Was kostet Kompetenz, bevor sie entsteht? Lernzeit und Lernkosten a priori berechnen

Evidenzbasierte Schätzverfahren aus Lernpsychologie, Berufspädagogik, Instructional Design, Bildungsökonomie und Prognoseforschung – mit Benchmarktabellen, Rechenmodell, Faktencheck und einem Leitfaden für den Mittelstand

Dr. Walter Lieberei
03.11.2026

Kaum eine Frage hört die Personalentwicklung so oft wie diese: „Was kostet uns das – und wie lange dauert es?“ Meist antwortet sie mit Seminartagen und Teilnahmegebühren. Damit beschreibt sie aber die *Dauer eines Angebots*, nicht die *Zeit, bis Beschäftigte etwas können*. Der größte Kostenblock bleibt dabei oft unsichtbar: die Arbeitszeit der Lernenden. Hinzu kommt ein ernüchternder Befund. Eine allgemeingültige Tabelle, die „Lerntiefe mal Lernumfang“ in Minuten umrechnet, gibt es nicht. Die im Umlauf befindlichen Faustwerte halten einer Quellenprüfung meist nicht stand.

Dieser Beitrag zeigt, was stattdessen trägt. Seine Leitthese: *Eine seriöse A-priori-Kalkulation ist keine Formel, sondern ein Prognoseprozess.* Sie beginnt beim prüfbaren Kompetenzziel und beim Vorwissen der Zielgruppe. Entwicklung und Durchführung rechnet sie getrennt. Die Unsicherheit weist sie als Bandbreite aus, und nach einem Pilotdurchgang wird sie neu kalibriert. Wer so vorgeht, kann dem Management keine scheinbar exakte Zahl nennen, aber ein Budget, das mit hoher Wahrscheinlichkeit hält.

1. Einleitung: Warum die Vorab-Kalkulation jetzt wettbewerbsrelevant ist

Erstens beschleunigt sich der Kompetenzwandel. Generative Künstliche Intelligenz (KI), neue Regulierung und technologische Sprünge verkürzen die Halbwertszeit von Fachwissen. Wer Kompetenz zu spät oder zu flach aufbaut, verliert Angebote, Qualität und Fachkräfte. Wer sie zu breit und zu teuer aufbaut, verliert Geld und Akzeptanz.

Zweitens wächst der regulatorische Druck. Seit dem 2. Februar 2025 verpflichtet Art. 4 der KI-Verordnung (EU) 2024/1689 Anbieter und Betreiber von KI-Systemen, für ein „ausreichendes Maß an KI-Kompetenz“ ihres Personals zu sorgen. Eine Stundenzahl nennt die Verordnung nicht. Gerade deshalb müssen Unternehmen begründen können, welches Kompetenzniveau sie anstreben und mit welchem Aufwand sie es erreichen. Die EU-Kommission hat im November 2025 im Rahmen ihres Digital-Omnibus-

Pakets eine Abschwächung dieser Pflicht vorgeschlagen. Der aktuelle Gesetzgebungsstand ist vor Umsetzungsentscheidungen zu prüfen.

Drittens ist die Lernzeit die teuerste und zugleich am schlechtesten geschätzte Größe. In der europäischen Unternehmenserhebung CVTS6 entfielen in Deutschland 57 Prozent der Kosten betrieblicher Weiterbildungskurse auf Personalausfallkosten, also auf die Arbeitszeit der Teilnehmenden (Statistisches Bundesamt, 2022). Wer die Lernzeit um ein Drittel unterschätzt, verfehlt sein Budget daher nicht um Kleinigkeiten.

Viertens klafft eine Lücke zwischen Versprechen und Evidenz. In Praxisleitfäden kursieren Faktoren wie „Bloom-Stufe mal zwei", „Selbststudium gleich Präsenz mal drei" oder „Arbeitsausfall macht 60 bis 80 Prozent der Kosten aus". Kapitel 8 prüft sie. Grundlage dieses Beitrags sind die Originalstudie von Chapman (2010), der ECTS-Leitfaden der Europäischen Kommission (2015), Meta-Analysen zu Lernen, Transfer und Vergessen, die Kostenanalyse-Standards der US-Bildungsforschung (Institute of Education Sciences, 2020; Hollands et al., 2021) sowie deutsche Amtsstatistik und Erhebungen bis 2026. Der Beitrag knüpft an unsere Fachbeiträge Nr. 35 (Kompetenzmodelle rechtskonform entwickeln), Nr. 37 (Anforderungsanalyse), Nr. 52 (Selbstentwicklungs-Loop) und Nr. 54 (360°-Feedback) an.

Praxisfall: Die Kessler Fluidtechnik GmbH (fiktiv)

Der Fall ist erfunden, aber typisch; er begleitet uns durch den Beitrag. Kessler ist ein familiengeführter Hersteller von Hydraulikkomponenten mit 520 Beschäftigten und einem Betriebsrat. Ein Wettbewerber erstellt technische Angebote neuerdings mit KI-Unterstützung in zwei statt zehn Arbeitstagen. Die Geschäftsführung beschließt: „Unsere 48 Vertriebs- und Applikationsingenieurinnen und -ingenieure müssen KI-gestützte Angebots- und Auslegungsprozesse bis Ende Q2 2027 sicher beherrschen – und ich will vorher wissen, was das kostet." Die Personalleiterin soll in vier Wochen eine Entscheidungsvorlage liefern.

2. Begriffe: Was genau soll vorab berechnet werden?

2.1 Fünf Zeitgrößen, die nicht verwechselt werden dürfen

Viele Fehlkalkulationen beginnen damit, dass unterschiedliche Zeitgrößen in einer Formel landen. Eine einstündige Lerneinheit für 100 Personen erzeugt bei vollständiger Bearbeitung mindestens 100 Personenstunden Lernaufwand. Ihre Entwicklung hat womöglich 80 Personalstunden gekostet. Die Durchführung kann dennoch an einem einzigen Kalendertag stattfinden.

Tabelle 1: Zeitgrößen der Bildungsplanung

Größe	Definition	Einheit	Typischer Fehler
Angebotsdauer	nominelle Dauer von Seminar, Modul oder Video	Unterrichtsstunden, Minuten	wird als Lernzeit verbucht
Lernaufwand	Bearbeitung, Übung, Prüfung, Wiederholung und Transfer	Personenstunden je Teilnehmendem	Selbststudium und Transfer fehlen
Zeit bis zum Kriterium	Aufwand, bis das Kompetenzziel nachweislich erreicht ist	Personenstunden bis zum Nachweis	wird gar nicht geplant
Entwicklungsaufwand	Arbeit aller Rollen zur Erstellung des Angebots	Personalstunden je Rolle	wird mit Lernzeit verwechselt
Kalenderdauer	Zeit zwischen Start und Meilenstein	Wochen	wird aus Stunden linear abgeleitet

Quellen: Chapman (2010); Europäische Kommission (2015); eigene Zusammenstellung. Alle Stunden in diesem Beitrag sind 60-Minuten-Stunden; Unterrichtseinheiten à 45 Minuten sind vor jeder Rechnung umzurechnen.

2.2 Entwicklungskosten und Durchführungskosten

Für die Managemententscheidung braucht es zwei getrennte Rechnungen. Die **Entwicklungskosten** fallen einmalig beziehungsweise je Version an: Analyse, Design, Produktion, Qualitätssicherung, Pilotierung, Freigaben und Pflege. Sie haben eine *Dauer*, also Kalenderwochen bis zur Einsatzreife, und einen *Preis*, also Personalstunden mal Kostensatz plus Fremdleistungen. Die **Durchführungskosten** hängen von Zielgruppe, Lernzeit, Lerntiefe und Gruppengröße ab. Dazu zählen die Zeit der Lernenden, Trainerinnen und Trainer, Lizenzen, Betreuung, Prüfungen, Transferbegleitung und Administration. Ein Teil davon ist variabel je Person, ein Teil sprungfix je Gruppe.

2.3 Lernzieltiefe operationalisieren

Die revidierte Lernzieltaxonomie nach Bloom unterscheidet sechs kognitive Prozesse vom Erinnern bis zum Erschaffen sowie vier Wissensarten (Krathwohl, 2002). Sie beschreibt Ziele und Prüfungsaufgaben, aber keine Zeitbedarfe. Für die Kalkulation braucht es mehr: ein **prüfbares Kompetenzziel**, das Handlung, Bedingungen, Qualitätsmaßstab und Prüfzeitpunkt benennt. Als Zielanker helfen domänenspezifische Niveaumodelle, etwa der europäische Referenzrahmen für digitale Kompetenz DigComp 2.2 mit acht Kompetenzstufen (Vuorikari et al., 2022).

„Mit hoher Sicherheit“ kann zweierlei heißen, und beides wird getrennt geplant. **Zielerreichungssicherheit** bedeutet zum Beispiel: Mindestens 90 Prozent der Zielgruppe bestehen eine realitätsnahe Arbeitsprobe. **Budgetsicherheit** bedeutet zum Beispiel: Die Kosten werden mit 80 Prozent Wahrscheinlichkeit nicht überschritten. Diesen Wert nennt man das 80. Perzentil oder kurz P80.

Kasten 1: Fünf Prüffragen – ist mein Lernziel überhaupt kalkulierbar?

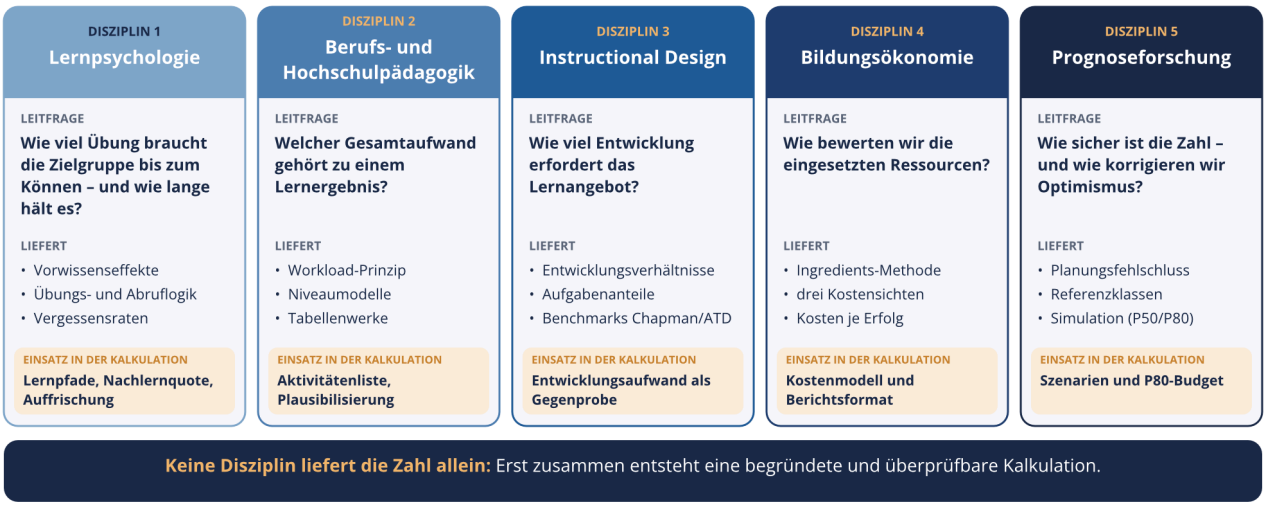
1. Ist die Zielhandlung beobachtbar und an einer realen Arbeitsaufgabe prüfbar?
2. Sind Bedingungen und Hilfsmittel festgelegt, etwa „mit“ oder „ohne“ KI-Unterstützung?
3. Gibt es einen Qualitätsmaßstab, und sind kritische Fehler gesondert definiert?
4. Ist der Zeitpunkt des Nachweises bestimmt – unmittelbar und nach einigen Wochen?
5. Ist das Vorwissen der Zielgruppe aufgabenbezogen erhoben?

Eigene Zusammenstellung auf Basis von Krathwohl (2002) und Europäische Kommission (2015).

3. Evidenzbasierte Ansätze aus fünf Disziplinen

Keine Disziplin allein liefert eine verlässliche Vorab-Zahl. Jede beantwortet aber eine Teilfrage. Erst zusammen ergeben sie eine Kalkulation, die man begründen und später überprüfen kann.

Abbildung 1: Fünf Disziplinen, fünf Teilfragen



Eigene Darstellung.

3.1 Lernpsychologie: Lernzeit ist eine bedingte Größe

Den begrifflichen Grundstein legte Carroll (1963): Der Lernerfolg hängt vom Verhältnis der *aufgewendeten* zur *benötigten* Zeit ab. Die benötigte Zeit wiederum hängt von Eignung, Unterrichtsqualität und Verständnisfähigkeit ab. Das Modell ist ein Denkrahmen, keine fertige Rechenformel. Es hat aber eine klare Konsequenz: Eine Lernzeitangabe ohne Bezug zu Zielgruppe, Aufgabe und Unterstützung ist nicht überprüfbar.

Bloom (1968) zog daraus die Folgerung des *Mastery Learning*: Das Ziel bleibt fest, die Zeit variiert. Die Meta-Analyse von Kulik, Kulik und Bangert-Drowns (1990) über 108 kontrollierte Evaluationen fand positive Effekte auf Prüfungsleistungen, besonders bei schwächeren Lernenden. Sie zeigt aber auch, dass solche Programme mehr Lernzeit beanspruchen können. Für die Kalkulation heißt das: Wer ein festes Kompetenzziel „mit hoher Sicherheit“ erreichen will, muss Nachlernschleifen einplanen.

Neuere lernanalytische Befunde schärfen das Bild. Koedinger, Carvalho, Liu und McLaughlin (2023) werteten rund 1,3 Millionen Beobachtungen aus 27 Datensätzen digital gestützten Übens aus. Die Lernenden unterschieden sich stark in ihrem Anfangswissen, aber deutlich weniger in ihrer modellierten Lernrate. Im Mittel brauchten sie rund sieben Übungsgelegenheiten je Wissenskomponente. Eine Reanalyse derselben Daten zeigt allerdings, dass die geschätzte Streuung der Lernraten davon abhängt, welche Beobachtungen einbezogen werden (Lee et al., 2026). Die praktische Lehre ist dennoch robust: Das Vorwissen erklärt einen großen Teil der Zeitunterschiede. Es gehört deshalb in die Eingangsdiagnose, und unterschiedliche Lernpfade sind wirtschaftlicher als ein Einheitspfad.

Die Cognitive Load Theory begründet, warum Novizinnen und Novizen Anleitung und ausgearbeitete Beispiele brauchen, während dieselben Hilfen Fortgeschrittene bremsen können (Sweller et al., 2019). Daraus folgen verschiedene Lernpfade, aber kein pauschaler Zeitaufschlag. Zwei Lerntechniken sind besonders gut belegt: Abrufübungen und über die Zeit verteiltes Lernen (Dunlosky et al., 2013). Wer nach einem Lerndurchgang einen Test absolviert, behält langfristig mehr als jemand, der den Stoff erneut liest (Roediger & Karpicke, 2006). Der günstigste Abstand zwischen Lernepisoden hängt davon ab, wie lange das Wissen halten soll (Cepeda et al., 2006). Solche Wiederholungen werden als eigene Aktivitäten geplant. Sie kosten wenige Stunden, verlängern aber die Kalenderdauer.

Für eng umrissene, wiederholbare Teilaufgaben lassen sich Übungskurven aus Pilotdaten schätzen. Dabei passen exponentielle Funktionen individuelle Verläufe oft besser an als das klassische Potenzgesetz (Heathcote et al., 2000). Ohne eigene Daten sind solche Kurven jedoch Spekulation. Auch die populäre „10.000-Stunden-Regel“ trägt nicht. In der Meta-Analyse von Macnamara, Hambrick und Oswald (2014) erklärte gezielte Übung über alle Domänen hinweg etwa 12 Prozent der Leistungsunterschiede, in Berufen weniger als 1 Prozent. Das relativiert die These, dass vor allem Übung Expertise erklärt (Ericsson et al., 1993).

Schließlich ist Können nicht von Dauer. Ohne Anwendung verfällt es, und zwar umso stärker, je länger die Pause währt. Die Effektstärke des Verlusts steigt von nahezu null unmittelbar nach dem Training auf d = −1,4 nach mehr als einem Jahr Nichtgebrauch (Arthur et al., 1998). Eine neue Meta-Analyse über 457 Berichte schätzt für prozedurale Fertigkeiten: Die Hälfte des Lerngewinns ist nach rund 6,5 Monaten (Genauigkeit) beziehungsweise 11 bis 13 Monaten (gemischte Maße und Geschwindigkeit) verloren (Tatel & Ackerman, 2025). Schon Ebbinghaus’ Vergessenskurve hat sich in einer modernen Replikation im Kern bestätigt (Murre & Dros, 2015). Wer eine Kompetenz *dauerhaft* sichern will, muss daher Auffrischung und Anwendungsgelegenheiten mitkalkulieren.

3.2 Berufs- und Hochschulpädagogik: Workload denken, Tabellenwerke richtig lesen

Das ECTS-System verknüpft Lernergebnisse mit dem typischen Gesamtarbeitsaufwand. Ein Leistungspunkt entspricht 25 bis 30 Stunden, und zwar einschließlich Selbststudium, Projekten und Prüfungsvorbereitung (Europäische Kommission, 2015). Für die betriebliche Planung zählt vor allem das Prinzip: *Alle notwendigen Lernaktivitäten zählen*, nicht nur die Präsenzzeit. Ebenso wichtig ist, was der Leitfaden *nicht* enthält: einen verbindlichen Faktor zwischen Kontaktzeit und Selbststudium oder eine Umrech-

nung von Taxonomiestufen in Stunden. Er verlangt vielmehr, den geschätzten Workload nachträglich zu überprüfen und bei Abweichungen anzupassen. Lehrende unterschätzen ihn nämlich regelmäßig.

Echte Tabellenwerke für Lernzeiten gibt es durchaus, aber nur dort, wo Ziel, Prüfung und Population standardisiert sind. Das bekannteste Beispiel sind die Richtwerte für angeleitete Lernstunden im Fremdsprachenerwerb.

Tabelle 2: Tabellenwerke und ihr Geltungsbereich

Tabellenwerk	Bezugsgröße	Beispielwerte	Geltungsbereich und Grenzen
Guided Learning Hours, Cambridge (GER-Niveaus)	kumulierte angeleitete Stunden ab null	A2: 180–200 h · B1: 350–400 h · B2: 500–600 h · C1: 700–800 h	Englisch als Fremdsprache; ausdrücklich nur Richtwerte, individuell abweichend
ECTS-Leitfaden	Gesamtworkload je Leistungspunkt	25–30 h	Hochschule; keine Umrechnung von Taxonomiestufen
Rahmenpläne geregelter Fortbildungen	empfohlene Lehrgangsstunden	je Rahmenplan	beschreiben Angebotsdauer, nicht individuelle Lernzeit
Chapman (2010); ATD-Studien	Entwicklungsstunden je fertiger Lernstunde	siehe Tabelle 3	nur Entwicklungsaufwand, nicht Lernzeit

Quellen: Cambridge Assessment English (o. J.); Europäische Kommission (2015); Chapman (2010); eigene Zusammenstellung.

Tabellenwerke ersetzen also keine Schätzung. Sie sind aber wertvolle *Referenzklassen*, an denen man die eigene Schätzung plausibilisiert (Kapitel 3.5). Für komplexe berufliche Kompetenzen liefert die Instruktionsdesign-Forschung zudem eine brauchbare Mengeneinheit. Im Vier-Komponenten-Modell (4C/ID) werden ganze Aufgaben in Aufgabenklassen mit abnehmender Unterstützung geübt (van Merriënboer & Kirschner, 2018). Die *Zahl der Lernaufgaben bis zur selbstständigen Ausführung*, multipliziert mit ihrer Bearbeitungsdauer, ist eine solidere Basis für die Lernzeit als Seiten- oder Folienzahlen.

3.3 Instructional Design: Was die Entwicklung kostet

Die meistzitierte Benchmarkstudie zum Entwicklungsaufwand stammt von Chapman (2010). Sie beruht auf 249 Organisationen, die zusammen 3.947 Fachkräfte der Lernentwicklung repräsentierten. Gemessen wurde, wie viele Personalstunden aller beteiligten Rollen zusammen nötig sind, um *eine fertige Lernstunde* zu erstellen. Das umfasst Analyse, Design, Medienproduktion, Reviews durch Fachexpertinnen und -experten, Qualitätssicherung und Projektmanagement.

Tabelle 3: Entwicklungsstunden je fertiger Lernstunde

Format	einfach	typisch	aufwendig	Ø interne Kosten je fertiger Stunde (2010)
Präsenztraining	22	43	82	5.934 USD
E-Learning Stufe 1 (Seiten, Grafiken, Test)	49	79	125	10.054 USD
E-Learning Stufe 2 (≥ 25 % interaktive Übungen, Medien)	127	184	267	18.583 USD
E-Learning Stufe 3 (Simulation, Serious Game)	217	490	716	50.371 USD

Quelle: Chapman (2010, S. 28 und Kostenübersicht). Einzelne Befragte nannten für sehr aufwendige Simulationen mehr als 2.000 Stunden.

Für die Nutzung gelten fünf Lesehinweise.

1. Die Werte messen *Entwicklung*, nicht Lernzeit.
2. „Einfach" und „aufwendig" sind Durchschnitte von Projektkategorien, keine Konfidenzgrenzen.
3. Skaleneffekte bei längeren Kursen hat die Studie ausdrücklich nicht untersucht.
4. Die Angaben zu Blended Learning beruhen auf getrennten Fragen und lassen sich laut Chapman nicht zu einem Stundenpreis verrechnen.
5. Es handelt sich um eine Branchenbefragung ohne Peer Review mit US-Dollarwerten von 2010.

Nützlich bleibt die Aufgabenverteilung als Gliederung einer eigenen Schätzung. Beim interaktiven E-Learning entfallen etwa 9 Prozent auf die Analyse, 13 Prozent auf das didaktische Design, 18 Prozent auf Autorenarbeit und Programmierung, 6 Prozent auf Reviews durch Fachexpertinnen und -experten und 4 Prozent auf den Piloten.

Die Association for Talent Development (ATD) hat die Frage seit 2003 in mehreren Befragungswellen verfolgt. In der Welle 2021 wurde die Bezugsgröße von der „Lernstunde" auf typische Lerneinheiten umgestellt, weil kaum ein Lernprodukt genau eine Stunde dauert (Defelice, 2021). Ältere und neuere Werte sind deshalb nicht unmittelbar vergleichbar.

Offen ist, wie stark KI die Entwicklung beschleunigt. In einer ATD-Befragung gaben 37 Prozent der Instructional Designer, die KI nutzen, an, KI habe die Zeit für das Kursdesign „stark" verkürzt (Association for Talent Development, 2025b). Das ist eine Selbstauskunft und keine gemessene Stundenersparnis. Der zusätzliche Aufwand für fachliche Prüfung und Freigabe wird dabei selten mitgerechnet. Ein pauschaler „KI-Abschlag" auf die Chapman-Werte ist daher nicht begründet. KI-Effekte gehören als eigenes Szenario in die Kalkulation.

3.4 Bildungsökonomie und Bildungscontrolling: Ressourcen richtig bewerten

Die Bildungsökonomie hat für die Kostenseite ein bewährtes Verfahren: die *Ingredients-Methode* (Levin et al., 2018). Zuerst werden alle eingesetzten Ressourcen *mengenmäßig* erfasst: Personenstunden je Rolle, Räume, Lizenzen, Material. Erst danach werden sie aus einer vorab festgelegten Perspektive mit Preisen bewertet, einschließlich bereits vorhandener Ressourcen, mit dokumentierten Annahmen und

einer Sensitivitätsanalyse (Institute of Education Sciences, 2020; Hollands et al., 2021). Für Unternehmen folgen daraus drei Kostensichten, die nebeneinander berichtet und nicht vermischt werden:

- die **Auszahlungssicht**: Welche zusätzlichen Zahlungen löst die Maßnahme aus?
- die **interne Ressourcensicht**: Welche Kapazität wird gebunden, bewertet mit definierten Kostensätzen?
- die **Opportunitätssicht**: Welcher Deckungsbeitrag entgeht, weil die Zeit nicht anders genutzt wird?

Fortgezahltes Gehalt ist in der Regel keine *zusätzliche* Auszahlung. Es ist aber ein sinnvoller Maßstab für die gebundene Arbeitskapazität. Der interne Stundensatz wird aus den jährlichen Arbeitgeberpersonalkosten und einer konsistent definierten Zahl verfügbarer Stunden gebildet. Ein Kundenverrechnungssatz ist dafür kein Ersatz, denn er enthält Gewinn- und Vertriebsanteile.

Tabelle 4: Referenzwerte für Deutschland und international

Kennzahl	Wert	Quelle
Anteil der Personalausfallkosten an den Kosten betrieblicher Weiterbildungskurse	57 % (2020)	Statistisches Bundesamt (2022), CVTS6
Weiterbildungsinvestition je Beschäftigtem (inkl. Lernzeit)	1.347 € (2022)	Seyda et al. (2024)
Weiterbildungsstunden je Beschäftigtem	20,3 h (2022)	Seyda et al. (2024)
Arbeitskosten je geleistete Stunde, Produzierendes Gewerbe und Dienstleistungen	45,00 € (2025)	Statistisches Bundesamt (2026)
Kostensätze je Teilnehmerstunde zugelassener Weiterbildungsmaßnahmen	Bundesdurchschnitt je Maßnahmeart, gültig 07/2026–06/2028	Bundesagentur für Arbeit (2026)
Formale Lernstunden je Beschäftigtem; Kosten je genutzter Lernstunde	13,7 h; 165 USD (2024)	Association for Talent Development (2025a)

Eigene Zusammenstellung. Die Werte dienen der Einordnung; sie sind keine Prognose für eine Einzelmaßnahme.

Für den Mittelstand ist zusätzlich die Förderung relevant. Seit dem 1. April 2024 bezuschusst die Bundesagentur für Arbeit nach § 82 SGB III die Weiterbildung Beschäftigter gestaffelt nach Betriebsgröße. Betriebe unter 50 Beschäftigten erhalten bis zu 100 Prozent der Lehrgangskosten und bis zu 75 Prozent Zuschuss zum Arbeitsentgelt. Bei 50 bis 499 Beschäftigten sind es jeweils 50 Prozent, ab 500 Beschäftigten jeweils 25 Prozent (Bundesagentur für Arbeit, 2024). Voraussetzungen sind unter anderem eine Maßnahmedauer von mehr als 120 Stunden und eine Zulassung nach der Akkreditierungs- und Zulassungsverordnung Arbeitsförderung (AZAV). Ausgeschlossen sind Schulungen, zu denen der Arbeitgeber rechtlich verpflichtet ist. Die Förderung ist eine Ermessensleistung.

3.5 Prognoseforschung und Projektmanagement: Wie sicher ist die Zahl?

Die vielleicht wichtigste Erkenntnis für die Vorab-Kalkulation stammt nicht aus der Pädagogik, sondern aus der Entscheidungspsychologie. Menschen unterschätzen systematisch, wie lange ihre eigenen Vorhaben dauern. Das gilt sogar dann, wenn sie wissen, dass ähnliche Vorhaben früher länger gedauert haben. Buehler, Griffin und Ross (1994) nannten dieses Muster den *Planungsfehlschluss*. Kahneman und

Lovallo (1993) führten ihn auf die „Innensicht“ zurück: Man plant vom eigenen Projekt aus und blendet die Erfahrung mit vergleichbaren Projekten aus. Als Gegenmittel hat sich die **Referenzklassenprognose** bewährt (Flyvbjerg, 2008). Die eigene Schätzung wird an der Verteilung tatsächlicher Werte vergleichbarer, abgeschlossener Projekte ausgerichtet. Für die Personalentwicklung sind die Tabellenwerke aus Kapitel 3.2 und 3.3 und vor allem eigene frühere Maßnahmen diese Referenzklasse.

Aus dem Projektmanagement kommen zwei Werkzeuge. Erstens die *Drei-Punkt-Schätzung*: Für jede wichtige Annahme werden ein günstiger, ein wahrscheinlicher und ein ungünstiger Wert festgelegt. Zweitens die *Monte-Carlo-Simulation*: Ein Rechner zieht viele tausend Male zufällige, plausible Kombinationen dieser Werte. Daraus entsteht eine Kostenverteilung, aus der man P50 (Median) und P80 ablesen kann. Beide Verfahren erzeugen keine Gewissheit. Sie machen aber sichtbar, wie unsicher eine Zahl ist. Schließlich gilt für wiederholte Durchführungen ein Lernkurveneffekt. Wright (1936) beschrieb ihn für die Flugzeugfertigung: Mit jeder Verdopplung der Stückzahl sinkt der Aufwand um einen annähernd konstanten Anteil. Für die zweite und dritte Trainingswelle sinken Vorbereitungs- und Organisationsaufwand entsprechend.

Tabelle 5: Was jede Disziplin zur Kalkulation beiträgt

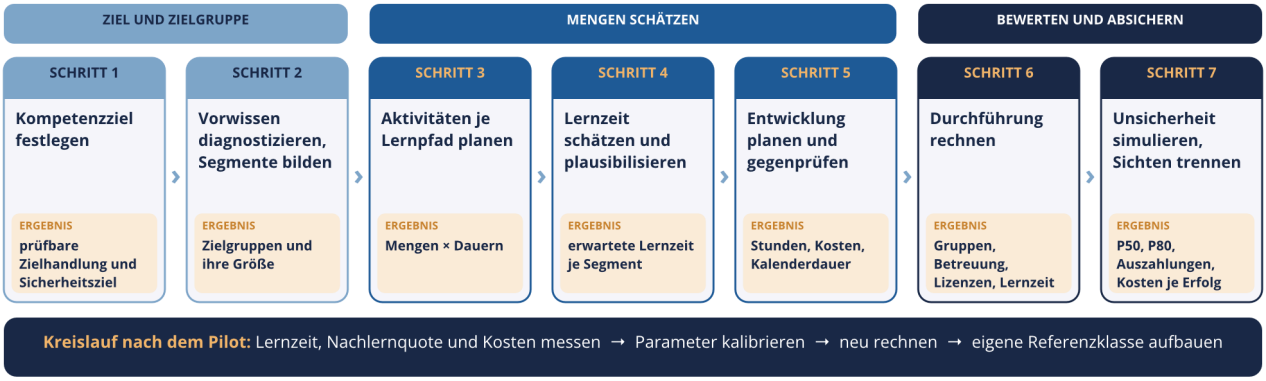
Disziplin	liefert	liefert *nicht*	Einsatz in der Kalkulation
Lernpsychologie	Vorwissenseffekte, Übungslogik, Vergessensraten	Minutenwerte für konkrete Ziele	Lernpfade, Nachlernquote, Auffrischung
Berufs- und Hochschulpädagogik	Workload-Prinzip, Niveaumodelle, Tabellenwerke	Taxonomie-Zeitfaktoren	Aktivitätenliste, Plausibilisierung
Instructional Design	Entwicklungsverhältnisse, Aufgabenanteile	Lernzeit der Teilnehmenden	Entwicklungsaufwand (Gegenprobe)
Bildungsökonomie	Bewertungsregeln, Kostensichten	universelle Kostensätze	Kostenmodell, Berichtsformat
Prognoseforschung	Korrektur von Optimismus, Bandbreiten	die „wahre“ Zahl	Szenarien, P80-Budget

Eigene Darstellung.

4. Das Rechenmodell: Von der Lernzieltiefe zum Budget

Aus den Bausteinen lässt sich ein Rechenmodell in sieben Schritten ableiten. Es ist eine eigene, anwendungsorientierte Synthese, keine aus einer einzelnen Studie übernommene Standardformel.

Abbildung 2: Die Rechenarchitektur in sieben Schritten



Eigene Darstellung. Nach dem Pilotdurchgang beginnt der Kreislauf mit kalibrierten Werten erneut (Kapitel 6).

4.1 Ziel festlegen und Zielgruppe segmentieren

Am Anfang steht das prüfbare Kompetenzziel. Wenn möglich, wird es aus einer Anforderungsanalyse der Zielrolle abgeleitet (vgl. Fachbeitrag Nr. 37). Danach folgt eine kurze, aufgabenbezogene Eingangsdiagnose, aus der Vorwissenssegmente entstehen, etwa „Einsteiger", „Fortgeschrittene" und „Vorerfahrene". Die Diagnose kostet selbst Zeit. Sie lohnt sich, sobald die eingesparte Lernzeit der Vorerfahrenen ihren Aufwand übersteigt, und das ist meist schon bei kleinen Gruppen der Fall (Kapitel 5.5).

4.2 Lernzeit aktivitätenbasiert schätzen

Die Lernzeit eines Segments *g* ist die Summe seiner Aktivitäten, jeweils Menge mal Dauer:

$$\hat{T}_g = \Sigma_j \, x_{gj} \cdot \hat{d}_{gj}$$

Typische Aktivitäten sind Eingangsdiagnose, Lesen oder Video einschließlich Verständnisarbeit, Fallbearbeitung, Übung bis zum Kriterium, Feedback und Reflexion, Prüfung, verteilte Abrufübungen, Transferaufgaben und Organisation. Für das erstmalige Lesen gibt die Meta-Analyse von Brysbaert (2019) einen Anhaltspunkt von durchschnittlich rund 238 Wörtern pro Minute. Das gilt für englische Sachtexte und Erwachsene und ist keine Norm für anspruchsvolles Fachlernen; Verständnis- und Übungszeit kommen hinzu. Bei einem festen Kompetenzziel wird das Nachlernen als Erwartungswert ergänzt:

$$E[T_g] = T_{g,\text{Grundpfad}} + p_g \cdot T_{g,\text{Nachlernschleife}}$$

Dabei ist p_g der erwartete Anteil der Personen, die nach dem Grundpfad eine weitere Schleife benötigen. Soll die Kompetenz dauerhaft halten, kommen Auffrischungen hinzu (Kapitel 3.1). Die so gewonnene Schätzung wird anschließend an Referenzklassen plausibilisiert, also mit der „Außensicht" abgeglichen.

4.3 Entwicklungskosten: Dauer und Personalkosten

Die Entwicklung wird von unten nach oben geplant (*bottom-up*): Arbeitspakete mal Stunden je Rolle mal Kostensatz, zuzüglich Fremdleistungen.

$$K_{\text{Entwicklung}} = \Sigma_r \, h_r \cdot s_r + K_{\text{extern}}$$

Die Benchmarks aus Tabelle 3 dienen als Gegenprobe von oben (*top-down*): Entwicklungsstunden ≈ Entwicklungsverhältnis × fertige Lernstunden, jeweils mit der Spanne von „einfach" bis „aufwendig". Weicht die eigene Planung deutlich nach unten ab, muss man das begründen können, etwa mit vorhandenem Material oder erprobten Vorlagen. Die *Entwicklungsdauer* ergibt sich nicht aus den Stunden allein. Man teilt die Stunden durch die tatsächlich verfügbare Teamkapazität und addiert Wartezeiten für Reviews, Freigaben durch Fachbereich, Datenschutz und Betriebsrat sowie den Piloten. Bei mehrjährig genutzten Angeboten kommen Pflege und Aktualisierung hinzu.

4.4 Durchführungskosten: Lernzeit, Lerntiefe und Gruppengröße

Den größten Posten bildet in der Regel die bewertete Lernzeit:

$$K_{Lernende} = \Sigma_g \, n_g \cdot E[T_g] \cdot s_g$$

Synchrone Formate erzeugen *sprungfixe* Kosten. Bei n Personen und höchstens m Personen je Gruppe braucht man $G = \lceil n / m \rceil$ Gruppen. Eine zusätzliche Person kann eine ganze weitere Gruppe auslösen. Hinzu kommen Trainerinnen und Trainer samt Vor- und Nachbereitung, Lizenzen, Lernplattform, Betreuung, Prüfungen und Bewertung, die Zeit der Führungskräfte für Transfergespräche, Administration und gegebenenfalls Reisen. In kompakter Form gilt:

$$K(n) = F + G(n) \cdot K_{Gruppe} + \Sigma_g \, n_g \cdot (v_g + E[T_g] \cdot s_g + b_g)$$

F steht für die Fixkosten, v_g für variable Sachkosten je Person und b_g für die individuelle Betreuung.

4.5 Unsicherheit ausweisen und das Budget absichern

Jede wichtige Annahme erhält drei Werte, und die Szenarien müssen zueinander passen. Komplexere Inhalte erhöhen meist Entwicklungsaufwand, Lernzeit *und* Nachlernquote zugleich. Eine Monte-Carlo-Simulation liefert daraus P50 und P80. Zwei Feinheiten sind wichtig. Erstens ist das 80. Perzentil der *individuellen* Lernzeit, multipliziert mit der Personenzahl, nicht das 80. Perzentil der *Summe*. Zweitens zeigt eine einfache Sensitivitätsrechnung, wo genaueres Wissen am meisten wert ist:

$$\partial K / \partial T_g = n_g \cdot s_g$$

Bei 48 Personen und 75 Euro je Stunde kostet jede zusätzliche Lernstunde je Person 3.600 Euro. Eine genauere Lernzeitschätzung ist deshalb meist wertvoller als eine harte Verhandlung über eine kleine Lizenzposition.

Kasten 2: Die Formeln auf einen Blick

Symbol	Bedeutung	Einheit
$x_{gj} \cdot d_{gj}$	Menge × Dauer der Aktivität *j* im Segment *g*	Stunden
p_g	Anteil mit Nachlernschleife	0–1
$E[T_g]$	erwartete Lernzeit je Person	Stunden
n_g	Personen im Segment	Anzahl
s_g, s_r	Kostensatz Lernende bzw. Rolle *r*	€/Stunde
$G = \lceil n/m \rceil$	Zahl der Gruppen bei max. *m* Personen	Anzahl
F, v_g, b_g	Fixkosten; variable Sachkosten; Betreuung	€ bzw. €/Person

Eigene Darstellung. Jeder Parameter erhält einen Herkunftshinweis: Messung, Referenzklasse oder Annahme.

Kasten 3: Sieben typische Rechenfehler

1. Chapman-Verhältnisse als Lernzeit der Teilnehmenden lesen.
2. Angebotsdauer statt Lernaufwand ansetzen und Selbststudium und Transfer vergessen.
3. Einmalige Entwicklungsstunden mit der Zahl der Teilnehmenden multiplizieren.
4. Personalkosten *und* vollen Produktionsausfall für dieselbe Stunde addieren.
5. Nachlernen zugleich als Pauschalaufschlag und als eigene Aktivität zählen.
6. Abbrechende und Nichtbestehende aus den Kosten herausrechnen.
7. Nur „Kosten je Teilnehmendem“ statt auch „Kosten je erfolgreicher Person“ berichten.

Eigene Zusammenstellung.

5. Praxisfall Kessler: Die Kalkulation

Der folgende Fall zeigt die Rechenlogik mit fiktiven, aber realistischen Planannahmen. Er enthält keine erhobenen Unternehmensdaten oder Marktpreise.

5.1 Ziel, Lerntiefe und Zielgruppe

Die Personalleiterin übersetzt den Auftrag zusammen mit dem Vertriebsleiter in ein prüfbares Ziel. Jede Ingenieurin und jeder Ingenieur erstellt für drei neue, reale Kundenanfragen eigenständig ein KI-gestütztes Angebot einschließlich technischer Auslegung. Die kritischen Prüfschritte müssen dabei vollständig korrekt sein: technische Plausibilität, Normbezüge, Vertraulichkeit der Kundendaten und Kennzeichnung KI-generierter Inhalte. Acht Wochen später folgt eine Wiederholungsprüfung. Die angestrebte Lerntiefe reicht damit bis zum *Anwenden, Analysieren und Bewerten* im Sinne der Taxonomie, denn KI-Ergebnisse müssen fachlich beurteilt werden. Zugleich ist der rollenbezogene Teil der KI-Kompetenz nach Art. 4 KI-Verordnung abgedeckt. Als Sicherheitsziele legt die Geschäftsführung fest: Mindestens 90 Prozent bestehen spätestens nach einer Nachlernschleife, und das Budget wird auf P80 freigegeben. Eine 30-minütige Eingangsdiagnose ergibt drei Segmente: 20 Einsteigerinnen und Einsteiger, 20 Fortgeschrittene und 8 Vorerfahrene.

5.2 Lernzeit

Der Grundpfad ist ein Blended-Learning-Programm. Einsteigerinnen und Einsteiger bearbeiten vier interaktive E-Learning-Module (Laufzeit je 30 Minuten, Bearbeitung einschließlich Übungen rund eine Stunde). Sie besuchen einen Grundlagenworkshop (6 Stunden) und einen Anwendungsworkshop mit eigenen Fällen (6 Stunden), üben fünf Stunden an einer Fallbibliothek und erledigen drei Transferaufgaben im Tagesgeschäft mit Feedback. Fortgeschrittene überspringen den Grundlagenworkshop und Teile der Module. Vorerfahrene arbeiten nur die Module zu Recht und Prüfregeln, Fallübungen und Transferaufgaben durch.

Tabelle 6: Lernzeit nach Segment

Segment	Personen	Grundpfad je Person	Summe
Einsteigerinnen und Einsteiger	20	24 h	480 h
Fortgeschrittene	20	16 h	320 h
Vorerfahrene	8	8 h	64 h
Gemeinsam: Diagnose 0,5 h, Arbeitsprobe 2 h, Wiederholungsprüfung 1 h, Organisation 0,5 h	48	4 h	192 h
Nachlernen: 25 % × 4 h (Coaching und zweite Arbeitsprobe)	12	4 h	48 h
Summe	**48**	**Ø 23 h**	**1.104 h**

Fiktive Planannahmen. Die Transferaufgaben zählen nur mit ihrer zusätzlichen Übungs- und Reflexionszeit, nicht mit der ohnehin geleisteten Angebotsarbeit.

Der interne Kostensatz der Zielgruppe beträgt 75 Euro je Stunde (rund 112.500 Euro jährliche Arbeitgeberpersonalkosten ÷ 1.500 verfügbare Stunden). Die bewertete Lernzeit kostet damit 82.800 Euro. Zur Plausibilisierung dient die IW-Weiterbildungserhebung: Durchschnittlich 20 Weiterbildungsstunden je Beschäftigtem und Jahr für *alle* Themen zeigen, dass ein einzelnes, tiefes Kompetenzziel einen Jahresdurchschnitt allein beanspruchen kann. Das muss man gegenüber dem Management offen sagen.

5.3 Entwicklung

Die Bottom-up-Planung ergibt 800 Personalstunden:

- 60 Stunden für Analyse, Zielformulierung und Arbeitsprobe
- 80 Stunden für das didaktische Design
- 220 Stunden für die E-Learning-Produktion durch eine Agentur
- 120 Stunden für Workshopkonzeption und -material
- 90 Stunden für eine Fallbibliothek mit 15 anonymisierten Angebotsfällen
- 40 Stunden für das Prüfungsdesign
- 50 Stunden für Reviews durch Fachexpertinnen und -experten
- 30 Stunden für Datenschutz, Betriebsrat und KI-Nutzungsrichtlinie
- 60 Stunden für die Pilotbegleitung
- 50 Stunden für das Projektmanagement

Bewertet mit 85 Euro je interner Stunde und 95 Euro je Agenturstunde, entstehen **70.200 Euro**. Die Gegenprobe mit Chapman ergibt für zwei Stunden interaktives E-Learning und zwölf Stunden Präsenztraining einen Korridor von 518 („einfach“) über 884 („typisch“) bis 1.518 Stunden („aufwendig“). Die Planung liegt also im plausiblen Bereich, leicht unter dem typischen Wert. Begründet wird das mit vorhandenen Produktunterlagen und KI-gestützter Texterstellung, deren Prüfaufwand eigens eingeplant ist.

Die **Entwicklungsdauer** beträgt rund 20 Wochen: sechs Wochen Analyse und Design, acht Wochen Produktion parallel zum Aufbau der Fallbibliothek, zwei Wochen Freigaben und vier Wochen Pilot mit acht Personen aus der Zielgruppe. Bei Start Mitte November 2026 beginnt der Rollout im April 2027. Bei drei nutzbaren Lernstunden pro Woche brauchen die Teilnehmenden rund acht Wochen. Die beiden Wellen erreichen die Arbeitsprobe bis Ende Juni. Die Wiederholungsprüfung fällt in den August, und das wird der Geschäftsführung offen mitgeteilt.

5.4 Durchführung

Tabelle 7: Durchführungskosten

Position	Ansatz	Kosten
Lernzeit der Teilnehmenden	1.104 h × 75 €	82.800 €
Externe Trainerin (Workshops)	2 + 4 Gruppen à max. 12 Personen = 6 Gruppentage × 1.800 €	10.800 €
Interne Co-Trainer (Fachexperten)	36 h × 85 €	3.060 €
Raum und Verpflegung	6 × 150 €	900 €
Nachlern-Coaching	12 Personen × 1,5 h × 85 €	1.530 €
Transferbegleitung durch Führungskräfte	48 × 1,5 h × 95 €	6.840 €
Bewertung der Arbeitsproben	54 h × 85 €	4.590 €
KI-Lernumgebung (Lizenzen Lernphase)	48 × 30 € × 6 Monate	8.640 €
Plattform, Administration, Evaluation	Pauschale, 30 h × 60 €, 20 h × 85 €	4.700 €
Summe Durchführung		**123.860 €**

Fiktive Planannahmen. Die Zeit der Lernenden in Transfergesprächen ist bereits in deren Lernzeit enthalten; die Zeit der Führungskräfte ist eine zweite, eigene Ressource.

5.5 Ergebnis für die Geschäftsführung

Im Basisszenario kostet das Vorhaben **194.060 Euro**: 70.200 Euro Entwicklung und 123.860 Euro Durchführung. Das sind rund 4.040 Euro je teilnehmender Person. Erreichen wie geplant 44 von 48 Personen das Ziel, sind es rund **4.410 Euro je erfolgreicher Person**. Nur rund 42.400 Euro davon sind zusätzliche Auszahlungen (Agentur, Trainerin, Lizenzen, Verpflegung, Plattform). Der Rest bewertet vorhandene interne Kapazität. Der Anteil der Lernzeit an den Gesamtkosten liegt hier bei 43 Prozent, also deutlich unter dem CVTS-Durchschnitt von 57 Prozent. Der Grund: Bei kleinen Zielgruppen fällt die Entwicklung stärker ins Gewicht.

Tabelle 8: Szenarien und Simulation

Kennzahl	günstig	Basis	ungünstig
Grundpfade (Faktor)	0,85	1,00	1,30
Nachlernquote	15 %	25 %	40 %
Entwicklungsstunden	700	800	1.000
Lernstunden gesamt	955 h	1.104 h	1.392 h
Gesamtkosten	**173.800 €**	**194.100 €**	**233.600 €**
Monte-Carlo (20.000 Läufe): P50 / P80 / P90		203.300 / 212.900 / 218.300 €	

Eigene Berechnung mit fiktiven Annahmen. Szenarien sind Bedingungen, keine Wahrscheinlichkeiten; erst die Simulation liefert Quantile.

Ein Befund überrascht Geschäftsführungen regelmäßig: Der Median der Simulation (P50) liegt über dem Basisszenario. Der Grund ist, dass die Risiken asymmetrisch sind. Lernzeit und Nachlernquote können stärker nach oben als nach unten abweichen. Das „wahrscheinlichste" Szenario ist daher nicht das mittlere Budget. Die Personalleiterin beantragt deshalb **213.000 Euro (P80)** und weist 194.000 Euro als Basiswert aus.

Zwei Vergleichsrechnungen machen den Nutzen des Verfahrens sichtbar. Ohne Eingangsdiagnose, wenn also alle 48 Personen den Einsteigerpfad durchlaufen, stiegen die Lernstunden um 288 und die Kosten um 21.600 Euro. Und die Förderung? Kessler fiele mit 520 Beschäftigten in die Stufe mit 25 Prozent Zuschuss. Mit 23 Stunden liegt die Maßnahme jedoch weit unter der Mindestdauer von 120 Stunden, und ihr KI-VO-Pflichtanteil wäre ohnehin ausgeschlossen. Ein Förderantrag lohnt sich hier nicht. Das sollte man früh wissen, statt es spät zu erfahren.

6. Pilotierung und Kalibrierung: Aus der Prognose wird ein Lernsystem

Die Vorab-Kalkulation ist eine Hypothese. Ein gut angelegter Pilotdurchgang prüft sie, bevor das große Budget gebunden ist. Er bildet die Segmente der Zielgruppe ab, läuft in der späteren Arbeitsumgebung und prüft die Zielerreichung an neuen Aufgaben ohne unzulässige Hilfen. Gemessen wird die *aktive* Lernzeit. Ein offener Browser-Tab ist keine Lernzeit. Bewährt hat sich eine Kombination aus kurzem Lerntagebuch und Plattformdaten mit einer vorab festgelegten Inaktivitätsregel. Die Erfassung bleibt zweckgebunden und wird aggregiert ausgewertet. Eine technische Zeiterfassung berührt das Mitbestimmungsrecht nach § 87 Abs. 1 Nr. 6 BetrVG und den Grundsatz der Zweckbindung nach Art. 5 Abs. 1 lit. b DSGVO (vgl. Fachbeitrag Nr. 54, Kapitel 10).

Ob die Prognose taugte, zeigen wenige Kennzahlen. Der *mittlere signierte Fehler* zeigt, ob systematisch zu niedrig geschätzt wurde. Der *mittlere absolute Fehler* zeigt die typische Abweichung in Stunden. Die Fehler je Segment decken verdeckte Verzerrungen auf. Die *Trefferquote der Bandbreiten* zeigt, ob die angegebene Unsicherheit realistisch war. Zielerreichung und Transfer verhindern, dass eine kürzere Lernzeit allein als Erfolg gilt, denn ein kurzer Kurs kann einfach anspruchsvolle Übungen weggelassen

haben. Plan-Ist-Abweichungen werden nach Ursachen zerlegt: mehr Teilnehmende, anderer Segmentmix, längere Aufgaben, höhere Kostensätze, mehr Nachlernen oder geänderte Inhalte. Personen, die abbrechen oder das Ziel verfehlen, bleiben in der Kostenrechnung. Andernfalls erscheint die Maßnahme günstiger, als sie war.

Mit jeder durchgeführten Maßnahme wächst die eigene Referenzklasse. Bewährt hat sich ein Stufenmodell:

1. ein transparenter Plan mit drei Szenarien,
2. die Kalibrierung nach dem Piloten,
3. eine kleine Vergleichsdatenbank eigener Maßnahmen nach Ziel, Zielgruppe, Format und Aufwand,
4. erst bei ausreichender Datenbasis statistische Prognosemodelle wie Lernkurven oder Knowledge Tracing (Corbett & Anderson, 1994).

Diese Reihenfolge verhindert, dass ein technisch ausgefeilter Rechner auf unbelegten Faktoren beruht.

Praxisfall Kessler (2): Der Pilot

Acht Personen aus allen drei Segmenten durchlaufen das Programm im März 2027. Die Einsteigerinnen und Einsteiger brauchen für die Fallübungen 30 Prozent länger als geplant. Die Vorerfahrenen kommen mit sechs statt acht Stunden aus. Nur einer von acht benötigt eine Nachlernschleife. Die Fallbibliothek wird um fünf Fälle erweitert (20 Stunden zusätzlich). Unter dem Strich bleibt der Basiswert mit rund 194.800 Euro nahezu gleich. Weil die Unsicherheit über die Nachlernquote gesunken ist, fällt das P80 jedoch auf rund 209.400 Euro, und der Puffer zum beantragten Budget wächst. Der Termin der zweiten Welle verschiebt sich um drei Wochen; die Geschäftsführung akzeptiert das.

7. Digitale und KI-gestützte Lernangebote: Was sich an der Kalkulation ändert

Digitale Formate verändern Zugang, Wiederverwendbarkeit, Terminabhängigkeit und Messbarkeit. Die Lernzeit bestimmen sie aber nicht automatisch. Die Meta-Analyse von Sitzmann, Kraiger, Stewart und Wisher (2006) zeigt: Unterschiede zwischen webbasierter und klassischer Instruktion hängen stark davon ab, wie viel Übung, Feedback und Steuerung das jeweilige Format bietet. Ein pauschaler Zeitersparnisfaktor für E-Learning ist daher nicht haltbar. Bei großen Zielgruppen können kleine Unterschiede pro Kopf wirtschaftlich bedeutsam werden, bei kleinen dominieren Entwicklung und Pflege.

KI-Tutoren zeigen in kontrollierten Studien beachtliche Effekte. Das gilt vor allem, wenn sie didaktisch sorgfältig gestaltet sind (Kestin et al., 2025). Ein Feldexperiment mit fast tausend Schülerinnen und Schülern zeigt aber auch die Kehrseite: Ungeschützter KI-Einsatz verbesserte die Leistung während des Übens, verschlechterte jedoch die spätere Leistung ohne KI. Didaktische Schutzmechanismen milderten diesen Effekt deutlich (Bastani et al., 2025). Daraus folgt eine einfache, aber oft übersehene Planungsregel: *Vorab festlegen, ob die Zielhandlung später mit oder ohne KI beherrscht werden soll, und genau unter diesen Bedingungen prüfen.* Bei Kessler ist es ausdrücklich das Arbeiten *mit* KI. Die Fähigkeit, KI-Ergebnisse *ohne* KI fachlich zu prüfen, wird deshalb eigens abgefragt.

Auf der Kostenseite kommen bei KI-Angeboten Einrichtung, didaktische Anpassung, fachliche Qualitätssicherung, laufende Lizenzen, Pflege bei Modellwechseln und Governance hinzu. Eine kürzere Erstproduktion kann durch zusätzlichen Prüfaufwand teilweise aufgezehrt werden. Beides wird als eigenes Arbeitspaket geführt, nicht als Abschlag versteckt.

8. Faktencheck: Verbreitete Faustregeln auf dem Prüfstand

Tabelle 9: Behauptungen und Befunde im Vergleich

Verbreitete Behauptung	Befund der Forschung
„Nach Chapman braucht man für eine Lernstunde 79 Stunden."	Die 79 Stunden sind *Entwicklungs*stunden je fertiger Stunde E-Learning der einfachsten Stufe (Stand 2010), nicht die Lernzeit der Teilnehmenden (Chapman, 2010).
„Bloom-Stufen ergeben Zeitfaktoren 1 – 2 – 3."	Die Taxonomie klassifiziert Ziele und Aufgaben; empirische Zeitfaktoren folgen daraus nicht (Krathwohl, 2002).
„Nach ECTS gilt: Selbststudium = Präsenz × 1,5 bis × 3."	Einen solchen Standard gibt es nicht; der Leitfaden nennt 25–30 Stunden Gesamtworkload je Leistungspunkt und verlangt dessen Überprüfung (Europäische Kommission, 2015).
„Arbeitsausfall macht 60–80 Prozent der Kosten aus."	In der CVTS6 waren es 57 Prozent (Statistisches Bundesamt, 2022); im Praxisfall 43 Prozent. Der Anteil ist projektspezifisch zu berechnen.
„Für Rüst- und Ablenkungszeiten pauschal 15–20 Prozent aufschlagen."	Aus der Cognitive Load Theory nicht ableitbar (Sweller et al., 2019); Organisation und Unterbrechungen messen oder als benanntes Szenario führen.
„Kirkpatrick verlangt viermal 15 Minuten Transfergespräche."	Kirkpatrick ist ein Evaluationsrahmen (Kirkpatrick & Kirkpatrick, 2016); Umfang der Transferbegleitung ist eine Gestaltungsentscheidung, deren Wirkung vom Arbeitsumfeld abhängt (Blume et al., 2010).
„70 Prozent lernt man informell, 10 Prozent im Training."	Die 70-20-10-Regel ist empirisch nicht belegt; ihr Ursprung sind Befragungen von Führungskräften (Clardy, 2018).
„10.000 Stunden machen zur Expertin."	Gezielte Übung erklärt in Berufen weniger als 1 Prozent der Leistungsunterschiede (Macnamara et al., 2014).
„E-Learning spart grundsätzlich 40–60 Prozent Lernzeit."	Kein genereller Formateffekt; entscheidend sind Übung, Feedback und Steuerung (Sitzmann et al., 2006).
„KI halbiert die Entwicklungszeit."	Bisher überwiegend Selbstauskünfte und Fallberichte; Prüf- und Freigabeaufwand wird selten mitgerechnet (Association for Talent Development, 2025b).
„Einmal gelernt ist gelernt."	Bei prozeduralen Fertigkeiten ist ohne Anwendung etwa die Hälfte des Lerngewinns nach 6,5 bis 13 Monaten verloren (Tatel & Ackerman, 2025).

Eigene Zusammenstellung verbreiteter Praxisfaustregeln; Bewertung auf Basis der genannten Quellen.

9. Leitfaden für den Mittelstand

Mittelständische Unternehmen kalkulieren unter besonderen Bedingungen. Die Zielgruppen sind klein, sodass die Entwicklung die Kosten dominiert. Es gibt kaum Vergleichsdaten. Fachexpertinnen und -experten sind knapp, und das Tagesgeschäft bestimmt, wie viel Lernzeit pro Woche tatsächlich verfügbar ist. Daraus folgt nicht, auf eine Kalkulation zu verzichten. Ihre Tiefe sollte aber zur Entscheidung passen.

Tabelle 10: Entscheidungsmatrix – welche Kalkulationstiefe passt?

Situation	Empfohlene Kalkulationstiefe	Hinweis zu Kaufen oder Selbst entwickeln
weniger als 15 Personen, Standardthema	Angebotsvergleich und Lernzeit in drei Szenarien; Anbieterwerte als Referenzklasse	einkaufen; Eigenentwicklung lohnt selten
15 bis 80 Personen, wettbewerbsrelevantes Thema	vollständige Aktivitätenkalkulation, drei Szenarien, Pilot	Mischform: eingekaufte Grundlagen, eigene Fallbibliothek und Arbeitsprobe
mehr als 80 Personen oder mehrere Wellen	zusätzlich Simulation, Gewinnschwelle, Lernkurveneffekt	Eigenentwicklung prüfen
Pflichtschulung (z. B. KI-Kompetenz nach Art. 4 KI-VO)	rollenbezogenes Mindestziel und Nachweis	keine Förderung nach § 82 SGB III; Aufwand schlank halten
Kompetenz muss dauerhaft gesichert sein	Auffrischung alle 6 bis 12 Monate einplanen	Lizenz- und Pflegekosten über den Lebenszyklus rechnen

Eigene Ableitung aus den in diesem Beitrag dargestellten Befunden.

Bei der Frage „Kaufen oder selbst entwickeln?“ hilft die Gewinnschwelle. Für zwei Alternativen A und B mit Fixkosten F und Kosten je Person $v + T \cdot s$ gilt:

$$n^* = (F_B - F_A) / [(v_A + T_A \cdot s) - (v_B + T_B \cdot s)]$$

Im Kessler-Fall hätte die Alternative A, ein offenes Zweitagesseminar mit interner Nacharbeitung, Fixkosten von 8.000 Euro und je Person 3.390 Euro verursacht (Gebühr und Reise 1.740 Euro, 22 Lernstunden). Die Eigenentwicklung B hat Fixkosten von 89.660 Euro und je Person 2.175 Euro. B wird damit erst ab 68 Personen günstiger; bei 48 Personen kostet A 170.720 Euro, B 194.060 Euro. Der reine Kostenvergleich führt hier jedoch in die Irre, wenn beide Alternativen nicht dasselbe erreichen. Unterstellt man für A nur 70 Prozent Zielerreichung in derselben Arbeitsprobe, weil Fälle aus dem eigenen Haus fehlen, kehrt sich das Bild um: rund 5.020 Euro je erfolgreicher Person bei A gegenüber 4.410 Euro bei B. Die Mehrkosten von B betragen rund 2.330 Euro je *zusätzlicher* erfolgreicher Person. Diese Kennzahl, das inkrementelle Kosten-Wirksamkeits-Verhältnis, ist für die Geschäftsführung aussagekräftiger als jeder Stundensatz (Levin et al., 2018). Erst wenn der Nutzen einer zusätzlich erfolgreichen Person verlässlich in Euro beziffert werden kann, ist eine Rendite-Rechnung (ROI) sinnvoll. Die Zurechnung bleibt dann die eigentliche Schwierigkeit (Salas et al., 2012).

Praxisfall Kessler (3): Bilanz nach neun Monaten

Die Ist-Kosten liegen bei 201.500 Euro, also 3,8 Prozent über dem Basiswert und unter dem P80-Budget. 44 von 48 Personen bestehen die Arbeitsprobe, 39 davon im ersten Anlauf. Die Einsteigerinnen und Einsteiger brauchten im Mittel 26 statt 24 Stunden. Die mittlere Durchlaufzeit eines technischen Angebots sinkt von zehn auf vier Arbeitstage. Die Wiederholungsprüfung zeigt Schwächen bei selten genutzten Normprüfungen. Kessler führt deshalb eine einstündige Quartalsauffrischung ein, die jährlich rund 14.400 Euro bewertete Lernzeit kostet, und nimmt sie in das Folgebudget auf. Alle Plan- und Ist-Werte werden als erster Datensatz der eigenen Referenzklasse gespeichert. Die nächste Kalkulation, KI im Einkauf, beginnt bereits mit der „Außensicht“.

10. Fazit und Ausblick: Sieben Thesen zur A-priori-Kalkulation

Was kostet Kompetenz, bevor sie entsteht? Die Antwort der Forschung lautet: *Man kann es wissen*, aber nicht aus einer Tabelle, sondern aus einem nachvollziehbaren Prognoseprozess.

1. **Lernzeit ist berechenbar, aber nur bedingt.** Sie gilt für eine Zielgruppe, eine Aufgabe und ein Kriterium. Eine Zahl ohne diese drei Angaben ist nicht überprüfbar.
2. **Das Kompetenzziel ist die wichtigste Stellschraube.** Ohne prüfbares Ziel gibt es keine seriöse Kalkulation, nur eine Angebotsdauer.
3. **Vorwissen schlägt Format.** Eine kurze Eingangsdiagnose mit segmentierten Lernpfaden spart mehr als die Wahl des Mediums.
4. **Entwicklung und Durchführung sind zwei Rechnungen.** Benchmarks wie Chapman helfen bei der ersten, nicht bei der zweiten.
5. **Die Lernzeit der Beschäftigten ist der größte Kostenhebel.** Darum lohnt es sich, sie im Piloten zu messen.
6. **„Hohe Sicherheit" braucht Bandbreiten.** Das Budget wird auf P80 freigegeben statt auf eine Punktschätzung, und gemessen wird die Zielerreichungsquote statt der Teilnahmequote.
7. **Wer nicht nachmisst, schätzt beim nächsten Mal wieder falsch.** Eigene Referenzklassen sind der eigentliche Wettbewerbsvorteil einer lernenden Personalentwicklung.

Der Ausblick ist zwiespältig. Lernanalytik und KI-Tutoren werden Lernzeit genauer messbar und Lernpfade individueller machen. Damit werden Vorab-Kalkulationen präziser, und Lernen, das bisher als Kostenblock erschien, wird als Investition mit messbarem Ergebnis steuerbar. Zugleich stellen sich neue Fragen: Wie viel Messung ist mit Datenschutz und Mitbestimmung vereinbar? Und welche Kompetenz soll ein Mensch künftig ohne, welche mit KI beherrschen? Die Kalkulation wird damit vom einmaligen Budgetakt zu einem laufenden Steuerungsinstrument der Personalentwicklung. Ob ein Unternehmen das ernst meint, zeigt sich nicht an der ersten Schätzung, sondern daran, ob es die zweite besser macht.

Kasten 4: Checkliste in zwölf Schritten

1. Entscheidungsfrage klären: Was kostet ein Programm, eine Kompetenz oder was ist im Budget erreichbar?
2. Kompetenzziel prüfbar formulieren: Handlung, Bedingungen, Maßstab, Zeitpunkt.
3. Lernzieltiefe mit Taxonomie und Niveaumodell verankern; Sicherheitsziel festlegen.
4. Vorwissen diagnostizieren und Segmente bilden.
5. Aktivitäten je Lernpfad planen; Mengen mal Dauern schätzen; Nachlernen und Auffrischung einrechnen.
6. Schätzung an Referenzklassen plausibilisieren: Tabellenwerke, Anbieterwerte, eigene Vorprojekte.
7. Entwicklung bottom-up planen, mit Benchmarks gegenprüfen und die Kalenderdauer ableiten.
8. Durchführung mit Gruppen-, Trainer-, Lizenz-, Prüfungs- und Transferkosten rechnen.
9. Kostensätze sauber bilden; Auszahlungen, interne Ressourcen und Opportunitäten getrennt ausweisen.
10. Szenarien und Simulation rechnen; Budget auf P80 und Ziel als Erfolgsquote freigeben.
11. Förderung, Mitbestimmung und Datenschutz früh prüfen.
12. Pilot messen, Prognose kalibrieren, Ist-Werte als eigene Referenzklasse sichern.

Literaturverzeichnis

Primärstudien, die nur nach Sekundärquellen genannt werden, sind nicht gesondert verzeichnet. Webquellen wurden am 3. November 2026 abgerufen.

Arthur, W., Jr., Bennett, W., Jr., Stanush, P. L., & McNelly, T. L. (1998). Factors that influence skill decay and retention: A quantitative review and analysis. *Human Performance, 11*(1), 57–101. https://doi.org/10.1207/s15327043hup1101_3

Association for Talent Development. (2025a). *Benchmarks and trends from the 2025 State of the Industry report* [Blogbeitrag]. https://www.td.org/content/atd-blog/benchmarks-and-trends-from-the-2025-state-of-the-industry-report

Association for Talent Development. (2025b). *ATD research: AI tools are benefitting instructional designers* [Pressemitteilung]. https://www.td.org/content/press-release/atd-research-ai-tools-are-benefitting-instructional-designers

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences, 122*(26), Artikel e2422633122. https://doi.org/10.1073/pnas.2422633122

Bloom, B. S. (1968). Learning for mastery. *Evaluation Comment, 1*(2), 1–12.

Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010). Transfer of training: A meta-analytic review. *Journal of Management, 36*(4), 1065–1105. https://doi.org/10.1177/0149206309352880

Brysbaert, M. (2019). How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of Memory and Language, 109*, Artikel 104047. https://doi.org/10.1016/j.jml.2019.104047

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology, 67*(3), 366–381. https://doi.org/10.1037/0022-3514.67.3.366

Bundesagentur für Arbeit. (2024). *Förderübersicht Beschäftigtenqualifizierung 2024* [Übersicht]. https://www.arbeitsagentur.de/vor-ort/datei/foerderuebersichtbeschaeftigtenqualifizierung2024_ba186764.pdf

Bundesagentur für Arbeit. (2026). *Bundesdurchschnittskostensätze für Maßnahmen der beruflichen Weiterbildung (B-DKS), gültig ab 1. Juli 2026* [Tabelle].

Cambridge Assessment English. (o. J.). *Guided learning hours*. Abgerufen am 3. November 2026, von https://support.cambridgeenglish.org/hc/en-gb/articles/202838506-Guided-learning-hours

Carroll, J. B. (1963). A model of school learning. *Teachers College Record, 64*(8), 723–733. https://doi.org/10.1177/016146816306400801

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354–380. https://doi.org/10.1037/0033-2909.132.3.354

Chapman, B. (2010). *How long does it take to create learning?* [Forschungsstudie]. Chapman Alliance.

Clardy, A. (2018). 70-20-10 and the dominance of informal learning: A fact in search of evidence. *Human Resource Development Review, 17*(2), 153–178. https://doi.org/10.1177/1534484318759399

Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4*(4), 253–278. https://doi.org/10.1007/BF01099821

Defelice, R. A. (2021, 13. Januar). *How long does it take to develop training? New question, new answers* [Blogbeitrag]. Association for Talent Development. https://www.td.org/content/atd-blog/how-long-does-it-take-to-develop-training-new-question-new-answers

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363–406. https://doi.org/10.1037/0033-295X.100.3.363

Europäische Kommission. (2015). *ECTS users' guide 2015*. Amt für Veröffentlichungen der Europäischen Union. https://doi.org/10.2766/87192

Flyvbjerg, B. (2008). Curbing optimism bias and strategic misrepresentation in planning: Reference class forecasting in practice. *European Planning Studies, 16*(1), 3–21. https://doi.org/10.1080/09654310701747936

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review, 7*(2), 185–207. https://doi.org/10.3758/BF03212979

Hollands, F. M., Pratt-Williams, J., & Shand, R. (2021). *Cost analysis standards & guidelines 1.1*. Cost Analysis in Practice (CAP) Project.

Institute of Education Sciences. (2020). *Cost analysis: A starter kit* (IES 2020-001). U.S. Department of Education.

Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science, 39*(1), 17–31. https://doi.org/10.1287/mnsc.39.1.17

Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. *Scientific Reports, 15*, Artikel 17458. https://doi.org/10.1038/s41598-025-97652-6

Kirkpatrick, J. D., & Kirkpatrick, W. K. (2016). *Kirkpatrick's four levels of training evaluation*. ATD Press.

Koedinger, K. R., Carvalho, P. F., Liu, R., & McLaughlin, E. A. (2023). An astonishing regularity in student learning rate. *Proceedings of the National Academy of Sciences, 120*(13), Artikel e2221311120. https://doi.org/10.1073/pnas.2221311120

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice, 41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Kulik, C.-L. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research, 60*(2), 265–299. https://doi.org/10.3102/00346543060002265

Lee, H., Lichand, G., Barnard, C., Klotz, L., Thille, C., Kim, Y., & Domingue, B. W. (2026). *Revisiting the regularity of student learning rate: Sensitivity to which observations are included* [Preprint]. arXiv. https://arxiv.org/abs/2605.01690

Levin, H. M., McEwan, P. J., Belfield, C., Bowden, A. B., & Shand, R. (2018). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis* (3. Aufl.). SAGE.

Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science, 25*(8), 1608–1618. https://doi.org/10.1177/0956797614535810

Murre, J. M. J., & Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting curve. *PLOS ONE, 10*(7), Artikel e0120644. https://doi.org/10.1371/journal.pone.0120644

Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The science of training and development in organizations: What matters in practice. *Psychological Science in the Public Interest, 13*(2), 74–101. https://doi.org/10.1177/1529100612436661

Seyda, S., Köhne-Finster, S., Orange, F., & Schleiermacher, T. (2024). IW-Weiterbildungserhebung 2023: Investitionsvolumen auf Höchststand. *IW-Trends, 51*(2), 3–23. https://doi.org/10.2373/1864-810X.24-02-01

Sitzmann, T., Kraiger, K., Stewart, D., & Wisher, R. (2006). The comparative effectiveness of web-based and classroom instruction: A meta-analysis. *Personnel Psychology, 59*(3), 623–664. https://doi.org/10.1111/j.1744-6570.2006.00049.x

Sozialgesetzbuch (SGB) Drittes Buch (III) – Arbeitsförderung, § 82 Förderung beschäftigter Arbeitnehmerinnen und Arbeitnehmer. https://www.gesetze-im-internet.de/sgb_3/__82.html

Statistisches Bundesamt. (2022, August). *Gut drei Viertel der Unternehmen bieten berufliche Weiterbildung an* (Pressemitteilung Nr. 349). https://www.destatis.de/DE/Presse/Pressemitteilungen/2022/08/PD22_349_215.html

Statistisches Bundesamt. (2026, 29. April). [Pressemitteilung zu den Arbeitskosten je geleistete Stunde 2025] (Nr. 148).

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Tatel, C. E., & Ackerman, P. L. (2025). Procedural skill retention and decay: A meta-analytic review. *Psychological Bulletin, 151*(6), 696–736. https://doi.org/10.1037/bul0000481

van Merriënboer, J. J. G., & Kirschner, P. A. (2018). *Ten steps to complex learning: A systematic approach to four-component instructional design* (3. Aufl.). Routledge.

Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Verordnung über künstliche Intelligenz). *Amtsblatt der Europäischen Union, L*, 12.07.2024.

Vuorikari, R., Kluzer, S., & Punie, Y. (2022). *DigComp 2.2: The digital competence framework for citizens*. Amt für Veröffentlichungen der Europäischen Union. https://doi.org/10.2760/115376

Wright, T. P. (1936). Factors affecting the cost of airplanes. *Journal of the Aeronautical Sciences, 3*(4), 122–128. https://doi.org/10.2514/8.155

Hinweis: Dieser Beitrag stellt allgemeine Fachinformationen bereit und ersetzt keine individuelle Rechts-, Förder- oder Steuerberatung. Die Ausführungen zu Rechtsrahmen und Förderung geben den Stand bei Redaktionsschluss (November 2026) wieder. Die Kessler Fluidtechnik GmbH ist ein fiktives Fallbeispiel; alle Fallzahlen sind Planannahmen.

KI-Footer: Erstellt mit KI-Unterstützung. Geprüft und verantwortet von Dr. Walter Lieberei.