
Good enough needs a test

What evidence should make us trust a workflow?

Use authorised information. Keep personal records in your own controlled workspace.

Task and acceptance rule

Representative ordinary cases

Ambiguous, missing and difficult cases

Expected answer or human rubric

Failure type and consequence

Review time and operating cost

Configuration and re-test trigger

Use the example. Then test your own.

INVENTED TEACHING EXAMPLE

Ten cases and two acceptance rules

Seven fictional outputs meet strict structure and support rules. One has supported facts but a different structure. Two invent commitments. Strict acceptance gives 7/10. Permitting the structure variation gives 8/10; both unsupported commitments still fail. These exercise rates are not production accuracy estimates.

Before treating this as evidence

Record the real inputs, source version and acceptance rules. Compare observed outcomes with the requirement. Keep unknowns and failure consequences visible.

What did you observe?

What remains unknown?

What will you do next—and who owns it?

Method and references: [the web edition's sources and scope](#). This is a practical editorial worksheet, not a certification or validated assessment instrument.