

Before You Trust It, Ask It to Check Itself

A Self-Guided Practice Exercise

AI as Your Thinking Partner

We did not get to this demo during our session, so here it is as a self-guided exercise. Work through it at your own pace. The goal is simple: you are going to ask the AI a health question, and then you are going to ask it to cross-examine its own answer.

Before you dive in, a reminder about context. This technique is for Level 3 questions, the ones we talked about where the stakes are real and a wrong answer actually costs you something. It is not for "what is a good movie on Netflix tonight."

The prompts here use a medical example. That is not an accident. Health questions are precisely where you cannot afford to just nod along because the answer sounds organized and sure of itself.

This matters because AI delivers solid facts, shaky evidence, and educated guesses all in the same confident tone. You cannot tell them apart just by reading. These prompts fix that.

The 4-Step Method

We are going to ask one question and then follow it up with three cross-examination prompts. Copy and paste each prompt exactly as written.

PROMPT 1 — ASK YOUR QUESTION

Paste this into your AI tool of choice and press Enter:

```
I am 72 years old and have mild osteoarthritis in my hands and knees. Morning stiffness bothers me the most.
```

```
I have heard that turmeric or curcumin supplements may help arthritis pain.
```

```
Please explain in plain English:
```

- ```
1. What the evidence says about turmeric or curcumin for osteoarthritis
2. What benefits people might realistically expect
3. What risks or medication interactions I should know about
```

```
I take low-dose aspirin and atorvastatin. I am not asking for a personal medical recommendation. I want to understand whether turmeric or curcumin might actually help me and how I should think through the decision before trying them.
```

**What to look for:**

- The answer will look well-organized and confident
- It may reference studies or cite research
- This is the moment most people stop and head to the vitamin aisle
- We are not stopping here

**PROMPT 1A — SORT FACTS FROM GUESSES**

Stay in the same conversation. Paste this in as your next message:

Before I rely on your answer, separate it into two lists:

1. Claims you are highly confident are factual
2. Claims that may be uncertain, incomplete, oversimplified, or based on general pattern rather than specific evidence

Use simple language. Be honest about what you are less sure about.

**What to look for:**

- The AI will start splitting its own answer into two categories
- Some things that sounded very solid will move to the uncertain list
- This is normal, and it is useful

**PROMPT 1B — FIND THE HOLES**

Now push a little harder. Still in the same conversation, send this one next:

Now review your answer as if you were trying to find mistakes.

What parts of your answer could be wrong, outdated, exaggerated, or missing important context?

Give me a short checklist of what I should verify before I act on this information.

**What to look for:**

- You are asking the AI to play prosecutor against its own answer
- It will not volunteer this kind of self-doubt on its own
- You always have to ask for it

**PROMPT 1C — FLAG THE EXPIRATION DATES**

One more. Same conversation, same message box:

Which parts of your answer depend on current or changing information?

For example: studies, medical guidelines, supplement safety warnings, drug interactions, or dosage recommendations.

Tell me what should be checked with an up-to-date source.

**What to look for:**

- Every AI has a knowledge cutoff date, a point after which it has not learned anything new
- Drug interactions and safety warnings can be updated after that date
- This prompt forces the AI to tell you exactly where its knowledge may be stale

*At this point you have used four prompts to pull a much more trustworthy answer out of the AI. More importantly, you now have a short checklist of things to verify with a pharmacist or your doctor before acting on any of it.*

**The Shortcut: All Four Questions in One Prompt**

---

Once you have practiced the 4-step method a few times, you can combine everything into a single prompt. Here it is with a different question, because plenty of us are wide awake at 3 a.m. whether we like it or not.

Open a new chat and paste this entire prompt at once:

I am 70 and waking up at 3 a.m. most nights. What does the evidence say about melatonin for older adults?

Please cover:

1. What melatonin actually does
2. What benefits older adults might realistically expect
3. Typical dosage ranges
4. Possible side effects or interaction concerns

In your answer:

- Separate facts from assumptions
- Tell me what I should verify
- Tell me what you are unsure about
- Avoid sounding certain when the evidence is weak or mixed

#### **What to look for:**

- You should see the AI hedging more than usual
- It should flag some things as uncertain or worth verifying
- Compare this to what a plain question would have produced

## **Try It With Your Own Question**

---

The all-in-one format works for any health, money, or legal question. The pattern is always the same:

1. Ask your real question with enough context (who you are, what you take, what you actually want)
2. Ask it to separate confident facts from uncertain claims
3. Ask it to find what could be wrong or missing in its own answer
4. Ask it to flag what might be outdated and needs a current source

### **Key Takeaway**

You are not asking the AI to be an oracle. You are hiring it as a research assistant who has to show its work.

The mistakes are always in there. They will not announce themselves. But the moment you ask the right follow-up questions, the AI will usually

come clean.

After all that cross-examination, the final check is still with a real human professional who knows your specific situation.