

Paper-Autopsie

Vom Forschungsartikel zur methodischen Diagnose: eine strukturierte Checkliste zur Studienkritik

Dr. Rüdiger-Philipp Rackwitz · philipp-rackwitz.de · Version 08/2026

Leitfrage: Was behauptet die Studie, was wurde tatsächlich gemacht und trägt die Datengrundlage die gezogene Schlussfolgerung?

Die Checkliste eignet sich für quantitative, qualitative und Mixed-Methods-Studien. **Nicht jeder Punkt ist für jede Studie gleichermaßen relevant.** Entscheidend ist nicht, möglichst viele Fehler zu finden, sondern die Logik der Studie nachvollziehbar zu rekonstruieren und ihre Aussagekraft angemessen zu beurteilen.

1. Erkenntnisanspruch und Forschungsfrage

- ☐ Welches Problem oder welche Forschungslücke wird behauptet?
- ☐ Ist die Forschungsfrage klar, empirisch bearbeitbar und aus Theorie bzw. Forschungsstand hergeleitet?
- ☐ Soll beschrieben, verglichen, ein Zusammenhang untersucht, erklärt, rekonstruiert oder eine Wirkung geprüft werden?
- ☐ Passt der spätere Schluss tatsächlich zu diesem ursprünglichen Erkenntnisanspruch?

2. Begriffe, Theorie und Operationalisierung

- ☐ Welche zentralen Begriffe oder Konstrukte tragen die Argumentation?
- ☐ Wie werden diese Begriffe definiert und theoretisch begründet?
- ☐ Wie werden sie empirisch beobachtbar gemacht: Variablen, Items, Kategorien, Indikatoren, Beobachtungseinheiten?
- ☐ Geht durch die Operationalisierung etwas Wesentliches verloren oder wird ein komplexer Begriff zu stark verkürzt?

3. Design und Untersuchungslogik

- ☐ Welches Studiendesign liegt tatsächlich vor?
- ☐ Passt das Design zur Forschungsfrage und zur behaupteten Reichweite?
- ☐ Welche Alternativerklärungen kann das Design reduzieren und welche bleiben bestehen?
- ☐ Werden Kausalität, Entwicklung oder Veränderung behauptet, obwohl das Design nur Unterschiede oder Zusammenhänge zeigt?

4. Stichprobe, Fälle und Feldzugang

- ☐ Welche Population bzw. welches Feld soll untersucht werden?
- ☐ Wie wurden Personen, Fälle, Dokumente oder Situationen ausgewählt?
- ☐ Welche Ein- und Ausschlusskriterien gelten?
- ☐ Welche Selektions- oder Ausfallprozesse könnten die Ergebnisse verzerren?
- ☐ Ist die Fallzahl bzw. Fallauswahl für die gewählte Analyse- und Aussageform plausibel begründet?

5. Datenerhebung und Datenqualität

- ☐ Wie wurden Daten erhoben und unter welchen Bedingungen?
- ☐ Sind Instrumente, Leitfäden, Kategorien oder Beobachtungsverfahren nachvollziehbar beschrieben?
- ☐ Welche Formen von Messfehler, Reaktivität, sozialer Erwünschtheit oder Forschendeneinfluss sind denkbar?
- ☐ Wie werden Reliabilität, Validität, Nachvollziehbarkeit, Reflexivität oder andere passende Gütekriterien behandelt?

6. Auswertung

- ☐ Ist der Auswertungsweg so beschrieben, dass die zentralen Entscheidungen nachvollziehbar werden?
- ☐ Passt die Analyse zum Datenniveau, Design und zur Forschungsfrage?
- ☐ Wie werden fehlende Daten, Ausreißer, Mehrfachtests, Modellannahmen oder uneindeutige Fälle behandelt?
- ☐ Wurde die Auswertung offenbar nach Sichtung der Ergebnisse verändert oder selektiv berichtet?

Methodenspezifischer Kurzcheck

Studententyp	Besonders zu prüfen
Quantitativ	Messqualität; Stichprobenlogik; Testvoraussetzungen; Effektgrößen; Konfidenzintervalle; statistische Unsicherheit; multiple Tests; Modellpassung; praktische Bedeutsamkeit
Qualitativ	Samplinglogik; Feldzugang; Forschendenrolle; Dokumentation des Kodier-/Interpretationswegs; Gegenbeispiele; Reflexivität; Nachvollziehbarkeit; Reichweite der Verallgemeinerung
Mixed Methods	Begründung der Kombination; zeitliche/logische Verknüpfung; Integrationspunkt; Umgang mit widersprüchlichen Befunden; Mehrwert gegenüber zwei unverbundenen Teilstudien

7. Ergebnisse und Bedeutsamkeit

- ☐ Welche Ergebnisse sind zentral und welche nur Nebenfunde?
- ☐ Werden statistische Signifikanz und inhaltliche bzw. praktische Bedeutsamkeit sauber getrennt?
- ☐ Sind Effektgrößen, Unsicherheiten oder Spannweiten erkennbar, soweit dies zum Design passt?
- ☐ Werden Nullbefunde, widersprüchliche Ergebnisse oder abweichende Fälle angemessen berücksichtigt?

8. Interpretation und Schlussfolgerungen

- ☐ Welche Schlussfolgerung ziehen die Autorinnen und Autoren aus den Daten?
- ☐ Ist diese Schlussfolgerung enger, gleich weit oder weiter als das Design erlaubt?
- ☐ Welche alternative Interpretation wäre ebenfalls mit den Daten vereinbar?
- ☐ Werden Korrelation und Kausalität, Gruppenunterschied und individuelle Entwicklung oder Stichprobe und Population verwechselt?

9. Limitationen, Transparenz und Interessenkonflikte

- ☐ Werden zentrale Grenzen der Studie benannt oder eher ritualisiert abgehandelt?
- ☐ Sind Daten, Materialien, Analyseentscheidungen oder Präregistrierung verfügbar bzw. transparent dokumentiert?
- ☐ Welche Finanzierungsquellen oder Interessenkonflikte werden genannt?
- ☐ Welche methodische Schwäche hätte bei anderem Ergebnis vermutlich mehr Aufmerksamkeit erhalten?

10. Gesamtdiagnose

- ☐ Was ist die stärkste methodische Seite der Studie?
- ☐ Was ist die wichtigste Grenze für die Interpretation?
- ☐ Welche Aussage ist durch die Studie gut gedeckt?
- ☐ Welche häufig naheliegende Aussage wäre eine Überinterpretation?
- ☐ Wie müsste eine Anschlussstudie aussehen, um die wichtigste offene Frage besser zu beantworten?

Kurzformel für die Studienkritik: Forschungsfrage → Operationalisierung → Design → Datengrundlage → Auswertung → Interpretation. An jeder Verbindung kann Aussagekraft gewonnen oder verloren werden.

Ein-Satz-Diagnose

Die Studie zeigt belastbar ...	
Sie zeigt nicht ...	
Das größte methodische Risiko ist ...	
Eine sinnvollere Anschlussstudie würde ...	

Ziel ist keine Fehlerjagd, sondern eine begründete Einschätzung der Reichweite empirischer Erkenntnis.