

The Live-Decision Pressure Test: A ChatGPT Action Guide

Turn one agreeable ChatGPT answer into a decision map you can inspect, correct, and test. By the end, you will have identified the decision's crux, priced the next action, and built a repeatable pressure-test process.

WHAT'S INSIDE

- 01 — Run your first pressure test
- 02 — Audit claims before accepting them
- 03 — Price the favorable path
- 04 — Find the crux and design a credible test
- 05 — Run the corrected second pass
- 06 — Follow a worked example
- 07 — Save the repeatable protocol
- 08 — Complete the guide and use the guardrails

01 Run your first pressure test

ChatGPT can return the posture built into your question. Your first job is not to ask for a better verdict. It is to assign distinct responsibilities, expose assumptions, and make unsupported certainty visible.

Before entering a real decision, remove confidential, personal, regulated, security-sensitive, or proprietary information unless you have determined that using it is permitted and appropriate. Anonymized or synthetic material is often enough for practice.

1

NOW

Open the tool

Open ChatGPT in the website or app and start a new chat. You will paste the practice scenario below into the message field and submit it.

2

FIRST RUN

Use the complete practice scenario

Start with the supplied scenario so you do not have to invent a decision before learning the method. It is illustrative, not a claimed client engagement or result.

PROMPT

Decision: Consider offering a paid 30-day AI workflow implementation sprint to small professional-services firms.

Supplied facts and constraints:

- Delivery cannot exceed 24 hours of provider time.
- The engagement cannot depend on custom software.
- A customer may handle sensitive client information.
- The customer's team must be able to maintain the resulting workflow.
- There is interest in the concept, but no supplied evidence of willingness to pay.

Assess this decision from four separate perspectives: advocate, critical skeptic, execution-focused operator, and prospective buyer.

For each perspective, identify the strongest legitimate value, material concerns, load-bearing assumptions, credible failure conditions, and evidence that could change the assessment. Give every material concern a mechanism: what could happen, why it could happen, why it matters, which assumption it challenges, and what evidence would strengthen or weaken it.

Label consequential claims as Supplied fact, Inference, or Unknown. Supplied fact means only that the claim appeared in my decision context; it does not mean independently verified.

Keep the four assessments separate. Do not count them as votes and do not recommend anything yet. After the separate assessments, identify the likely crux and the least costly credible test that could materially change the decision.

1

AFTER IT ANSWERS

Look for the required structure

Confirm that you received four meaningfully different assessments, explicit claim labels, at least one concern with a causal mechanism, a crux, and a proposed test. You are checking structure, not accepting the content as true.

2

IF IT BLENDS THE ROLES

Make it separate the work

Paste the repair prompt below. Early synthesis often turns real disagreement into variations of “proceed carefully,” which is not much of a decision tool.

PROMPT

Revise the response. Keep each perspective in a separate section and remove any combined recommendation. Each perspective must protect a different consequence: value creation, logical validity, execution reliability, or buyer impact. Do not give two perspectives the same concern in different language. Preserve disagreements until the final crux analysis.

Your first success is complete when the answer has stopped behaving like a verdict and started behaving like a map of arguments, assumptions, consequences, and evidence needs. Do not finish by asking, “So, is it good?”

02 Audit claims before accepting them

Multiple roles increase coverage, not corroboration. Every voice still comes from the same AI and may share the same missing context or flawed premise. Audit the map before using it.

PROMPT

Audit your previous response claim by claim.

1. List every consequential claim that affects value, demand, feasibility, trust, cost, capacity, or risk.
2. Show its current label: Supplied fact, Inference, or Unknown.
3. Quote or identify the exact supplied context supporting anything labeled Supplied fact.
4. Relabel any generated assumption that was phrased as fact.
5. For every material concern, show the mechanism, consequence, challenged assumption, and evidence path.
6. Remove claims that are merely plausible atmosphere and do not affect the decision.

Do not add a recommendation.

Apply the human correction

Compare every claimed supplied fact with the original decision packet. If you did not provide it, it is not a supplied fact. Then use this exact correction on the practice scenario:

SCRIPT

Revise the decision map with these corrections:

- Willingness to pay is Unknown.
- Any hour-by-hour delivery allocation is an Inference and planning estimate, not an established fact.
- Possible handling of sensitive information is supplied. The claim that confidentiality will be a decisive buyer objection remains an Inference unless buyer evidence supports it.
- General interest is supplied, but it is not evidence that anyone will purchase this offer.

Show which conclusions became weaker, stronger, or unchanged after these corrections. Keep the perspectives separate and do not issue a verdict.

CHECKLIST

- A supplied fact can be traced to the decision context.
- An inference is plausible but not treated as established.
- An unknown is connected to evidence that could resolve it.
- A material concern explains what happens and why.
- Polished wording is not being counted as evidence.
- The four roles are being used as search probes, not votes.

03 Price the favorable path

Now assume the strongest positive case deserves further investigation. Make the operator price it. A recommendation is not operationally useful until it names ownership, displaced work, dependencies, and stopping conditions.

PROMPT

Take the strongest evidence-compatible version of the favorable case and apply an operator's test.

Identify:

- The next concrete action and its owner
- The capacity or budget it consumes
- The existing work that moves, shrinks, or stops
- Skills, access, approvals, or other dependencies
- Dependencies outside the owner's control
- The result that would justify expansion
- The result that would trigger revision or cancellation

Use Unknown wherever the decision context does not supply an answer. Do not invent people, budgets, access, approvals, capacity, or delivery estimates. Separate supplied constraints from proposed planning assumptions.

What good looks like

CHECKLIST

- One person or accountable role owns the immediate next action.
- The answer names what the proposed work would displace.
- External dependencies are visible rather than buried.
- Planning estimates remain labeled as inferences.
- Expansion requires an observable result.
- Revision or cancellation has an observable trigger.

If the answer says to run a pilot but cannot say who runs it, what it consumes, what it interrupts, or what would stop it, the pilot has not been priced. “Start small” is not a plan until small has boundaries.

04 Find the crux and design a credible test

A load-bearing assumption can materially change scope, timing, investment, or the decision itself. The crux is the unresolved assumption most likely to change what you do next.

PROMPT

Identify no more than five load-bearing assumptions in the corrected decision map. Apply this flip test to each one: If this assumption proved false, would the scope, timing, investment, or decision materially change?

Rank the assumptions by likely decision impact and identify one crux. Explain why the other assumptions are secondary for the next step.

For the crux, propose three possible tests. For each test, state:

- The evidence it could produce
- What it would cost or expose the decision-maker to
- Why the evidence would or would not be credible enough to change the decision
- What result would support continuing
- What result would require narrowing, revising, or stopping

Select the least costly credible test, not merely the cheapest activity. Reject tests that are inexpensive but too weak to affect the decision.

Use the right kind of test

CRUX	WEAK PROXY	MORE CREDIBLE SMALL TEST
Willingness to pay	General enthusiasm or compliments	Discuss the named costly problem with prospective buyers and, if supported, use a tightly bounded paid pilot to test purchase behavior.
Delivery capacity	An optimistic hour estimate	Rehearse discovery, setup, testing, documentation, training, and handoff using non-sensitive test material.
Maintainability	Telling the customer that the workflow is simple	Test whether the intended owner can operate and maintain the bounded workflow after handoff.
Trust or access	Assuming that a disclaimer resolves concern	Define exact access boundaries, exclusions, approvals, and human review before exposing real information.
Policy effectiveness	Visibility or anecdotes	Name the demonstrated failure the policy is supposed to address and run a bounded test with success and stopping criteria.

These are test patterns, not automatic prescriptions. Legal, safety, financial, employment, privacy, security, and customer commitments require appropriate sources and qualified human judgment.

05 Run the corrected second pass

The second pass should use your corrections rather than allowing the AI to quietly rebuild the same assumptions. Its job is to show how the decision process changes.

PROMPT

Using only the supplied decision context, the corrected claim labels, and the operator analysis, produce a revised decision map.

Return these sections:

1. Strongest legitimate upside worth preserving
2. Scope that should be narrowed or clarified
3. Decision crux
4. Least costly credible test
5. Owner of the next action
6. Capacity consumed and work displaced
7. Dependencies
8. Evidence that would justify continuing or expanding
9. Evidence that would trigger revision or cancellation

Keep Unknowns visible. Do not convert a plausible inference into a fact. Do not make a final verdict when the required evidence is still missing.

Compare the first and second answers

The pressure test has earned its keep if it changed at least one part of the decision process:

- Scope became narrower or more explicit.
- Sequence changed so evidence comes before commitment.
- The evidence plan became more credible.
- The initial commitment became smaller or better bounded.
- Success criteria became observable.
- Ownership and displaced work became visible.
- Stopping conditions were added.

If nothing changed except the length of the answer, you may have generated language rather than leverage. Return to the claim audit and ask which unknown could actually change the decision.

06 Follow a worked example

This example is illustrative. It demonstrates the moves without claiming a real engagement or result.

Decision context

EXAMPLE

Decision: Add a paid priority-support option for existing customers.

Supplied facts and constraints:

- The same team would continue delivering standard support.
- No additional hire has been approved.
- Existing standard-support commitments must continue.
- There is no supplied evidence that customers will pay for priority support.
- The offer needs an accountable owner and clear service boundaries.

MOVE	WHAT THE PRESSURE TEST SURFACES	HUMAN JUDGMENT
Separate perspectives	The advocate sees a potentially clearer paid offer. The skeptic challenges demand. The operator asks what happens during competing requests. The customer perspective examines whether the promise solves a costly problem.	These are candidate arguments, not four independent findings.
Correct claim status	“Customers value faster support” may sound plausible. “The team can absorb the work” may sound operationally reasonable.	Both remain Unknown or Inference without evidence. Neither belongs in the business case as a supplied fact.
Apply the operator test	The offer needs an owner, service boundaries, a backup plan, capacity rules, and a definition of what standard work may not be displaced.	A broad launch is unpriced until those operating conditions are explicit.
Find the crux	The key question becomes whether customers have a costly enough delay to purchase a narrow support promise that the current team can reliably fulfill.	General interest in better service would not resolve that question.
Design evidence	Discuss specific costly delays with prospective buyers, rehearse delivery against a synthetic request queue, and use a bounded paid pilot only if the earlier evidence supports it.	The sequence tests demand and delivery without treating a full launch as the first experiment.
Record the change	The decision moves from “Should we launch priority support?” to “Can one bounded service promise create enough buyer value without degrading standard support?”	The output is now an evidence question with ownership and stopping conditions, not a polished opinion.

07 Save the repeatable protocol

Use the following prompt after you have practiced once. Replace the bracketed context with concise material from a real decision. Include only information you are permitted to enter.

TEMPLATE

DECISION: [One commitment under consideration]

SUPPLIED FACTS: [Only facts provided in the decision context]

CONSTRAINTS: [Time, capacity, budget, access, policy, trust, or delivery limits]

STAKES: [What may be affected if the decision is wrong]

Assess the decision through these separate responsibilities:

- Advocate: build the strongest evidence-compatible case and state the conditions required for value.
- Critical skeptic: identify weak logic, hidden costs, disconfirming evidence, and credible failure mechanisms.
- Execution-focused operator: account for sequence, capacity, skills, dependencies, maintenance, ownership, displaced work, and stopping conditions.
- Affected stakeholder: examine adoption, trust, inconvenience, risk, and unintended consequences from the receiving end.

For each perspective, provide the strongest legitimate upside, material concerns with mechanisms, load-bearing assumptions, failure conditions, and evidence that could change the assessment. Label consequential claims as Supplied fact, Inference, or Unknown.

Keep assessments separate and do not vote. Then audit the claim labels, apply the flip test, identify one crux, and compare possible tests. Select the least costly credible test capable of changing the decision.

Finish with a bounded next action that names the owner, capacity consumed, work displaced, dependencies, success criteria, and revision or cancellation conditions. If evidence is insufficient, say so instead of manufacturing a verdict.

Change the perspectives when consequences change

- For a large financial commitment, replace a duplicate role with a cash-flow perspective.
- For a workplace policy, consider employee, manager, customer-facing, and compliance responsibilities.
- For a product or workflow decision, include support, maintenance, and affected-customer responsibilities.
- Remove roles that produce the same concern in different words. More headings do not create more evidence.

Use the fast version for a smaller decision

PROMPT

Give me the strongest evidence-compatible case for this decision, the strongest case against it, and one credible test of the load-bearing assumption. Keep supplied facts, inferences, and unknowns separate. Do not give me a verdict. Show what would change the scope, sequence, commitment, or stopping conditions.

Practice the skill again

1. Run the protocol on a service or offer decision and emphasize demand, delivery, and maintenance.
2. Run it on a policy decision and emphasize who experiences the costs and who protects customer coverage.
3. Run it after receiving new evidence and ask which assumptions changed rather than requesting an entirely new opinion.
4. Compare an advocacy prompt such as “Help me explain why this will work” with the structured-opposition prompt. Notice how changing the assigned job changes the answer’s burden of proof.

08 Complete the guide and use the guardrails

CHECKLIST

- I ran separate assessments with meaningfully different responsibilities.
- I treated the perspectives as search probes rather than independent sources or votes.
- I checked supplied facts against the material I actually provided.
- I relabeled generated assumptions as Inference or Unknown.
- I required mechanisms and evidence paths for material concerns.
- I named the owner, capacity, displaced work, dependencies, and stopping conditions.
- I found a crux that passes the flip test.
- I selected the least costly credible test rather than the cheapest weak activity.
- I can name at least one change to scope, sequence, commitment, evidence, ownership, or stopping conditions.
- I saved the reusable protocol for another decision.

Guardrails that remain in force

- Four AI perspectives are still one AI, not four independent experts.
- Shared blind spots, invented details, and flawed premises can survive every role.
- Harsh language is not stronger reasoning. A concern needs a mechanism, consequence, and evidence path.
- Forced balance should not preserve a weak argument. Evidence determines weight.
- Review effort should reflect the stakes. A reversible, low-cost choice may not require this full process.
- Debate is not verification. Consequential facts require appropriate sources and accountable human judgment.
- Do not enter sensitive information unless doing so is permitted and appropriate for the system and account you are using.

The finish line is not an AI verdict. It is a decision you can describe in operational terms, an uncertainty capable of changing it, and a credible next test with clear boundaries.

Want the complete reasoning behind the method?

Read the companion article, “How to Pressure-Test ChatGPT When It Keeps Agreeing With You,” when available at <https://chuckgoetschel.com/ai-insights/pressure-test-chatgpt-answers>

INSTAGRAM · FACEBOOK · TIKTOK · YOUTUBE · CHUCKGOETSCHEL.COM

Agreement is an output. Better judgment begins when you expose what must be true and test it before you commit.