

# Der AKIH-Score

## Methodologische Grundlagen und Herleitung der Bewertungsformel

**Prof. Bernhard Strobl M.A.**

KPH Edith Stein | Dezember 2025

## Zusammenfassung

*Dieses Paper dokumentiert die methodologischen Grundlagen des AKIH-Scores. Es wird transparent dargelegt, wie die Formel entwickelt wurde, welche theoretischen Bezüge bestehen, welche Entscheidungen normativ getroffen wurden und welche empirische Validierung noch aussteht.*

## 1. Einleitung: Warum eine Formel?

Der AKIH-Score ist **kein empirisch validiertes Messinstrument**, sondern ein **normatives Bewertungswerkzeug**. Er basiert auf theoretischen Überlegungen und praktischen Erfahrungen – nicht auf statistischer Kalibrierung. Diese Transparenz ist selbst Teil des Qualitätsanspruchs.

Das Ziel: Abstrakte Qualitätsanforderungen (Transparenz, menschliche Kontrolle, Nachvollziehbarkeit) in eine anwendbare Struktur überführen, die Reflexion ermöglicht und Vergleichbarkeit schafft.

## 2. Theoretische Bezüge

### 2.1 Gütekriterien qualitativer Forschung

Die Qualitätsbewertung qualitativer Forschung unterscheidet sich fundamental von quantitativen Ansätzen. Lincoln und Guba (1985) etablierten mit dem Konzept der *Trustworthiness* vier alternative Kriterien: Credibility, Transferability, Dependability, Confirmability.

Diese Kriterien bilden den konzeptionellen Hintergrund. Sie wurden jedoch nicht 1:1 übernommen, sondern für KI-gestützte Forschung adaptiert und um neue Dimensionen ergänzt (KI-Transparenz, Reproduzierbarkeit).

### 2.2 Phasenmodelle qualitativer Forschung

Die Strukturierung in sieben Phasen orientiert sich an etablierten Darstellungen des Forschungsprozesses (vgl. Creswell & Poth, 2018; Mayring, 2015; Kuckartz, 2018). Der Prozess wird als iterativ verstanden, nicht als strikt linear.

### 3. Die Formel im Detail

$$\text{AKIH-Score} = \sum(w_i \times P_i) \times TI \times HV$$

#### 3.1 Komponente 1: Phasenqualität $P_i$

$P_i$  bezeichnet die Qualität der Phase  $i$ , bewertet auf einer Skala von 0-100. Jede Phase wird anhand spezifischer Kriterien bewertet (z.B. Quellvalidität in Phase 1, Kodierungsqualität in Phase 4).

#### 3.2 Komponente 2: Gewichtung $w_i$

$w_i$  ist die Gewichtung (weight) der Phase  $i$ . Die Summe aller Gewichtungen ergibt 1.0 (100%). Die Gewichtungen wurden normativ festgelegt basierend auf:

- Bedeutung der Phase für die Gesamtqualität
- Intensität der KI-Nutzung in dieser Phase
- Risikopotenzial für Qualitätsmängel

Phase	$w_i$	Begründung
1. Literaturrecherche	<b>0.15</b>	Fundament; hohes Halluzinationsrisiko bei KI
2. Forschungsdesign	<b>0.10</b>	Wichtig, aber meist vor KI-Nutzung abgeschlossen
3. Datenerhebung	<b>0.15</b>	Transkription oft KI-gestützt; Datenqualität kritisch
<b>4. Datenanalyse</b>	<b>0.20</b>	Kernstück; höchste KI-Interaktionsdichte
5. Interpretation	<b>0.15</b>	Genuin menschlich; KI unterstützt, ersetzt nicht
6. Schreiben	<b>0.15</b>	Verantwortung für jeden Satz; authentische Position
7. Dokumentation	<b>0.10</b>	Ermöglicht Nachvollziehbarkeit und Reproduktion
<b>Summe</b>	<b>1.00</b>	

#### 3.3 Komponente 3: Transparenz-Index $TI$

**TI** (Transparenz-Index) misst die Vollständigkeit der KI-Nutzungsdokumentation. Wertebereich: 0.5 bis 1.0

- 1.0 = Vollständige Dokumentation aller Prompts, Outputs, Entscheidungen
- 0.75 = Überwiegend dokumentiert mit Lücken
- 0.5 = Minimale oder fehlende Dokumentation

**Warum Multiplikator statt Addition?** Transparenz ist keine optionale Zusatzleistung, sondern Voraussetzung. Ein exzellenter Basis-Score ohne Transparenz verliert an Glaubwürdigkeit – daher die multiplikative Verknüpfung.

#### 3.4 Komponente 4: Human-Validierung $HV$

**HV** (Human-Validierung) misst den Anteil der menschlich geprüften KI-Outputs. Wertebereich: 0.5 bis 1.0

- 1.0 = Alle KI-Outputs wurden menschlich evaluiert
- 0.75 = Überwiegende Prüfung mit Stichproben
- 0.5 = Minimale oder keine systematische Prüfung

**Warum Untergrenze 0.5?** Ein Score von 0 würde die gesamte Arbeit entwerten, auch wenn die Phasenarbeit selbst gut war. Die Untergrenze von 0.5 wirkt dämpfend, nicht eliminierend.

## 4. Berechnungsbeispiel

Angenommen, eine Forschungsarbeit erreicht folgende Bewertungen:

Phase	$P_i$	$w_i$	$w_i \times P_i$
1. Literaturrecherche	85	0.15	12.75
2. Forschungsdesign	90	0.10	9.00
3. Datenerhebung	80	0.15	12.00
4. Datenanalyse	75	0.20	15.00
5. Interpretation	85	0.15	12.75
6. Schreiben	80	0.15	12.00
7. Dokumentation	70	0.10	7.00
<b>Basis-Score <math>\Sigma(w_i \times P_i)</math></b>			<b>80.50</b>

Multiplikatoren: TI = 0.90, HV = 0.85

AKIH-Score =  $80.50 \times 0.90 \times 0.85 = \mathbf{61.58}$

Interpretation: Akzeptabel (65-74 = Akzeptabel), aber Verbesserungspotenzial vorhanden – insbesondere bei Dokumentation und Human-Validierung.

## 5. Limitationen und offene Fragen

Der AKIH-Score hat bewusste Grenzen:

- 1. Keine empirische Validierung:** Die Gewichtungen basieren auf Plausibilität, nicht auf Daten.
- 2. Subjektivität der Bewertung:** Die Phasenscores erfordern Urteilsvermögen.
- 3. Kontextabhängigkeit:** Nicht alle Phasen sind in jedem Projekt gleich relevant.
- 4. Entwicklungsstand:** Version 1.0 – Weiterentwicklung basierend auf Praxiserfahrung geplant.

## 6. Literaturverzeichnis

- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE.
- Kuckartz, U. (2018). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung* (4. Aufl.). Beltz Juventa.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12. Aufl.). Beltz.

**Lizenzhinweis:** Dieses Werk steht unter CC BY 4.0. Zitiervorschlag: Strobl, B. (2025). Der AKIH-Score: Methodologische Grundlagen. KPH Edith Stein. <https://doi.org/10.5281/zenodo.17801547>