

The Reader's Kit

EIGHT QUESTIONS TO ASK OF ANY AI TEST SCORE

A BUDDAY BUDDERSON PRODUCTION

 CITE THIS PAGE

Say an AI model “scores 43%”. Before you repeat that number, ask it eight questions. Each one below is a rule this program holds itself to, and each links to where it is set out in full.

01

Is the rate quoted with its interval?

A rate quoted bare is misquoted: the interval is the part that makes it evidence.

[Definitions](#)

02

Which interval: on this test, or on tests like it?

The narrow one compares week on week. Any claim about the model itself needs the wide one.

[Two rates, two intervals](#)

03

Under what conditions was it measured?

Gateway, serving provider, temperature, token budget, date. A rate belongs to a model measured a particular way.

[The draws](#)

04

What is the test made of?

A bank built from items models fail is adversarial by construction. Its failure rate is not a general error rate.

[Failure rate](#)

05

Would it survive being taken twice?

Two independent halves set a noise floor. A movement smaller than that gap is not reported.

[Noise floor](#)

06

Who graded it, and did the graders agree?

No judge grades its own model family, and agreement is published against a floor. Agreement is not correctness.

[Verdicts](#)

07

What went missing?

Draws lost to truncation, refusal or unreadable output. Above 5%, the run is disqualified: what vanishes is never random.

[Attrition](#)

08

Can you check it yourself?

Every published figure is registered with the SHA-256 of its sources, and every issue's fingerprint is printed on its PDF.

[Receipts](#)

TRY IT YOURSELF

The test, and a finished issue

The 37 public probes, each with its written pass criterion, one per line:

[bank_v1_1_public.jsonl](#). Run any of them on any model, and grade the answer against the criterion.

A finished issue, to read with the eight questions beside it: [The Calibration Vol 004 \(PDF\)](#). The [archive](#) has every issue.

Keep it: [the Reader's Kit as a PDF](#), or print this page. It is free and needs no sign-up.

Protocol STAMP · A Budday Budderson Production

Protocol STAMP and bud.day are under Budday Budderson Studio LLC, a New Mexico company

2105 Vista Oeste NW, Suite #E3993, Albuquerque, NM 87120

Begun by a person · Drafted by AI · Edited together · Signed by a person

Lawrence Ho · s@bud.day

Chief Human Investigator, Evaluation Administrator + Instrument Maker

Protocol STAMP can make mistakes. Verify before you implement anything. What we have already got wrong is in [Corrections + changes](#).

[Methodology](#) · [Definitions](#) · [Independence](#) · [Support](#) · [Corrections + changes](#) · [Other STAMPs](#) · [Subscribe](#) · [RSS](#) · [Public bank](#) · [Terms](#) · [Privacy](#) · [AI Safety + Governance](#)