

Kriterienkatalog: Bewertung von OER- und KI-gestütztem Material zu BNE

Checkliste zur systematischen Bewertung von Bildungsmaterial zur Bildung für nachhaltige Entwicklung (BNE), das als Open Educational Resource (OER) veröffentlicht und/oder mit KI erstellt wurde. Die Kriterien dienen als Bewertungs- und Orientierungshilfe. **Es müssen nicht alle Kriterien erfüllt sein.** Kreuzen Sie die zutreffenden Kriterien an — die Analysespinne am Ende aktualisiert sich automatisch. Nicht jeder Bereich ist für jedes Material relevant: Mit dem Häkchen „Bereich einbeziehen“ im Kopf jedes Abschnitts können Sie einzelne Unterkategorien (A1, B3 etc.) gezielt aus der Bewertung und der Analysespinne herausnehmen.

A. Pädagogisch-didaktische Dimension

A1. Inhaltliche Dimension

☒ Bereich einbeziehen

- ☐ **Zielgruppenorientierung und Anwendungsbereich:**
Zielgruppe (Alter, Vorwissen, Kontext) und Einsatzbereich (Fach, Schulform, Setting) sind klar definiert und das Material ist darauf abgestimmt.
- ☐ **Fachliche Richtigkeit und Fundierung:**
Inhalte sind fachlich korrekt, aktuell und durch Quellen bzw. Fachpersonen belegt; bei KI-generierten Inhalten ist eine fachliche Prüfung erfolgt (siehe PRÜFE-Check, Abschnitt C3).
- ☐ **Vollständigkeit und Kohärenz:**
Das Thema wird in sich schlüssig und ohne inhaltliche Lücken oder Brüche behandelt.
- ☐ **Inhaltliche Verwendbarkeit:**
Inhalte sind so aufbereitet, dass sie direkt oder mit geringem Aufwand im Unterricht/in der Bildungsarbeit nutzbar sind.
- ☐ **Lehrplan- und Bildungsstandards:**
Es besteht ein erkennbarer Bezug zu curricularen Vorgaben und Bildungsstandards (z. B. BNE-Orientierungsrahmen, KMK-Vorgaben).

A2. Didaktische Dimension

☒ Bereich einbeziehen

- ☐ **Motivation/Vielfalt der Lernzugänge:**
Unterschiedliche Zugänge (visuell, handelnd, diskursiv, forschend) werden angesprochen.
- ☐ **Anwendung und Transfer:**

Es besteht die Möglichkeit zur Übertragung auf eigene Praxis bzw. reale Situationen.

☐ **Differenzierung und Support:**

Angebote in unterschiedlichen Niveaustufen, Erklärhilfen und Unterstützungsangebote sind vorhanden.

☐ **Kollaboration und Interaktion:**

Das Material ermöglicht Zusammenarbeit bzw. Austausch unter Lernenden.

☐ **Aufgabenstellung/Assessment und Feedback:**

Aufgaben sind klar formuliert, es gibt Möglichkeiten zur Selbst- bzw. Fremdüberprüfung und Feedback-Formate.

*Bereich A2.1 „Aufgabenstellung im KI-Zeitalter“ liegt in zwei Varianten vor — bitte **entweder** die KI-Aufgabenkultur nach Falck/Flick **oder** das Polmodell nach Barbara Geyer für die Bewertung einbeziehen.*

A2.1 Variante (1) — KI-Aufgabenkultur

☒ Diese Variante verwenden

Details zur Aufgabenstellung: Umsetzung der KI-Aufgabenkultur (nach Falck/Flick 2025). Zur Förderung von KI-Kompetenzen bei Lernenden siehe die Checkliste im Bereich C1.

☐ **KI-thematisierend:**

KI ist Lerngegenstand — Funktionsweise, Chancen/Risiken und gesellschaftlich-ethische Einordnung werden behandelt.

☐ **KI-integrierend:**

KI wird als Werkzeug eingesetzt; der Einsatz ist didaktisch begründet, an Lernziele gebunden und geregelt.

☐ **KI-reflektierend:**

Das Material regt an, KI-Ergebnisse zu prüfen, Verzerrungen (Bias) zu erkennen und eigene Nutzungsmotive zu hinterfragen.

☐ **KI-limitierend:**

An geeigneten Stellen wird bewusst auf KI verzichtet, um Kompetenzen (z. B. eigenständiges Argumentieren) sichtbar zu sichern.

☐ **Begründete Auswahl erkennbar:**

Es ist nachvollziehbar, warum diese Aufgabenart und dieser KI-Einsatzgrad für das Lernziel gewählt wurden.

A2.2 Variante (2) — Das Polmodell

☐ Diese Variante verwenden

Mit dem Polmodell nach Barbara Geyer (2026) lassen sich Aufgabenstellungen in drei Zonen einteilen. (1) Den **Präsenzpol** (Leistung wird direkt und beaufsichtigt überprüft), (2) den **integrierten Pol** (KI-Nutzung ist Voraussetzung und bewertet wird, was darüber hinausgeht) und (3) die dazwischenliegende „**gefährliche Mitte**“ — meist unbeaufsichtigte Fließtextaufgaben, bei denen der Entstehungsweg unsichtbar bleibt und ein schlüssiges Ergebnis kein Verstehen mehr beweist.

- ☐ **Leitfragen gestellt:**
Für die Aufgabenstellung wurde geprüft, ob die Leistung direkt (mündlich, praktisch oder unter Aufsicht) überprüft wird und ob eine reine, unbearbeitete KI-Abgabe erkennbar schlecht ausfallen würde.
- ☐ **Präsenzpol erfüllt (falls zutreffend):**
Die Leistung entsteht unter Aufsicht; KI kann höchstens zur Vorbereitung genutzt werden, nicht aber in der Überprüfungssituation selbst.
- ☐ **Integrierter Pol erfüllt (falls zutreffend):**
Die Nutzung von KI ist ausdrücklich vorausgesetzt und offengelegt; die KI-Ausgabe ist Rohmaterial, bewertet wird, was Lernende darüber hinaus leisten.
- ☐ **Gefährliche Mitte vermieden:**
Die Aufgabe ist keine unbeaufsichtigte, rein schriftliche Ausarbeitung, bei der weder Entstehungsweg noch tatsächliches Verständnis erkennbar sind.
- ☐ **Ausweg aus der Mitte gewählt (falls zutreffend):**
Liegt die Aufgabe dennoch in der Mitte, wird einer der drei Wege genutzt — 3P-Modell (Prozess, Produkt, Präsentation, nach Weißels 2025), Verlagerung zum integrierten Pol oder Verlagerung zum Präsenzpol.
- ☐ **Gestaltung statt Kontrolle:**
Der Fokus liegt auf der bewussten Gestaltung der Aufgabenstellung, nicht auf dem nachträglichen Versuch, KI-Nutzung zu erkennen oder zu verbieten.

B. Technische Dimension

B1. Zugänglichkeit

☒ Bereich einbeziehen

- ☐ **Barrierefreiheit:**
Material ist für Menschen mit unterschiedlichen Voraussetzungen (Sehen, Hören, Sprache, technische Ausstattung) zugänglich.
- ☐ **Technische Wiederverwendbarkeit/Anpassbarkeit:**
Material liegt in bearbeitbaren Formaten vor und kann angepasst werden.
- ☐ **Zuverlässigkeit und Kompatibilität (offene Dateiformate):**
Es werden offene, plattformunabhängige Dateiformate verwendet; Material funktioniert zuverlässig.

B2. Usability

☒ Bereich einbeziehen☐ **Struktur, Navigation und Orientierung:**

Aufbau ist übersichtlich, Navigation ist intuitiv.

☐ **Design und Lesbarkeit:**

Layout, Schriftgröße und Kontraste unterstützen die Lesbarkeit.

☐ **Interaktivität:**

Interaktive Elemente funktionieren zuverlässig und unterstützen das Lernen.

B3. Rechtliche Qualität

☒ Bereich einbeziehen☐ **Korrekte Quellenangaben:**

Alle verwendeten Quellen, Daten und Zitate sind vollständig und korrekt ausgewiesen.

☐ **Transparente Lizenzierung:**

Lizenz, Autorenschaft, Erstellungsdatum und Version sind eindeutig dokumentiert und passend zur beabsichtigten Weiterverwendung.

☐ **Keine Verwendung urheberrechtlich geschützter Inhalte:**

Es werden keine fremden, nicht lizenzierten Inhalte ohne Erlaubnis verwendet.

☐ **Kenntlichmachung urheberrechtlicher Einschränkungen:**

Einschränkungen, die sich z. B. aus den Nutzungsbedingungen eingesetzter KI-Tools ergeben, sind kenntlich gemacht.

☐ **KI-Einsatz ist transparent gekennzeichnet:**

Es ist erkennbar, ob, mit welchem Tool und in welchem Umfang KI bei der Erstellung beigetragen hat, sowie der Grad menschlicher Nachbearbeitung (vollständig KI-generiert / KI-unterstützt / KI-geprüft).

C. KI-Integration (wenn vorhanden und kenntlich gemacht)

C1. Geförderte KI-Kompetenzen

☒ Bereich einbeziehen

Prüft, ob und welche KI-Kompetenzen bei den Lernenden durch das Material aufgebaut werden — unabhängig davon, ob KI im Material selbst zum Einsatz kommt. Grundlage ist das gemeinsame KI-Kompetenzrahmenmodell von OECD und Europäischer Kommission für die schulische Bildung („Empowering Learners for the Age of AI“, 2026), das vier aufeinander aufbauende Domänen unterscheidet.

☐ **Mit KI bewusst umgehen:**

Material befähigt dazu, sich kritisch und verantwortungsvoll mit einer von KI geprägten Welt auseinanderzusetzen.

☐ **KI kreativ anwenden:**

Das Material regt dazu an, KI als kreativen Partner zu nutzen und dabei das selbstbestimmte menschliche Handeln zu bewahren.

☐ **KI gezielt einsetzen:**

Das Material motiviert die Haltung, Arbeit bewusst zwischen Menschen und KI-Systemen aufzuteilen und den KI-Einsatz gezielt zu steuern.

☐ **KI aktiv mitgestalten:**

Das Material regt dazu an, über die bloße Nutzung bestehender KI-Systeme hinauszugehen und die Funktionsweise dieser Systeme aktiv zu reflektieren und mitzugestalten.

C2. Niveau der Integration digitaler Technologien ☒ Bereich einbeziehen

Mit dem SAMR-Modell nach Ruben Puentedura (2006, Übersetzung Ekkehardt Brüggemann) lässt sich die Integrationstiefe von digitalen Technologien im didaktischen Kontext beurteilen.

Wichtig: Wenn R, M oder A angekreuzt werden, immer auch alle „niedrigeren“ Kategorien mitankreuzen.

☐ **S — Substitution, Ersatz:**

Technik ist direkter Ersatz für Arbeitsmittel ohne funktionale Änderung (z. B. Lesen von Texten auf digitalen Endgeräten).

☐ **A — Augmentation, Erweiterung:**

Technik ist direkter Ersatz für Arbeitsmittel, mit funktionaler Verbesserung (z. B. Online-Recherche statt Lexika).

☐ **M — Modification, Änderung:**

Technik ermöglicht beachtliche Neugestaltung von Aufgaben (nur mithilfe digitaler Technik bearbeitbar, z. B. Tabellenkalkulation).

☐ **R — Redefinition, Neubelegung:**

Technik ermöglicht das Erzeugen neuartiger Aufgaben, die zuvor unvorstellbar waren (z. B. Erstellung digitaler Endprodukte, Aufbereitung schwer verständlicher Inhalte).

Bereich C3 „Kritische Prüfung KI-generierter Inhalte“ liegt in zwei Varianten vor — bitte **entweder** das PRÜFE-Framework **oder** die Antwort-Autopsie für die Bewertung einbeziehen.

C3. Variante (1) — Kritische Prüfung KI-generierter Inhalte (PRÜFE)

☒ Diese Variante verwenden

Mit dem PRÜFE-Framework nach Barbara Geyer (2025):

☐ **P — Plausibilität:**

Die Aussagen, Zahlen und Struktur wirken beim ersten Lesen stimmig. Es gibt keine Widersprüche oder Auffälligkeiten.

- ☐ **R — Recherche:**
Zentrale Fakten/Daten wurden mit unabhängigen Quellen abgeglichen bzw. mit mehreren KI-Systemen quergeprüft.
- ☐ **Ü — Überzeugungen/Annahmen hinterfragt:**
Verdeckte Annahmen und mögliche Verzerrungen (kognitiv, strukturell, Präferenz) wurden identifiziert — z. B. einseitige Perspektiven auf Nachhaltigkeitsfragen.
- ☐ **F — Falsifizieren:**
Es wurde aktiv nach Gegenbeweisen/Gegenpositionen gesucht (nicht nur Bestätigung gesammelt).
- ☐ **E — Entscheiden:**
Es liegt eine dokumentierte, begründete Entscheidung vor (übernommen/angepasst/abgelehnt), idealerweise gegengeprüft durch eine zweite Person.

C3. Variante (2) — Kritische Prüfung KI-generierter Inhalte (Antwort-Autopsie)

- ☐ Diese Variante verwenden

Mit der Methode „Antwort-Autopsie“ nach Lucca Spohn (2026): „Eine KI-Antwort ist erst geprüft, wenn ich sagen kann, was sie behauptet, worauf sie sich stützt, was sie voraussetzt und was sie weglässt.“

- ☐ **Behauptungen:**
Es wurden Aussagen getroffen, die überprüfbar sind.
- ☐ **Belege:**
Es sind Quellen oder Nachweise angegeben, die die Aussagen stützen.
- ☐ **Annahmen:**
Den Aussagen liegen stille Annahmen zugrunde, es wird etwas als gegeben vorausgesetzt.
- ☐ **Lücken:**
Es fehlt etwas, wichtige Informationen oder Perspektiven sind nicht vorhanden.

D. BNE-fachliche Kriterien

D. BNE-fachliche Kriterien

☒ Bereich einbeziehen

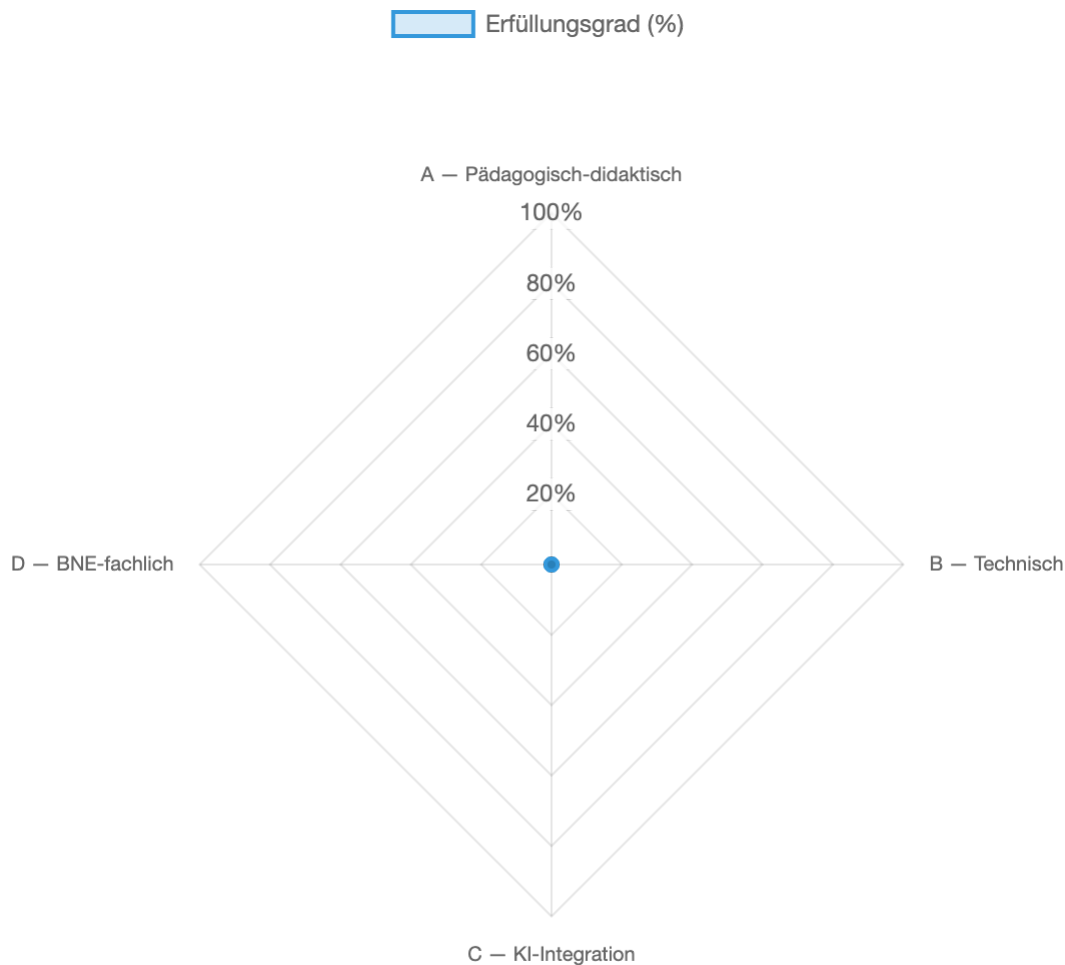
Für diesen Bereich wurden unterschiedliche Ansätze kombiniert, siehe Literaturverzeichnis (de Haan, 2008; Orientierungsrahmen Lernbereich Globale Entwicklung).

- ☐ **Bezug zu Nachhaltigkeitszielen (SDGs):**
Der SDG-Bezug ist erkennbar, exemplarische Fragestellungen sind relevant und mit Zukunftsbezug.

- ☐ **Gestaltungskompetenz im Fokus:**
Material fördert vorausschauendes Denken, interdisziplinäres Arbeiten und eigenständige/partizipative Lernprozesse.
- ☐ **Multiperspektivität:**
Ökologische, ökonomische und soziale Dimensionen von Nachhaltigkeit werden gemeinsam betrachtet; Interessenkonflikte werden thematisiert.
- ☐ **Handlungsorientierung:**
Reale Handlungsmöglichkeiten und Beteiligungsformen werden aufgezeigt, keine reine Problem- ohne Lösungsperspektive.
- ☐ **Globale und lokale Bezüge:**
Globale und lokale Perspektiven werden miteinander verknüpft (Globales Lernen).
- ☐ **Vermeidung von Vereinfachung/Alarmismus:**
Die Komplexität von Nachhaltigkeitsfragen wird angemessen abgebildet; Widersprüche, Dilemmata und offene Fragen werden didaktisch aufgegriffen.
- ☐ **Ökologischer Fußabdruck der KI-Nutzung:**
Ressourcen-/Energieverbrauch generativer KI wird, wo passend, als Teil des BNE-Themas mitgedacht bzw. thematisiert.

Ergebnisübersicht — Analysespinnne

Jede Achse zeigt den Erfüllungsgrad einer Hauptkategorie (A–D), gemittelt über die einbezogenen Unterkategorien. Ausgeschlossene Unterkategorien fließen nicht ein.



0 / 46 Kriterien

erfüllt (bezogen auf die einbezogenen Bereiche)

— noch keine Angaben —

Hinweis: Materialien, die bei Kriterien in Block **A1** (fachliche Richtigkeit) oder **B3** (rechtliche Qualität) schwach abschneiden, sollten vor dem Einsatz gezielt nachgebessert werden.



„Bewertung von OER- und KI-gestütztem Material zu BNE“ vom [ZfL Köln](#) auf Grundlage verschiedenster Quellen und mit KI-Unterstützung erstellt und steht unter der Lizenz [CC BY-SA 4.0](#). Bei Weitergabe oder Bearbeitung ist sowohl dieser Lizenzhinweis für die eigene Bearbeitung als auch die Einzelnachweise der verwendeten Quellen zu übernehmen.

Quellen

Becker, V., Schaper, F., Schulte-Buskase, A. & Tibbe, T. (2024): (Qualitäts-)Ansprüche an OER: Zur Entwicklung eines Kriterienkatalogs für OER im Kontext von inklusionsbezogener Lehrkräftebildung. In: *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*. Online verfügbar unter researchgate.net

de Haan, Gerhard (2008): Gestaltungskompetenz als Kompetenzkonzept der Bildung für nachhaltige Entwicklung. In: Bormann, Inka; de Haan, Gerhard (Hrsg.): *Kompetenzen der Bildung für nachhaltige Entwicklung. Operationalisierung, Messung, Rahmenbedingungen, Befunde*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 23–44.

Falck, J. & Flick, M. (2025): KI-Didaktik: Eine Planungsvorlage zur Aufgabenkultur im KI-Zeitalter. Online verfügbar unter joschafalck.de

Geyer, B. (2025): Nicht alles glauben, was KI sagt — Das PRÜFE-Framework für kritische KI-Bewertung. Online verfügbar unter barbarageyer.substack.com

Geyer, B. (2026): Das Polmodell — Welche Aufgabenstellungen in Zeiten von KI noch funktionieren. Online verfügbar unter barbarageyer.substack.com

Klein, A. (2026): KI, OER und OEP — für eine nachhaltige Openness. Learntec-Vortrag. Online verfügbar unter hnd-bw.de

Kultusministerkonferenz (KMK) & Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) (Hrsg.) (2016): *Orientierungsrahmen für den Lernbereich Globale Entwicklung im Rahmen einer Bildung für nachhaltige Entwicklung*. 2., aktualisierte und erweiterte Auflage. Bonn: Cornelsen.

Mayrberger, K. & Zawacki-Richter, O. (2018): Qualität von OER: Internationale Bestandsaufnahme von Instrumenten zur Qualitätssicherung von Open Educational Resources (OER) — Schritte zu einem deutschen Modell am Beispiel der Hamburg Open Online University. Online verfügbar unter researchgate.net

Medienzentrum Hofgeismar / Wolfhagen (o. J.): SAMR-Modell zur Integration von digitalen Technologien im Unterricht. Online verfügbar unter mz-hofgeismar.de/apps

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (MSB) (Hrsg.) (2019): *Leitlinie Bildung für nachhaltige Entwicklung*. Düsseldorf.

Weßels, Doris (2025): Das wissenschaftliche Schreiben als Bildungsziel an Hochschulen: Ein Opfer von ChatGPT und Co.? In: Krauss, Christian; Kurdelski, Lutz-Peter; Kütz, Klaas Ole (Hrsg.): *AKWI 2025*. Thesis Paper, Kiel, Germany, 09.–11.09.2025. Bonn: Gesellschaft für Informatik e.V., S. 185–192. pISSN: 2944-7682. DOI: [10.18420/AKW2025_20](https://doi.org/10.18420/AKW2025_20). Online verfügbar unter dl.gi.de

Spohn, Lucca (2026): Eine KI-Antwort ist kein Text. Online verfügbar unter luccaspohn.substack.com

Anhang: Wie wird die Empfehlung berechnet?

Die abschließende Einordnung („Empfohlen“ / „Mit Anpassungen empfohlen“ / „Nicht empfohlen“) wird automatisch aus den angekreuzten Kriterien berechnet. Grundlage sind zwei Werte:

1. Gesamterfüllungsgrad

Über alle **einbezogenen** Unterkategorien (nicht ausgeschlossene Bereiche) hinweg wird gezählt, wie viele Kriterien angekreuzt sind, im Verhältnis zu allen möglichen Kriterien:

Gesamtwert = angekreuzte Kriterien ÷ mögliche Kriterien (in den einbezogenen Bereichen)

Das ist derselbe Wert, der oben als „X / Y Kriterien“ angezeigt wird.

2. Sonderprüfung der Bereiche A1 und B3

Unabhängig vom Gesamtwert wird zusätzlich geprüft, wie die Bereiche **A1 (fachliche Richtigkeit)** und **B3 (rechtliche Qualität)** jeweils für sich abschneiden — beide gelten als besonders kritisch für die Einsatzfähigkeit des Materials. Liegt einer der beiden Werte unter 60 %, wird dies als „schwach“ markiert. Ist ein Bereich bewusst ausgeschlossen, fließt er nicht in diese Prüfung ein.

3. Entscheidungstabelle

Bedingung	Ergebnis
Gesamtwert $\geq$ 80 % und A1/B3 nicht schwach	 Empfohlen
Gesamtwert $\geq$ 50 % und A1/B3 nicht schwach	 Mit Anpassungen empfohlen
Gesamtwert $\geq$ 50 %, aber A1 oder B3 schwach	 Mit Anpassungen empfohlen (+ Hinweis auf A1/B3)
Gesamtwert $<$ 50 %	 Nicht empfohlen

Kurz gesagt: Der **Gesamtprozentsatz bestimmt die Grundstufe** (80 %-/50 %-Schwelle), während **A1 und B3 als „Veto“** wirken — selbst bei hoher Gesamterfüllung verhindern schwache Werte dort die Bestnote „Empfohlen“.

Sind noch keine Kriterien angekreuzt, erscheint „— noch keine Angaben —“. Die Schwellenwerte (80 % / 50 % / 60 %) sind eine Orientierungshilfe und können bei Bedarf angepasst werden.