubereiten.

Dieser Beitrag geht deshalb der Frage nach, in welcher methodischen Qualität Weiterbildung konkret evaluiert wird und wie die Evaluationsergebnisse in Unternehmen genutzt werden. Das Ziel der explorativen Interviewstudie besteht darin, vertiefte Einblicke in die Praxis betrieblicher Weiterbildungsevaluation zu erhalten und die daraus folgenden Implikationen zu diskutieren. Damit soll dazu beigetragen werden, dass Weiterbildungsmassnahmen effektiv und effizient sind und dadurch für Einzelpersonen, Teams, Unternehmen und Gesellschaft einen Nutzen stiften (Aguinis & Kraiger, 2009).

2 Stand der Forschung

2.1 Betriebliche Weiterbildung

Die vorliegende Arbeit bezieht sich auf die betriebliche Weiterbildung von Beschäftigten, welche in der Regel vom Arbeitgebenden finanziert wird und überwiegend während der Arbeitszeit stattfindet. Durch Weiterbildungen sollen Beschäftigte ihre Arbeitsleistung bei bisherigen oder neuen Aufgaben verbessern bzw. in die Lage versetzt werden, definierte Arbeitsaufgaben auszuführen. Auch erwarten Beschäftigte selbst, dass sie durch Weiterbildung bei ihrer beruflichen Weiterentwicklung unterstützt werden. Durch Weiterbildung soll sich die Innovations- und Wettbewerbsfähigkeit von Unternehmen erhöhen. Weiterbildungsangebote unterscheiden sich nach Inhalten (z.B. Trainings zu Sozialkompetenz, Sprache oder technischen Fähigkeiten), Methodik (z.B. online, Präsenz) oder zeitlichem Umfang (von einer Stunde bis mehrere Monate) (Bundesministerium für Bildung und Forschung, 2024). In der wissenschaftlichen Literatur werden die grosse Fülle an Weiterbildungsmassnahmen sowie deren Unterschiede bisher jedoch kaum beleuchtet.

Die praktische Relevanz betrieblicher Weiterbildungsangebote spiegelt sich beispielsweise in der Weiterbildungsumfrage des Instituts der deutschen Wirtschaft wider, wonach 87,9 % aller Unternehmen sich aktiv an der Weiterbildung ihrer Belegschaft beteiligen. Beschäftigte nahmen 2019 durchschnittlich 18,3 Stunden an Weiterbildungen teil und damit eine Stunde mehr als 2016. Die geschätzten betrieblichen Ausgaben für Weiterbildung

lagen 2019 in Deutschland bei 41,3 Milliarden Euro und haben sich damit im Vergleich zum Jahr 2016 um 23 % erhöht (Seyda & Placke, 2020).

Forschungsbefunde unterstreichen die Wirksamkeit von Weiterbildung, welche individuellen, betrieblichen und gesellschaftlichen Erfolg fördern kann. Dabei zeigen sich mittlere Effektstärken, allerdings auch beträchtliche Varianz zwischen den Wirksamkeitsstudien (Twitchell et al., 2000; Aguinis & Kraiger, 2009; Kauffeld, 2016).

Um sicherzustellen, dass betriebliche Weiterbildung effektiv und effizient gestaltet wird, ist daher die Evaluation dieser Weiterbildungsangebote unerlässlich.

2.2 Evaluation von Weiterbildung

Die Evaluation von Weiterbildung ist ein systematischer Prozess, bei dem Daten erhoben werden, um den Erfolg von Weiterbildung hinsichtlich festgelegter Kriterien zu bewerten. Die Bewertungskriterien leiten sich aus den Zielen, Formen und Inhalten der Weiterbildung ab (Saks & Burke, 2012; Kauffeld, 2016). Die verwendeten Erhebungsverfahren müssen abbilden, wie weit diese Ziele erreicht wurden, und das Untersuchungsdesign muss sicherstellen, dass die ermittelten Ergebnisse direkt auf die Weiterbildung und nicht auf andere Faktoren zurückzuführen sind (Peters et al., 2011; Rohmann et al., 2017).

Als Rahmenmodell für die Wirksamkeit von Weiterbildung hat sich in der Forschung das gängige Vier-Ebenen-Modell der Trainingsevaluation von Kirkpatrick (Kauffeld, 2016; Kirkpatrick & Kirkpatrick, 2016) etabliert. Es unterscheidet vier Ebenen, anhand derer sich der Erfolg von Weiterbildung zeigt und erfassen lässt. Die erste Ebene beinhaltet die emotionalen und motivationalen *Reaktionen* der Teilnehmenden auf die Weiterbildung. Die zweite Ebene, das *Lernen*, zeigt, wie gut sich die Teilnehmenden Weiterbildungsinhalte angeeignet haben in Form von Wissen, Fähigkeiten oder Einstellungen. Die dritte Ebene zielt auf das *Verhalten* ab und damit, wie umfassend die Weiterbildungsinhalte am Arbeitsplatz angewendet werden, also auf den Transfer. Die vierte und letzte Ebene der *Resultate* betrachtet die Auswirkungen der Trainingsmassnahmen auf betrieblich relevante Erfolgsgrößen. Es wird angenommen, dass die Ebenen hierarchisch aufeinander aufbauen, also z.B. eine positive Evaluation einer unteren Ebene Voraussetzung für die positive Evaluation der nachfolgenden Ebenen ist. Um die Effekte aller Ebenen möglichst direkt auf die Weiterbildung zurückzuführen, sollten Masse für die vier Ebenen vor und nach der Weiterbildung erhoben und die Veränderungen mit einer Kontrollgruppe ohne Weiterbildung verglichen werden.

Dieses Vier-Ebenen-Modell wurde in den letzten Jahrzehnten umfassend geprüft und diskutiert. Kritik richtete sich vor allem darauf, dass die hierarchische Beziehung der Ebenen empirisch nicht vollständig belegt werden konnte und auf die Trennschärfe der Ebenen. Weiterhin ergänzte beispielsweise Phillips (2003) den Return on Investment (ROI) als fünfte Ebene, welche den monetären Nutzen und die Kosten der Weiterbildung ins Verhältnis setzt. Trotz dieser Kritikpunkte erwies sich das Modell in Forschung und Praxis als sehr praktikable und robuste Heuristik für eine aussagekräftige Weiterbildungsevaluation und liegt bis heute den meisten empirischen Studien und Metaanalysen zur Wirksamkeit von Weiterbildung zugrunde (Giangreco et al., 2010, Hauser et al., 2020).

2.3 Betriebliche Evaluationspraxis

Wie häufig und auf welchen Ebenen Unternehmen im Sinne des Vier-Ebenen-Modells von Kirkpatrick betriebliche Weiterbildung evaluieren, wurde bereits in internationalen Studien betrachtet (z.B. Yadapadithaya & Stewart, 2003; Huerta et al., 2006; Becker et al., 2008). Dabei nimmt die Evaluationshäufigkeit von der ersten bis zur vierten Ebene deutlich ab. Insbesondere die Verhaltens- und die Resultatsebene werden am seltensten evaluiert. Dies steht im Kontrast dazu, dass die verhaltens- und resultatsbezogenen Ebenen 3 und 4 die grösste betriebliche Relevanz haben (Kennedy et al., 2014). Auch korreliert die Häufigkeit der Evaluation der letzten beiden Ebenen positiv mit dem Transfer des Gelernten (Saks & Burke, 2012).

In Verbindung damit wurde in diesen und weiteren Studien eruiert, welche Faktoren mit dem Prozess der betrieblichen Weiterbildungsevaluation zusammenhängen, insbesondere der Ebenen 3 und 4. Förderlich für die Durchführung einer Evaluation scheint es, wenn beispielsweise die zu evaluierende Weiterbildung grosse betriebliche Bedeutung hat (Bächmann et al., 2019) und teuer ist (Liechti & Abraham, 2011), wenn das Management des Unternehmens die Evaluation unterstützt (Kennedy et al., 2014) oder wenn alternative Weiterbildungsanbietende auf dem Markt verfügbar sind (Liechti & Abraham, 2011). Dagegen wird angenommen, dass Organisationen seltener Weiterbildungen evaluieren, wenn die Evaluation mit hohen Kosten verbunden ist, Trainingseffekte schwierig zu isolieren und Evaluationskriterien komplex zu messen sind (Kucherov & Manokhina, 2017). Ebenso ist es möglicherweise nachteilhaft, wenn das Management die Evaluationen nicht einfordert, personelle und zeitliche Ressourcen knapp sind oder die Zuständigen wenig Erfahrung und Fachwissen haben (Kennedy et al., 2014, Williams & Nafukho, 2015).

Zugleich sind die bisherigen Studien jedoch sehr allgemein gehalten. Es bleibt unklar, in welcher methodischen Qualität die Evaluationen durchgeführt werden. Nur wenn die Indikatoren für den Erfolg einer Weiterbildung aussagekräftig erfasst worden sind und wenn aus den genutzten Untersuchungsdesigns kausale Schlüsse über die Wirksamkeit einer Weiterbildung gezogen werden können, lassen sich die Evaluationsergebnisse sinnvoll interpretieren und nutzen.

Auch bleibt in den meisten Studien unbeantwortet, inwieweit die Ergebnisse in der Praxis systematisch und strategisch genutzt werden, um Konsequenzen für das betriebliche Weiterbildungsangebot zu ziehen, z.B. ob sie fortgesetzt oder verbessert werden sollen. Bober und Bartlett (2004) konnten zeigen, dass nur wenige der befragten Verantwortlichen verstehen, dass es wichtig ist, die erhaltenen Evaluationsergebnisse auch zu nutzen.

Inzwischen regeln Normen (ISO 9001) oder betriebliche Vereinbarungen, dass Unternehmen Weiterbildungsmassnahmen evaluieren sollen, um Qualitätsauflagen zu erfüllen, wobei dort allerdings methodische Standards kaum konkretisiert werden. Auch ermöglicht die zunehmende Digitalisierung, Evaluationsdaten schnell und komfortabel zu erheben und aufzubereiten, etwa in Verbindung mit HR-Analytics oder in digitalisierten Lernumgebungen. Vor diesem Hintergrund wäre zu erwarten, dass inzwischen häufiger und umfassender evaluiert wird als in der Vergangenheit.

Angesichts der steigenden betrieblichen Investitionen für Weiterbildung und der praktischen Relevanz für die Wettbewerbsfähigkeit von Unternehmen ist es wichtig, mehr über das konkrete Vorgehen von Unternehmen unter wissenschaftlichen Gesichtspunkten zu erfahren.

Stärker als in den bisherigen Studien steht also im Mittelpunkt, in welcher methodischen Qualität und unter welchen strategischen Gesichtspunkten Unternehmen ihre Weiterbildung evaluieren. Deshalb wurde ein explorativer Ansatz mittels eines qualitativen Erhebungs- und Analyseverfahrens gewählt, das hierzu aussagekräftigere Informationen liefert als eine standardisierte Fragebogenerhebung.

Im vorliegenden Beitrag sollen die folgenden Fragen beantwortet werden.

- F1: *Wie wird das Vier-Ebenen-Modell von Kirkpatrick in den Unternehmen umgesetzt?*
- F2: *Welche Erhebungsmethoden werden für die Weiterbildungsevaluation genutzt?*
- F3: *Was macht für die Zuständigen die Qualität einer Evaluation aus?*
- F4: *Welche Auswirkungen haben die Evaluationsergebnisse im Unternehmen?*
- F5: *Welche Faktoren beeinflussen aus Sicht der Verantwortlichen die betriebliche Weiterbildungsevaluation, und wie gehen die Unternehmen mit diesen um?*

3 Methode

3.1 Studienteilnehmende

Von Dezember 2022 bis April 2023 wurden 18 Personen interviewt, die für die betriebliche Weiterbildungsevaluation in ihrem Unternehmen zuständig sind. Sie stammen aus Unternehmen in Deutschland (16), Österreich (1) und der Schweiz (1). Die Unternehmen bestanden durchschnittlich seit 89 Jahren (Range 24–188), hatten zwischen 125 und 180.000 Mitarbeitende, einen Umsatz von 18 Millionen bis 40 Milliarden, und 15 der 18 Unternehmen befanden sich im Familienbesitz. In 16 der 18 Unternehmen verpflichteten betriebliche Vereinbarungen zur Evaluation von Weiterbildung. Die Befragten waren zwischen 26 und 60 Jahren alt und hatten durchschnittlich 12 Jahre Erfahrung (Range 6–22) mit der Weiterbildungsevaluation. Sie waren Mitglied der Personalabteilung, der Personalentwicklung oder von Kompetenzzentren, die für die Weiterbildung von speziellen Fachgebieten eingerichtet sind. Zehn der 18 Interviewten waren Frauen, 78 % der Befragten hatten das Abitur und einen akademischen Abschluss.

3.2 Durchführung der Interviews

Mit den Befragten wurden halbstandardisierte Leitfadeninterviews durchgeführt. Im Mittelpunkt standen vier Themenblöcke, welche durch insgesamt 24 Unterfragen spezifiziert wurden:

- Ablauf der betrieblichen Weiterbildungsevaluation
- Ebenen und Erhebungsverfahren der Weiterbildungsevaluation
- Eigenschaften einer guten Evaluation, Verbesserungspotenzial im eigenen Unternehmen
- Hinderliche und förderliche Faktoren für die Weiterbildungsevaluation

Die Interviews wurden per Telefon oder Videokonferenz durchgeführt, aufgezeichnet und transkribiert. Sie dauerten zwischen 20 und 55 Minuten. Die soziodemographischen Daten und Unternehmensdaten wurden über einen Fragebogen erhoben.

3.3 Datenanalyse

Die Interviewtranskripte wurden mit einer strukturierenden qualitativen Inhaltsanalyse nach Mayring (2022) ausgewertet. Damit lassen sich wesentliche Interviewaspekte

beschreiben und deren Häufigkeiten ermitteln. Das schrittweise entwickelte Kategoriensystem beinhaltet deduktiv gebildete Kategorien, welche sich aus vorliegenden Studien (Twitchell et al., 2000, Kennedy et al., 2014, Ho et al., 2016, Rohmann et al., 2017, Srimannarayana, 2017) und dem Vier-Ebenen-Modell von Kirkpatrick (Kirkpatrick & Kirkpatrick, 2016) ableiten. Ergänzend wurden aus dem Material induktive Kategorien entwickelt. Voraussetzung für die Bildung einer Kategorie war das Vorkommen dieser in mindestens drei Interviews. Dadurch sind 105 Kategorien entstanden, welche sich in 20 Ober-, 76 Unterkategorien sowie 9 Unter-Unterkategorien gliedern und sich in fünf Themenschwerpunkte unterteilen lassen: Verwendung des Vier-Ebenen-Modells, Erhebungsmethode, Qualität der Evaluation, Auswirkung der Evaluation sowie evaluationsbeeinflussende Faktoren.

Bei der Kodierung wurden alle Textstellen berücksichtigt, die relevant für die Fragestellungen waren. Dabei wurden für sich stehende, thematisch zusammenhängende Worte oder Sätze als Kodiereinheit festgelegt. Inhaltsgleiche Aussagen innerhalb eines Interviews wurden der jeweiligen Kategorie nur einmal zugewiesen. Insgesamt wurden 996 Textstellen mit MAXQDA kodiert.

Zur Bestimmung der Interraterreliabilität wurden mit dem Kategoriensystem ein Drittel des Materials durch zwei Personen kodiert. Hierbei wurden die relevanten Textstellen durch die hauptkodierende Person auf Oberkategorieebene markiert und durch die zweite Person den Unterkategorien zugeordnet. Der auf dieser Basis ermittelte Kappa-Koeffizient (Brennan & Prediger, 1981) beträgt 0,89. Diese hohe Übereinstimmung dokumentiert eine sehr gute Zuverlässigkeit des Kategoriensystems. Die Kategorien, Kategoriebeschreibungen und Kodierhäufigkeiten werden in den Tabellen 1 bis 5 detailliert dargestellt. Aus Darstellungsgründen wurden einzelne Nennungen unter den Kategorien „Sonstiges“ sprachlich zusammengefasst (z.B. Verbesserung von a, b, c etc.).

4 Ergebnisse

Nachfolgend werden die Ergebnisse der qualitativen Untersuchung entlang der fünf Forschungsfragen dargestellt.

4.1 Verwendung des Vier-Ebenen-Modells

Zunächst wurden die Interviewten befragt, wieweit das Vier-Ebenen-Modell der Trainingsevaluation nach Kirkpatrick bekannt ist und angewendet wird (vgl. Tabelle 1) (Kirkpatrick & Kirkpatrick, 2016). Während zwei Unternehmen das Vier-Ebenen-Modell explizit nutzen, geben 15 Befragte an, dass das Modell in ihren Unternehmen nicht zum Einsatz kommt. Der Hälften der Befragten war das Modell nicht bekannt. Fünf interviewte Personen hatten von dem Modell schon gehört, konnten es jedoch nicht wiedergeben. Vier der 18 Studienteilnehmenden konnten das Modell beschreiben.

Im Anschluss daran wurde im Interview das Vier-Ebenen-Modell erläutert und mit den interviewten Personen die verwendeten Instrumente den einzelnen Ebenen zugeordnet (vgl. Tabelle 1). Hier zeigte sich, dass 18 Unternehmen die Reaktionsebene erfassen. Dies geschieht weitgehend schriftlich durch die Selbsteinschätzung der Trainingsteilnehmenden, in den meisten Fällen über Funktionen im Lernmanagementsystem. Vereinzelt kommen Feedbackgespräche zum Einsatz. Ein Unternehmen gibt an, in der Vergangenheit die Reaktionsebene sehr ausführlich erhoben zu haben, inzwischen jedoch nur noch die Weiteremp-

fehlungsrate abzufragen, da die gewonnenen Erkenntnisse zu keinem Mehrwert geführt hätten.

Die Lernebene wird von 13 Unternehmen erhoben. Als Instrument werden mehrheitlich Wissenstests genannt. Diese werden beispielsweise durch Quizze über Online-Tools umgesetzt, durch schriftliche Prüfungsleistungen oder über Fragebogen und Gespräche. Einige Unternehmen erfassen Fertigkeiten durch Demonstration. So müssen Teilnehmende nach einem Schweißkurs die gelernten Schweißtechniken demonstrieren, oder bei einer Vertriebsschulung werden die erlernten Kenntnisse durch einen „Lockvogel“ getestet

Oberkategorie	Beschreibung	Unterkategorie + Häufigkeit
Bekanntheit 4-Ebenen-Modell*	Bekanntheitsgrad des 4-Ebenen-Modell nach Kirkpatrick bei der interviewten Person	<ul style="list-style-type: none"> ■ Modell ist bekannt (4) ■ Modell ist bekannt, kann jedoch nicht wiedergegeben werden (5) ■ Modell ist nicht bekannt (9)
Bewusste Anwendung 4-Ebenen-Modell*	Anwendung des 4-Ebenen-Modells nach Kirkpatrick	<ul style="list-style-type: none"> ■ ja (2) ■ nein (15) ■ zum Teil (1)
Ebene 1 – Reaktion*	Bewusste oder unbewusste Messung von Ebene 1	<ul style="list-style-type: none"> ■ ja (18) <ul style="list-style-type: none"> – Instrumente*** (14) – Fragebogen (11) – Sonstige: mündliche Befragung (2); Weiterempfehlungsrate
Ebene 2 – Lernen*	Bewusste oder unbewusste Messung von Ebene 2	<ul style="list-style-type: none"> ■ ja (13) <ul style="list-style-type: none"> – Instrumente*** (12): <ul style="list-style-type: none"> – Wissenstest/Quiz (6) – Fragebogen (4) – Sonstige: Anwendung der Kenntnisse (2); Transfergespräch; Präsentation einer Projektarbeit ■ nein (4) ■ fehlende Information (1)
Ebene 3 – Verhalten*	Bewusste oder unbewusste Messung von Ebene 3	<ul style="list-style-type: none"> ■ ja (17) <ul style="list-style-type: none"> – Instrumente*** (16): <ul style="list-style-type: none"> – Mitarbeitendengespräche (8) – Fragebogen (7) – Tests (3) – Sonstige: Zielvereinbarung (2); Reflexionsrunde zu Verhaltensveränderung; Feedback vom Vorgesetzten; Projektaufgabe; Arbeitsausführung ■ nein (1)
Ebene 4 – Ergebnisse*	Bewusste oder unbewusste Messung von Ebene 4	<ul style="list-style-type: none"> ■ ja (4) <ul style="list-style-type: none"> – Instrumente*** (3): Kennzahlen (2); Betrachtung der Ergebnisse in Reflexionsrunde ■ nein (14)

Anmerkung: * Zuordnung des Interviews nur zu einer Unterkategorie innerhalb der Oberkategorie

*** Unterkategorie, welche nur bei der Auswahl von „ja“ befüllt wird; Zuordnung des Interviews innerhalb dieser Unterkategorie zu mehreren Unter-Unterkategorien möglich

Tabelle 1: Kategorien und Kodierhäufigkeiten – Verwendung des Vier-Ebenen-Modells

(P 11). In einem Unternehmen fertigten die Teilnehmenden einer grösseren Schulungsreihe schriftliche Ausarbeitungen zum Schulungsthema an und präsentierten diese vor dem Vorstand und weiterem Publikum. Ein anderes Unternehmen vergibt beim Bestehen des Wissenstests Zertifikate. Punktuell berichten Unternehmen davon, die Lernebene nicht „flächendeckend“, sondern „sofern es sich lohnt“ (P 10) zu evaluieren. Kommunikationsthemen, Blended Learning Ansätze oder interne Ausbildungsreihen werden von diesen Unternehmen als Anwendungsbereiche hierfür identifiziert. Weder von den evaluierenden Unternehmen noch den nicht-evaluierenden Unternehmen dieser Ebene werden Gründe für die (Nicht-)Anwendung genannt.

Bis auf ein Unternehmen erfassen alle die Verhaltensebene. Dabei setzen viele auf eine Wirksamkeitsbeurteilung durch Vorgesetzte über einen Fragebogen. Bei einem Unternehmen muss die Führungskraft einen Nachweis erbringen und hat hierfür drei Möglichkeiten: Ermittlung der Wirksamkeit über die Arbeitsausführung, ein Gespräch oder einen Test. Zudem finden häufig Transfergespräche zwischen der teilnehmenden Person und ihrer Führungskraft statt. Einzelne Unternehmen empfehlen ihren Führungskräften bereits vor der Weiterbildung Zielvereinbarungen zu treffen, um „dann auch hinterher nochmal zu schauen: Konnte der Teilnehmer das jetzt auch in der Praxis umsetzen?“ (P 3). Vereinzelt werden die regelmässigen Mitarbeitendengespräche zur Feststellung des Transfers genutzt. Unklar ist, inwiefern Führungskräfte diese Art der Erfassung unterstützen. So berichtet eine Person aus ihrem Unternehmen: „Die Wirksamkeitskontrolle, die hassen die Führungskräfte“ (P 8). Die Verhaltensebene wird bei einem Unternehmen auch über gemeinsame Reflexionsrunden mit den Teilnehmenden gemessen. Vereinzelt berichten Unternehmen davon, die Verhaltensebene nicht bei allen ihren Weiterbildungsangeboten bzw. unterschiedlich stark zu messen. Beispielsweise werden Führungskräftetrainings oder Angebote für Potenzialtragende verstärkt evaluiert. Zudem werden Schwierigkeiten bei der Messung, z.B. aufgrund weicher Faktoren, oder die mangelnde Dokumentation der Verhaltensmessung angesprochen. In einem Unternehmen, das die Verhaltensebene nicht erfasst, gibt die verantwortliche Person an, dass die Rückmeldung auf der Verhaltensebene „immer wieder über das Feedback“, beispielweise über die Kunden oder das „Beschwerdemanagement“ (P 12) an sie zurückgemeldet wird. Weitere Gründe für die Nichtanwendung werden nicht genannt.

Die Ergebnisebene wird in vier Unternehmen berücksichtigt. Eine befragte Person aus dieser Gruppe gibt an, dass sie sich mit der Frage „Hat die Verhaltensänderung auch zu entsprechenden Ergebnissen geführt?“ (P 1) beschäftigt, erläutert dieses Vorgehen jedoch nicht genauer. Zwei der vier Personen heben das „rudimentäre“ (P 4) Vorgehen sowie die „viele[n] Einflussmöglichkeiten“ auf die Evaluation der Ergebnisebene hervor (P 11). Bezüglich der Anwendungshäufigkeit geben die Unternehmen an, in „reduzierter Form“ (P 4), in „Einzelfällen“ (P 14) oder in „Abschlussarbeiten“ (P 1) die vierte Ebene zu messen. Als Gründe für die geringe Anwendung werden der hohe Aufwand und die Voraussetzung messbarer Ziele genannt. Von den Unternehmen, welche die vierte Ebene nicht evaluieren, werden Schwierigkeiten bei der Messbarkeit des Zusammenhangs von Ursache und Wirkung sowie der Zuordnung von Umsatzzuwachs/Gewinnzuwachs zur jeweiligen Schulung berichtet. Zudem wird die hohe Komplexität der Messung auf dieser Ebene hervorgehoben.

4.2 Erhebungsmethoden

Tabelle 2 gibt eine Übersicht über die Erhebungsmethoden bei den Unternehmen der Studie. Bei allen Unternehmen findet die Evaluation nach der Weiterbildung statt, der Zeitraum reicht von unmittelbar nach der Weiterbildung bis zu sechs Monaten später. Einige Unternehmen ($n=4$) evaluieren die Weiterbildung bereits währenddessen, vor allem bei mehrtägigen oder auf mehrere Zeitpunkte aufgeteilten Schulungen. Manche Unternehmen führen eine Bedarfsanalyse vor der Evaluation durch, in der Ziele, Erwartungen und Unterstützungsbedarf abgefragt werden. Nur selten wird vorab der Wissensstand für eine anschliessende Prä-Post-Messung erhoben.

Methodisch setzen alle Unternehmen auf schriftliche Befragungen. Die Mehrheit der Unternehmen gab an, die Weiterbildungsevaluation digital gestützt durchzuführen, z.B. über Lernmanagementsysteme oder Onlineworkflows. Mündliche Befragungen und Beobachtungen werden von zwölf bzw. zwei der Unternehmen genutzt. Die Evaluation basiert in allen Unternehmen auf Selbsteinschätzungen der Teilnehmenden. 14 Unternehmen berichten zusätzlich von Fremdeinschätzungen überwiegend durch die jeweilige Führungskraft, vereinzelt durch Mitarbeitende oder die schulende Person. Für kein Unternehmen wurden Kontrollgruppen oder generell experimentelle Untersuchungsdesigns berichtet.

Oberkategorie	Beschreibung	Unterkategorie + Häufigkeit
Zeitpunkt**	Zeitpunkt, zu dem die Weiterbildungsmassnahme evaluiert wird	<ul style="list-style-type: none"> ■ danach (18) ■ zwischen den Massnahmen/während dessen (4) ■ davor (6)
Erhebungsinstrumente**	Erhebungsinstrumente, die zur Evaluation eingesetzt werden	<ul style="list-style-type: none"> ■ schriftlicher Fragebogen (18) ■ mündliche Befragungen (12) ■ Beobachtungen (2)
Systemunterstützung*	Die Evaluation wird durch die Unterstützung eines digitalen Systems durchgeführt	<ul style="list-style-type: none"> ■ ja (12) ■ nein (3) ■ fehlende Information (3)
Erhebungsquelle**	Quelle, welche zur Evaluation der Weiterbildungsmassnahme verwendet wird	<ul style="list-style-type: none"> ■ Selbsteinschätzung (18) ■ Fremdeinschätzung (14)

Anmerkung: *Zuordnung des Interviews nur zu einer Unterkategorie möglich

**Zuordnung des Interviews zu mehreren Unterkategorien möglich

Tabelle 2: Kategorien und Kodierhäufigkeiten – Erhebungsmethode

4.3 Qualität der Evaluation

Die Interviewten reflektierten, wie eine gute Evaluation aussieht. Hier wurden unterschiedliche Punkte genannt, wobei die Individualität des Prozesses von den Befragten mit am häufigsten angesprochen wurde. Die Evaluation soll auf die einzelnen Weiterbildungsangebote abgestimmt sein und den Teilnehmenden ein individuelles Feedback, beispielsweise durch individuelle Kommentare, offene Fragen und Freitextfelder im Fragebogen, ermöglichen. Zudem legen mehrere Interviewte Wert darauf, dass die Evaluationsdaten statistisch verwertbar und aussagekräftig sind. Weitere gewünschte Eigenschaften waren unter anderem die zeitnahe Evaluierung sowie die Messung der Wirksamkeit bzw. des Transfers der Massnahme.

Verbesserungspotenzial sehen die Studienteilnehmenden bei der Transferüberprüfung sowie bei dem Verständnis und dem Bewusstsein der Teilnehmenden gegenüber der Evaluierung. Auch die Digitalisierung und Automatisierung ihrer Evaluation, die nachhaltige Beschäftigung mit den Trainingsinhalten, eine Lernerfolgskontrolle sowie die Einführung von Kennzahlen wurden in diesem Kontext angesprochen.

Für die Konzeption der Weiterbildungsevaluation ziehen Unternehmen verschiedene Informationsquellen heran. Am häufigsten wurden von den Unternehmen die Erfahrung und das Vorgehen anderer Unternehmen, Verbände etc. genutzt, ebenso Informationen aus Büchern, Magazinen und Zeitungen sowie aus dem Internet. In Tabelle 3 sind die Kategorien und Kodierhäufigkeiten im Hinblick auf die Qualität der Evaluation zusammenfassend dargestellt.

Kategorie	Beschreibung	Unterkategorie + Häufigkeit
Eigenschaften eines guten Evaluationsprozesses**	Eigenschaften, über die ein guter Evaluationsprozess verfügen sollte	<ul style="list-style-type: none"> ▪ individuell (7) ▪ liefert verwertbare und aussagekräftige Daten (7) ▪ misst die Wirksamkeit/den Transfer (3) ▪ zeitnah (3) ▪ Sonstige: valide (2); gute Frageformulierung (2); klare Struktur (2); vergleichbar (2); anonym; vertraulich; bedienungsfreundlich; transparent; kontinuierlich; mehrdimensional; hohe Rücklaufquote; klare Kommunikation; standardisierte Skalen; gutes Zeitmanagement; viele Sichtweisen; Beobachtungen durch Vorgesetzte; Feedbackunterscheidung nach Zielgruppen; gezielte Auseinandersetzung mit der Massnahme; zieht Veränderungen/Anpassungen nach sich; findet vor dem Training statt; beinhaltet eine Bedarfsanalyse; spiegelt Leistungen in die Organisation zurück; bewertet Trainingsinhalte, Dozierende, Rahmenbedingungen trennscharf voneinander
Verbesserungspotenzial eigenes Unternehmens**	Punkte, deren Veränderung den Evaluationsprozess im Unternehmen verbessern würden	<ul style="list-style-type: none"> ▪ Transfer überprüfen (5) ▪ Verständnis und Bewusstsein der Mitarbeitenden gegenüber der Evaluation schaffen (3) ▪ fehlende Information (1) ▪ Sonstige: digitalisieren und automatisieren (2); Training nachhalten (2); Einführung von Lernerfolgskontrollen (2) und Kennzahlen (2); kein Verbesserungspotenzial (2); Einführung von Arbeitsproben und Learning Management System; Führungskräfte stärker einbinden; Verbesserung von Messbarkeit, Vergleichbarkeit, Rücklaufquote, Datenaussagekraft, Evaluationsvorbereitung sowie des Evaluations- und Personalentwicklungsprozess; Schnelllebigkeit reduzieren
Informationsquellen**	Quellen, auf deren Basis der Evaluationsprozess im Unternehmen erstellt wurde	<ul style="list-style-type: none"> ▪ Erfahrung und/oder Vorgehen anderer (7) ▪ Bücher, Magazine, Zeitungen (5) ▪ Internet (5) ▪ fehlende Information (5) ▪ Sonstige: berufliche Erfahrung der Beteiligten (2); Wissen aus Ausbildung der Beteiligten (2); Stand der Forschung/der Wissenschaft (2); Wissen der Beteiligten; Abschlussarbeiten; Schulungsunterlagen; Rückmeldungen zur Evaluation; Ursprungsevaluationsbogen; Unterlagen von internen AuditorInnen

Anmerkung: **Zuordnung des Interviews zu mehreren Unterkategorien möglich

Tabelle 3: Kategorien und Kodierhäufigkeiten – Qualität der Evaluation

4.4 Auswirkung der Evaluationsergebnisse

Laut einem Grossteil der Befragten werden die Evaluationsergebnisse intern im Team verwertet. In einigen Unternehmen werden sie den TrainerInnen, der Führungskraft der für die Evaluation verantwortlichen Abteilung, der Führungskraft der Teilnehmenden sowie dem Management berichtet. In einem Unternehmen werden die Ergebnisse niemandem bekannt gegeben.

Kategorie	Beschreibung	Häufigkeit
Berichterstattung**	Personengruppe(n), an welche die Ergebnisse der Evaluation berichtet werden	<ul style="list-style-type: none"> ▪ Verantwortliche Abteilung/Prozess/Person (12) ▪ TrainerIn (6) ▪ Führungskräfte der evaluationsverantwortlichen Abteilung (5) ▪ Führungskräfte der Teilnehmenden (3) ▪ Sonstige: Geschäftsführung/Management (2); Teilnehmende; AuditorInnen; Content Lieferanten; Trainingsneed anmeldende Abteilung; budgetverantwortliche Person, Ergebnisse werden nicht berichtet
Auswirkungen**	Folgen, die durch die Ergebnisse der Evaluation entstehen	<ul style="list-style-type: none"> ▪ Zusammenarbeit/Einsatz von Schulungsanbietenden und/oder TrainerInnen (7) ▪ Überarbeitung Schulung (6) ▪ Anpassung Finanzierung/Budget (4) ▪ Fortbestand Schulung (3) ▪ nochmalige Vermittlung der Schulungsinhalte (3) ▪ fehlende Information (1) ▪ Sonstige: Ergebnisse haben keine Auswirkungen (2); erneutes Nachfragen bei den Teilnehmenden; Vertragsverhandlungen mit TrainerInnen; Überarbeitung der Evaluation; Auswirkungen auf Mitarbeitendenbindung; „Donnerwetter“ an externen Schulungsanbietende

Anmerkung: **Zuordnung des Interviews zu mehreren Unterkategorien möglich

Tabelle 4: Kategorien und Kodierhäufigkeiten – Verwendung und Auswirkung der Evaluationsergebnisse

Die Evaluationsergebnisse beeinflussen beispielweise die Zusammenarbeit mit Schulungsanbietenden und TrainerInnen, die Überarbeitung der Schulung, den Fortbestand der Schulung oder eine erneute Vermittlung der Schulungsinhalte (vgl. Tabelle 4). Zudem werden von Auswirkungen auf die Finanzierung der Schulungen und das Schulungsbudget berichtet, da „man [zum Beispiel] jetzt festgestellt hat, dass Massnahmen eben sehr gut angenommen worden sind, die man eventuell nochmal intensivieren sollte“ (P 5). Eine Person gab an, dass die Evaluationsergebnisse bei ihnen im Unternehmen keine Auswirkungen haben.

4.5 Evaluationsbeeinflussende Faktoren

Für Unternehmen im Allgemeinen wurden hier zeitliche, finanzielle, informationstechnologische und personelle Ressourcen am häufigsten genannt. Auch die Wichtigkeit, Wertigkeit und Bedeutung der Personalentwicklung, die Unternehmensgrösse, die Akzeptanz und Wahrnehmung der Evaluation durch die beteiligten Personen sowie die Kompetenz und das Wissen der verantwortlichen Stelle spielen nach Einschätzung der befragten Personen eine Rolle. Zudem werden in den Interviews die Auswirkungen einzelner auf andere evaluationsbeeinflussende Faktoren hervorgehoben. Beispielsweise werden die vorhandenen

Ressourcen mit der Überzeugung, dass die Evaluation „einen Mehrwert stiftet“ (P 1) oder die Wichtigkeit der Themen Personalentwicklung, Weiterbildung und Evaluation im Unternehmen in Verbindung gebracht. Auch wird angenommen: „Ein Unternehmen was kleiner ist... hat nicht die Ressourcen, vielleicht, um hinter jedem Training dann eine Evaluation zu setzen“ (P 9).

Auffällig ist, dass die interviewten Personen, bezogen auf ihr eigenes Unternehmen, andere evaluationsbeeinflussende Faktoren nennen als sie zuvor allgemein berichteten, bzw. sich die Häufigkeit der bei beiden genannten Faktoren deutlich unterscheidet. Bezogen auf das eigene Unternehmen spielen vor allem die Akzeptanz, die Wahrnehmung sowie das Verständnis der beteiligten Personen gegenüber der Evaluation, die Digitalisierung sowie die Struktur- und Prozessgestaltung der Evaluation eine grosse Rolle. Übereinstimmend mit den allgemeinen evaluationsbeeinflussenden Faktoren wurden auch im eigenen Unternehmen die vorhandenen Ressourcen am häufigsten genannt (vgl. Tabelle 5). In den Interviews werden vereinzelt Massnahmen genannt, wie den qualitätsbeeinflussenden Faktoren im Unternehmen begegnet wird bzw. begegnet werden könnte. Ein Unternehmen berichtet davon, die Akzeptanz seiner Führungskräfte zu erhöhen, indem es sich auf zwei Bewertungsfragen mit einem Freitext beschränkt. Eine weitere interviewte Person nimmt an, dass sich die Akzeptanz und Wahrnehmung der Mitarbeitenden gegenüber der Evaluation positiv verändern lässt, wenn die Evaluationsergebnisse und abgeleitete Konsequenzen sowie Lösungen kommuniziert werden würden. Weitere Unternehmen berichten davon, dass sich ausreichend Zeit und personelle Kapazitäten, die vorhandene Infrastruktur durch das Learning Management System sowie der hohe Stellenwert, den Lernen, Weiterbildung und Personalentwicklung im Unternehmen einnehmen, positiv auf die Evaluation auswirken. Die Erfahrung, dass sich die Einführung eines Learning Management Systems auch negativ auf die Rücklaufquoten der Fragebogen und damit auf die Evaluation auswirken kann, bringt eine andere Person ein. Die Integration der Evaluation bereits in die Schulungsmassnahme ist für ein weiteres Unternehmen eine Lösungsmöglichkeit.

Die evaluationsbeeinflussenden Faktoren Einstellung und Stellenwert sowie die Erfahrung der Unternehmen im Bereich der Evaluation wurden separat tiefergehend erfragt. Bezüglich des Stellenwerts gaben vier der interviewten Personen an, dass die Bedeutung bzw. Wichtigkeit der betrieblichen Weiterbildungsevaluation in ihrem Unternehmen als hoch angesehen werden kann. Von einem mittleren Stellenwert berichten 14 Interviewte. Im Gegensatz dazu wird die Einstellung der Belegschaft gegenüber der Evaluation nur in einem Unternehmen als positiv sowie in allen anderen Unternehmen als neutral beschrieben.

Die häufigsten Gründe für die Evaluation waren die Wirksamkeits- bzw. Transferüberprüfung der Weiterbildung, die Notwendigkeit für die Auditierung bzw. Zertifizierung des Unternehmens, die Sicherstellung der Qualität sowie die Feststellung des Optimierungsbedarfs. Weitere einzeln genannte evaluationsbeeinflussende Faktoren können Tabelle 5 entnommen werden.

Themenbeiträge

Kategorie	Beschreibung	Häufigkeit
Evaluationsbeeinflussende Faktoren – Unternehmen allgemein**	Allgemeine Faktoren, welche die Qualität der Evaluation positiv oder negativ beeinflussen können	<ul style="list-style-type: none"> ■ Ressourcen (12) ■ Wichtigkeit/Wertigkeit/Bedeutung der Personalentwicklung (10) ■ Unternehmensgrösse (5) ■ Akzeptanz/Wahrnehmung der beteiligten Personen (4) ■ Kompetenz/Wissen der verantwortlichen Stelle (3) ■ fehlende Information (1) ■ Sonstige: Weiterbildungs- und Lernkultur des Unternehmens (2); Zusammenarbeit mit dem Betriebsrat; Thema und Dauer der Schulung; Unternehmenskultur; Umfang der Massnahme; Automatisierungsgrad und Bedienungsfreundlichkeit der Evaluation; Einfluss auf Gestaltung der Massnahme; gegebene Infrastruktur; sprachliches Verständnis; Betriebsklima; Kollegialität untereinander; Vertrauen; flache Hierarchie
Evaluationsbeeinflussende Faktoren – Eigenes Unternehmen**	Faktoren, welche die Qualität der Evaluation positiv oder negativ im Unternehmen der befragten Personen beeinflussen	<ul style="list-style-type: none"> ■ Ressourcen (9) ■ Akzeptanz/Wahrnehmung/Verständnis beteiligter Personen (7) ■ Digitalisierung (6) ■ Struktur- und Prozessgestaltung (6) ■ Sonstige: Bedeutung/Stellenwert des Themas (2); Selbstreflexionsfähigkeit; Rücklaufquote; Thema und Dauer der Schulung; Unternehmenskultur; Folgen/Nutzen der Evaluation; Datenschutz; Feedback- und Lernkultur; innere und äussere Einflüsse; Kenntnisse der Fachabteilung; Knüpfung an Balanced Scorecard; Vorleben durch Vorgesetzte; flache Hierarchien; Grösse des Unternehmens; Unternehmensklima; negative Konsequenzen, Verpflichtung; Betriebsrat; Standardisierung über alle Länder hinweg
Stellenwert*	Bedeutung/Wichtigkeit der Evaluation von betrieblichen Weiterbildungsmassnahmen im Unternehmen	<ul style="list-style-type: none"> ■ mittel (14) ■ hoch (4)
Einstellung*	Einstellung der Belegschaft im Unternehmen gegenüber der Evaluation	<ul style="list-style-type: none"> ■ neutral (17) ■ positiv (1)
Gründe für Evaluation**	Gründe, weshalb betriebliche Weiterbildungsmassnahmen evaluiert werden	<ul style="list-style-type: none"> ■ Wirksamkeitsüberprüfung/Transferüberprüfung (6) ■ Audit/Zertifizierung (6) ■ Qualitätssicherstellung (6) ■ Optimierungspotenzial/Optimierungsbedarf (3) ■ Sonstige: gute/richtige TrainerInnen (2) und Lernlücken (2) identifizieren; Bedarfspassung (2); Qualität der AnbieterInnen feststellen und auf dieser Basis Entscheidungen treffen (2); Weiterempfehlung der Schulung (2); Zielgruppenorientierung (2); Dokumentation/Wissensdatenbank (2); Trainingsinhalte festlegen; Veränderungspotenzial, schlechte/richtige Massnahmen sowie gewinnbringende Schulungen identifizieren; Feedback zur Massnahme erhalten; Mitarbeiterzufriedenheit erhöhen; kosteneffiziente Weiterbildungsmassnahmen anbieten; Lern- und Entwicklungszuwachs; Referentenweiterentwicklung; Mehrwert für Mitarbeitende und Organisation schaffen; Reaktion auf innere und äussere Einflüsse; Relation Aufwand und Nutzen prüfen; Austauschmöglichkeiten und Vergleichbarkeit schaffen

Anmerkung: * Zuordnung des Interviews nur zu einer Unterkategorie möglich

** Zuordnung des Interviews zu mehreren Unterkategorien möglich

Tabelle 5: Kategorien und Kodierhäufigkeiten – Evaluationsbeeinflussende Faktoren

5 Diskussion

Die vorliegende Studie gibt Einblicke in die gelebte Praxis der Weiterbildungsevaluation in deutschsprachigen Unternehmen aus Sicht der Verantwortlichen. Im Folgenden werden die Ergebnisse in Bezug auf die einzelnen Forschungsfragen diskutiert.

Mit Blick auf die *Verwendung des Vier-Ebenen-Modells nach Kirkpatrick (F1)* wird deutlich, dass sich ein Grossteil der Unternehmen nicht explizit daran orientiert. Die Hälfte der befragten Personen kannte das in der Forschung seit langem etablierte Modell nicht. Dies spricht dafür, dass sich der Transfer zwischen Theorie und Praxis verbessern könnte. Erfreulich ist, dass die Unternehmen de Facto dennoch die Ebenen des Modells erheben, was sich bei der Zuordnung der genannten Instrumente zu den vier Ebenen zeigt. Im Vergleich zu früheren Studien (z.B. Becker et al., 2008) erfassten in dieser Studie deutlich mehr Unternehmen die Verhaltensebene. Ein Grund hierfür könnten Regularien und Normen (ISO 9001) sein, welche die Erhebung der Wirksamkeit einer Weiterbildungsmassnahme regeln. Dies könnte jedoch auch auf einem Selektionseffekt beruhen, da Personen aus Unternehmen, die in der Weiterbildungsevaluation versierter sind, möglicherweise eher bereit sind, an einem Interview über Weiterbildungsevaluation mitzuwirken als Zuständige aus Unternehmen, in denen die Evaluation eine weniger wichtige Rolle spielt. Zudem wurde durch die Interviews deutlich, dass nicht jede Ebene bei jeder Massnahme gemessen wird. In weiteren Studien sollte geprüft werden, bei welcher Art von Massnahmen und wie häufig die jeweiligen Ebenen erhoben werden. Ferner sind die Gründe, welche für oder gegen die Messung der Ebenen sprechen, zu untersuchen. Im Hinblick auf die von Phillips (2003) hinzugefügte Ebene „ROI“ ist interessant, dass der ROI von keinem der befragten Unternehmen erhoben wird. Lediglich in einem Interview wird der ROI überhaupt angeprochen: „Wir haben keinen Return-of-Invest [sic] oder Ähnliches, der berechnet wird durch die Zahlen, weil es einfach sehr komplex ist, das auszurechnen“ (P 9). Zukünftige Forschung sollte an dieser Stelle ansetzen und die Praktikabilität dieser Ebene untersuchen bzw. konkrete Umsetzungsvorschläge als Hilfestellung formulieren.

Bezüglich der *verwendeten Methoden (F2)*, setzt die Mehrheit der Unternehmen bei den Instrumenten auf schriftliche oder mündliche Befragungen. Hierbei werden vielfältige Erhebungsverfahren genutzt. Aufwändigere Verhaltensbeobachtungen, welche im Gegensatz zu Befragungen weniger stark subjektiven Verzerrungen unterliegen (Döring, 2023), werden von nur drei Unternehmen angewendet. Aus den Interviewdaten geht jedoch nicht hervor, wie häufig und regelmässig die Methoden im Unternehmen verwendet werden, also welche Weiterbildungsangebote z.B. durch anspruchsvollere Evaluationen gemessen werden.

Bemerkenswert ist, dass viele Unternehmen Fremdeinschätzungen zusätzlich zu den Selbsteinschätzungen ($n=14$) verwenden. Diese werden häufig für die Messung der Verhaltensebene angewendet. Dabei ist offen, wie strukturiert die Fremdeinschätzung stattfindet und wie belastbar daher die Ergebnisse sind. Grundsätzlich ist es positiv, dass externe Urteile mit einbezogen werden, da sie geringeren Verzerrungen unterliegen als Selbsteinschätzungen (Rohmann et al., 2017).

Im Unterschied zu früheren Studien wurden in der vorliegenden Untersuchung digitale Tools in Lernmanagementsystemen ausdrücklich erwähnt. Hieran zeigt sich, wie die vorauslaufende Digitalisierung den Evaluationsprozess verändern und ggf. erleichtern kann.

Bezüglich des Zeitpunkts gibt die Hälfte der Befragten an, die Massnahmen nur im Nachhinein zu evaluieren. Auf diese Weise lassen sich jedoch keine Veränderungen feststellen.

len, womit unklar bleibt, ob die Wirksamkeit der Weiterbildung korrekt beurteilt werden kann (Peters et al., 2011). Dies gilt umso mehr, da nirgendwo experimentelle Designs eingesetzt werden. Diese wären zwingend nötig, um kausale Effekte einer Massnahme zu prüfen.

Bei den *Qualitätskriterien einer Evaluation (F3)* wurden bei den Eigenschaften einer guten Evaluation zum einen Aspekte der Standardisierung, zum anderen der Individualisierung und Flexibilisierung angesprochen. Dies zu vereinbaren ist anspruchsvoll. An dieser Stelle ist zu prüfen, welche technischen Möglichkeiten bestehen und wie diese im Hinblick auf den Evaluationsprozess bereits eingesetzt sind, zukünftig verwendet sowie weiterentwickelt werden können. Hierbei fällt auch auf, dass methodische Standards kaum thematisiert werden.

Viele der Informationsquellen, welche zur Konzeption der Evaluation genutzt werden, basieren auf nichtwissenschaftlichen Quellen, wie die Erfahrung Dritter oder dem Internet. Dies kann kritisch gesehen werden, da auf diese Weise der Nutzen der eingesetzten Evaluationsmethoden im Sinn einer Evidenzbasierung nicht gesichert ist.

Es zeigt sich, dass die Transferüberprüfung der Weiterbildungsmassnahmen von den Unternehmen am häufigsten als Ansatzpunkt zur Verbesserung ihres Evaluationsprozesses identifiziert wurde. Dies lässt mutmassen, dass die Unternehmen, trotz der geringen Orientierung ihres Evaluationsprozesses an wissenschaftlichen Standards und Befunden, an dieser Stelle den Fokus richtig setzen.

Mit Blick auf *Auswirkungen der Evaluationsergebnisse in den Unternehmen (F4)* wird deutlich, dass die Ergebnisse häufig nur abteilungsintern verwendet und nicht allen Stakeholdern zur Verfügung gestellt werden. Dieses Vorgehen sollte beispielsweise im Hinblick auf die Lehrenden angepasst werden, da diese die Trainings gestalten. In nur drei Unternehmen werden Schulungsinhalte erneut vermittelt, wenn die Evaluationsergebnisse nicht zufriedenstellend sind. Auch werden die Evaluationsergebnisse in einigen befragten Unternehmen gar nicht oder nur begrenzt genutzt, um Konsequenzen zu ziehen. Dort, wo Konsequenzen genannt werden, bleibt unklar, wie weit dem eine systematische Datenlage zugrunde liegt. Zusammenfassend lässt sich feststellen, dass die gewonnenen Daten deutlich weniger strategisch und systematisch genutzt werden, als dies möglich und sinnvoll wäre, um die Effektivität von Weiterbildung sicherzustellen. Dies könnte daran liegen, dass die Evaluation hauptsächlich aufgrund von formalen Vorgaben z.B. im Rahmen von Zertifizierungen vollzogen wird und der eigentliche Nutzen der Evaluation nebensächlich ist. Dafür spricht, dass als Grund für die Evaluation häufig die Auditierung bzw. Zertifizierung des Unternehmens genannt wird. Auch könnte sein, dass die mit der Evaluation betrauten Personen nicht alle erforderlichen Kenntnisse haben, um die Evaluation methodisch sauber durchzuführen sowie deren Ergebnisse zielgerichtet zu verarbeiten.

Gefragt nach den *evaluationsbeeinflussenden Faktoren (F5)*, wurden für alle Unternehmen die vorhandenen Ressourcen als entscheidender Faktor genannt. Dies geht mit früheren Befunden einher, die ebenfalls zeitliche, personelle, und finanzielle Ressourcen als evaluationsbeeinflussend identifizieren (z.B. Huerta et al., 2006, Kennedy et al., 2014, Twitchell et al., 2014, Williams & Nafukho, 2015). Dies könnte darin begründet sein, dass die Evaluation betrieblicher Weiterbildungsmassnahmen bei der Mehrheit der befragten Unternehmen lediglich einen mittleren Stellenwert hat und die Belegschaft gegenüber der Evaluation überwiegend nicht positiv eingestellt ist. Aus den Interviews geht nur bedingt hervor, wie mit den evaluationsbeeinflussenden Faktoren im Unternehmen umge-

gangen wird. Zukünftige Forschung sollte sich mit der Frage beschäftigen, ob und warum die genannten Faktoren tatsächlich die Qualität der Evaluation beeinflussen und wenn ja, welche Massnahmen ergriffen werden können.

Mit Blick auf das Gesamtmaterial fällt zudem auf, dass bisher keine der Studien die Digitalisierung als evaluationsbeeinflussenden Faktor identifiziert hat, während diese in der aktuellen Studie häufig genannt wird. Interessant ist weiterhin, dass die interviewten Personen unterschiedliche allgemeine und auf das eigene Unternehmen bezogene evaluationsbeeinflussende Faktoren nennen. Die Interviews geben jedoch keinen Aufschluss darüber, warum dies der Fall ist. An diesem Punkt sollte zukünftige Forschung ansetzen. Ggf. könnte es damit zusammenhängen, dass die befragten Personen den Evaluationsprozess im eigenen Unternehmen anders wahrnehmen als den von anderen Unternehmen.

Insgesamt wird deutlich, dass die Unternehmen der Studie ihren Evaluationsprozess noch weiter professionalisieren könnten. Für die Evaluationspraxis in den Unternehmen wäre ein möglicher Ansatzpunkt, die Evaluationsdaten mit verschiedenen Methoden auf Basis unterschiedlicher Quellen zu erheben, um deren Validität zu erhöhen (Rossi et al., 2019). Weiterhin sollten der Evaluationsprozess sowie die eingesetzten Instrumente auf Basis wissenschaftlicher Erkenntnisse entwickelt werden. Zur Feststellung kausaler Zusammenhänge bieten sich experimentelle Designs an. Im Hinblick auf den Evaluationsprozess und die Datenauswertung könnten Digitalisierungs- und Automatisierungsmöglichkeiten die Qualität erhöhen bzw. den Arbeitsaufwand verringern. Auch die Nutzung von künstlicher Intelligenz z.B. bei der Auswertung freier Kommentarfelder wäre denkbar. Grosses Potenzial liegt zudem in der vermehrten Nutzung und Kommunikation der Evaluationsergebnisse. Letztendlich benötigt es für all diese Punkte ausreichende zeitliche, finanzielle und personelle Ressourcen, um eine qualitativ hochwertige und zielgerichtete Weiterbildungsevaluation durchzuführen.

Zusammenfassend liefert die Studie einen ersten Einblick in die aktuelle betriebliche Praxis der Trainingsevaluation von Unternehmen im deutschsprachigen Raum. Die Ergebnisse müssen jedoch vor dem Hintergrund verschiedener Limitationen betrachtet werden. Bei der Auswahl der interviewten Personen wurde versucht, das Feld möglichst breit abzubilden, um eine hohe Varianz innerhalb der Antworten zu erhalten. Es ist dennoch wahrscheinlich, dass nicht alle Aspekte und Phänomene durch das Datenmaterial aufgezeigt werden konnten. Weiterhin besteht, wie bereits oben erwähnt, die Möglichkeit der Selbstselektion der Befragten, so dass ggf. Unternehmen, in denen die Evaluation eine untergeordnete Rolle spielt, nicht ausreichend abgebildet werden konnten. Durch das explorative, qualitative Vorgehen ist die Stichprobe nicht repräsentativ. Auf Basis der Ergebnisse sollten daher zukünftige Studien die Bedeutung der einzelnen Aspekte in der Praxis anhand einer repräsentativen quantitativen Stichprobe genauer untersuchen. Durch das explorative, qualitative Vorgehen konnte jedoch ein breiter Einblick in den Status quo der betrieblichen Weiterbildungsevaluation gewonnen und interessante Aspekte adressiert werden, die in zukünftigen Studien weiter vertieft werden sollten.

6 Literatur

- Aguinis, H., & Kraiger, K. (2009). Benefits of training and development for individuals and teams, organizations, and society. *Annual review of psychology*, 60, 451–474. <https://doi.org/10.1146/annurev.psych.60.110707.163505>
- Bächmann, A.-C., Abraham, M. & Huber, M. (2019). When do firms evaluate further training measures?. *International Journal of Manpower*, 40(2), 190–210. <https://doi.org/10.1108/IJM-06-2017-0146>
- Becker, F., Meißner, A. & Werning, E. (2008). Evaluation externer Weiterbildungsmaßnahmen. Mehr als nur ‚Happy Sheets‘?. *Personalführung*, 41(12), 64–71.
- Bober, C. F., & Bartlett, K. R. (2004). The utilization of training program evaluation in corporate universities. *Human Resource Development Quarterly*, 15(4), 363–383. <http://dx.doi.org/10.1002/hrdq.1111>
- Brennan, R. L., & Prediger, D. J. (1981). Coefficient kappa: Some uses, misuses, and alternatives. *Educational and Psychological Measurement*, 41(3), 687–699. <https://doi.org/10.1177/001316448104100307>
- Bundesministerium für Bildung und Forschung. (2024). *Weiterbildungsverhalten in Deutschland 2022. Ergebnisse des Adult Education Survey – AES-Trendbericht*. BMBF. https://www.bmbf.de/SharedDocs/Publikationen/de/bmbf/1/26667_AES-Trendbericht_2022.pdf?__blob=publicationFile&v=4. Zugriffen: 17.06.2024
- Döring, N. (2023). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (6- Aufl.). Springer. <https://doi.org/10.1007/978-3-662-64762-2>
- Giangreco, A., Carugati, A., & Sebastiani, A. (2010). Are we doing the right thing? Food for thought on training evaluation and its context. *Personnel review*, 39(2), 162–177. <https://doi.org/10.1108/00483481011017390>
- Hauser, A., Weisweiler, S., & Frey, D. (2020). Because ‘happy sheets’ are not enough—a meta-analytical evaluation of a personnel development program in academia. *Studies in Higher Education*, 45(1), 55–70. <https://doi.org/10.1080/03075079.2018.1509306>
- Ho, A. D., Arendt, S. W., Zheng, T., & Hanisch, K. A. (2016). Exploration of hotel managers' training evaluation practices and perceptions utilizing Kirkpatrick's and Phillips's models. *Journal of Human Resources in Hospitality & Tourism*, 15(2), 184–208. <https://doi.org/10.1080/15332845.2016.1084861>
- Hoffmann, E. (2018). *Personalentwicklung und -controlling. Strategien für den Mittelstand*. Springer Gabler. <https://doi.org/10.1007/978-3-658-15551-3>
- Huerta, M. E., Audet, X. L. & Peregort, O. P. (2006). In-company training in Catalonia: organizational structure, funding, evaluation and economic impact. *International Journal of Training and Development*, 10(2), 140–163. <https://doi.org/10.1111/j.1468-2419.2005.00249.x>
- Kauffeld, S. (2016). *Nachhaltige Personalentwicklung und Weiterbildung*. Springer. <https://doi.org/10.1007/978-3-662-48130-1>
- Kennedy, P. E., Chyung, S. Y., Winiecki, D. J., & Brinkerhoff, R. O. (2014). Training professionals' usage and understanding of Kirkpatrick's Level 3 and Level 4 evaluations. *International Journal of Training and Development*, 18(1), 1–21. <https://doi.org/10.1111/ijtd.12023>
- Kirkpatrick, J. D., & Kirkpatrick, W. K. (2016). *Kirkpatrick's four levels of training evaluation*. atd press.

- Kucherov, D. & Manokhina, D. (2017). Evaluation of Training Programs in Russian Manufacturing Companies. *European Journal of Training and Development*, 41(2), 119–143. <https://doi.org/10.1108/EJTD-10-2015-0084>
- Liechti, L. & Abraham, M. (2011). Die Evaluation von betrieblichen Weiterbildungsmaßnahmen. *Zeitschrift für Betriebswirtschaft*, 81(3), 241–262. <https://doi.org/10.1007/s11573-011-0440-4>
- Mayring, P. (2022). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (13 Aufl.). Beltz.
- Phillips, J.J. (2003), *Return on Investment in Training and Performance Improvement Programs*, Butterworth-Heinemann.
- Peters, L., Baum, J., & Stephens, G. (2011). Creating ROI in leadership development. *Organizational Dynamics*, 2(40), 104–109.
- Rohmann, A., Froncek, B., Mazziotta, A., & Piper, V. (2017). Current evaluation practices of diversity trainers in German-speaking countries. *International Journal of training research*, 15(2), 148–159. <https://doi.org/10.1080/14480220.2017.1299864>
- Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2019). *Evaluation: A systematic approach*. Sage publications.
- Saks, A. M. & Burke, L. A. (2012): An investigation into the relationship between training evaluation and the transfer of training. *International Journal of Training and Development*, 16(2), 118–127. <https://doi.org/10.1111/j.1468-2419.2011.00397.x>
- Schäfer, E., & Ebersbach, A. (2021). *Die digitale Transformation in der Weiterbildung*. Springer. <https://doi.org/10.1007/978-3-662-64605-2>
- Seyda, S., & Placke, B. (2020). IW-Weiterbildungserhebung 2020: Weiterbildung auf Wachstumskurs. *IW-Trends-Vierteljahresschrift zur empirischen Wirtschaftsforschung*, 47(4), 105–123.
- Srimannarayana, M. (2017). From reactions to return on investment: a study on training evaluation practices. *Indian Journal of Industrial Relations*, 53(1), 1–20.
- Twitchell, S., Holton, E. F. & Trott, J. W., Jr. (2000). Technical Training Evaluation Practices in the United States. *Performance Improvement Quarterly*, 13(3), 84–109. <https://doi.org/10.1111/j.1937-8327.2000.tb00177.x>
- Williams, R. C. & Nafukho, F. M. (2015). Technical Training Evaluation Revisited: An Exploratory, Mixed-Methods Study. *Performance Improvement Quarterly*, 28(1), 69–93. <https://doi.org/10.1002/piq.21187>
- Yadapadithaya, P. S. & Stewart, J. (2003). Corporate training and development policies and practices: a cross-national study of India and Britain. *International Journal of Training and Development*, 7(2), 108–123. <https://doi.org/10.1111/1468-2419.00175>
- Yates, K. (2015). Managing, tracking and evaluating coaching Part 1: where are you now?. *Industrial and Commercial Training*, 47(1), 36–41. <https://doi.org/10.1108/ICT-08-2014-0057>

Selina Ebert ist Doktorandin am Lehrstuhl für Personalentwicklung am Fachbereich Wirtschaftswissenschaften der Universität Konstanz.

Anschrift: Universitätsstraße 10, 78464 Konstanz, Deutschland, Tel.: +49 (0)7531 88-5114,
E-Mail: selina.ebert@uni-konstanz.de

Caroline Bonnes, PD Dr., ist Wissenschaftliche Mitarbeiterin am Deutschen Institut für Erwachsenenbildung – Leibniz-Zentrum für Lebenslanges Lernen e.V. in Bonn.

Anschrift: Heinemannstr. 12–14, 53175 Bonn, Deutschland, Tel.: +49 (0)228 3294-309,
E-Mail: bonnes@die-bonn.de.

Sabine Hochholdinger, Prof., Dr., ist Inhaberin des Lehrstuhls für Personalentwicklung am Fachbereich Wirtschaftswissenschaften der Universität Konstanz.

Anschrift: Universitätsstraße 10, 78464 Konstanz, Deutschland, Tel.: +49 (0)7531 88-3015,
E-Mail: sabine.hochholdinger@uni-konstanz.de