WELT 4: PRÜFSTERN · MISSION 4.1

Das Bewertungsraster, das erklärt

«Inhalt: gut» – aber was heisst gut? Sie lassen die KI ein schwammiges Raster zerpflücken und bauen eines, das Lernende schon vor der Abgabe verstehen.

Zur Mission

DAUER 50 Min. KOSTEN Gratis WERKZEUGE ChatGPT, Claude (Web), Google Gemini DATENAMPEL Grün – unbedenklich

Das Wichtigste

- Schwammwörter (gut, sorgfältig, angemessen, überzeugend) ersetzen durch das, was man sieht oder zählen kann.
- Kriterien kommen aus dem Lernziel bzw. der Handlungskompetenz im Bildungsplan, nicht aus dem Bauchgefühl.
- Produkt, Vorgehen und Begründung getrennt bewerten – Vorgehen und Begründung machen Denken sichtbar.
- Stufenlogik: pro Kriterium genau 3 prüfbare Punkte, die Stufe ist die Anzahl erfüllter Punkte (3 bis 0).
- Spalte «So kommst du eine Stufe höher» macht aus dem Messinstrument ein Lernwerkzeug.
- Stresstest mit zwei fiktiven Arbeiten: Wo Sie und die KI abweichen, ist das Raster oft mehrdeutig.
- Raster vor der Arbeit abgeben, nicht erst mit der Note. Die Notenumrechnung entscheidet die Lehrperson.

Ihr Produkt

Ein **erklärendes Bewertungsraster** für einen Ihrer Aufträge: 3–5 beobachtbare Kriterien mit je 4 Stufen, einer Spalte «So kommst du eine Stufe höher» und einer Ich-kann-Version für Lernende.

Abnahme: Jede Stufe jedes Kriteriums beschreibt etwas Beobachtbares (kein «gut», «sorgfältig», «weitgehend» ohne Erklärung), die beiden fiktiven Arbeiten aus dem Stresstest lassen sich in unter 5 Minuten eindeutig einstufen, und die Ich-kann-Version für Lernende liegt vor.

So gehen Sie vor

1. Die Schwammwort-Jagd
2. Kriterien aus dem Lernziel ableiten
3. Stufen, die erklären
4. Stresstest, Teil 1: zwei fiktive Arbeiten
5. Stresstest, Teil 2: blind einstufen lassen
6. Die Lernendenversion

Start-Prompt

Ich bin Lehrperson an einer Schweizer Berufsfachschule. Hier ist ein Bewertungsraster für eine Kundenberatung im Rollenspiel (Detailhandel, 2. Lehrjahr):

Kriterium | gut (3 P.) | genügend (2 P.) | ungenügend (1 P.)
Inhalt | Inhalt gut und vollständig | Inhalt weitgehend korrekt | Inhalt mangelhaft
Auftreten | sicher und professionell | meist sicher | unsicher
Fachwissen | fundiert | solide | lückenhaft
Gesamteindruck | überzeugend | in Ordnung | nicht überzeugend

Markiere jedes Wort, das nicht beobachtbar ist. Sag mir bei jedem: Was müsste eine lernende Person konkret tun oder sagen, damit eine Lehrperson «gut» ankreuzt? Und wo würden zwei Lehrpersonen vermutlich unterschiedlich entscheiden?
Schweizer Rechtschreibung (immer «ss», kein Eszett).

Qualitätscheck

- Jedes Kriterium ist einem Lernziel bzw. einer Handlungskompetenz zugeordnet.
- Keine Stufe enthält ein Schwammwort ohne Erklärung (gut, sorgfältig, angemessen, weitgehend, überzeugend).
- Produkt, Vorgehen und Begründung sind mit je mindestens einem Kriterium abgedeckt.
- Die Spalte «So kommst du eine Stufe höher» enthält konkrete Handlungen, keine Ermahnungen.
- Die Umrechnung in eine Note haben Sie selbst festgelegt (bzw. nach Vorgabe der Schule), nicht die KI.
- Im Stresstest wurden nur fiktive Arbeiten verwendet.

Hängen Sie fest? Frag die KI

Ich bin Lehrperson an einer Schweizer Berufsfachschule. Was ist der Unterschied zwischen einem beobachtbaren und einem nicht beobachtbaren Bewertungskriterium? Gib mir 5 Beispielpaare (schwammig → beobachtbar) aus verschiedenen Berufen. Sag mir auch, woran ich selbst erkennen kann, ob ein Kriterium beobachtbar ist.

Klappt nicht? Jede Gratis-KI funktioniert, z. B. Microsoft Copilot, Mistral Vibe (früher Le Chat) oder Perplexity. Für den Stresstest (Teil 2) lohnt sich ein zweites Modell.

Stand 19.11.2026 · CC BY 4.0 Alpasana GmbH · Werkzeuge ändern sich – prüfen Sie selbst.