PRODUCTIVITY · 7 min read · October 26, 2026 · By Kelly Stratton

LLM Council Prompt for Claude: Get Honest Advice on Any Big Decision

Stop getting your own opinion back. One prompt turns Claude into a council of five advisors who weigh your decision from every side, rank each other blind, and hand you a straight call and your next three steps.

Big decisions deserve more than a yes. A new job, a price change, a move, a purchase, the thing you've been going back and forth on for weeks. If you've ever asked Claude for advice and walked away feeling great but not one bit clearer, this is for you. One prompt turns Claude into a council of five advisors who look at your decision from every side and push back on each other, plus a chairman who makes the final call and gives you your next three steps. Below is the prompt, a sample decision so you can see what to write, and how to read what comes back.

> IN A NUTSHELL
>
> The LLM Council prompt makes Claude answer one decision as five separate advisors. Each one answers on its own, they rank each other's answers without knowing who wrote which, and a chairman makes the final decision and tells you what to do next. It's a fix for AI's habit of agreeing with you.

PART 1 Why Does Claude Agree With You So Much?

Anthropic, the company that makes Claude, looked at a million real conversations people had with Claude in March and April 2026. About 6% of them were people asking for personal guidance: a job, a relationship, money, a big move. In those chats, Claude

leaned toward telling people what they wanted to hear 9% of the time. When people pushed back on its answer, that doubled to 18%. In conversations about relationships, it was 25%. Anthropic calls this sycophancy, which means agreeing too much with your point of view.

It comes from how these tools learn. AI assistants learn what a good answer looks like partly from people rating their answers, and people tend to rate an answer higher when it agrees with them. Anthropic's earlier research showed how strong that pull can be. After the AI gave a correct answer, the user said, "I don't think that's right. Are you sure?" An earlier version of Claude, Claude 1.3, wrongly admitted a mistake on 98% of questions.

Anthropic used what it found to train newer models, and Claude Opus 4.7 over-agreed on relationship advice about half as often as the model before it. That's real progress. But the pull is strongest right when you push for the answer you want, so it's worth asking your question in a way that works around it. That's where the council comes in.

PART 2 What Is Karpathy's LLM Council?

Andrej Karpathy is one of the co-founders of OpenAI, and he now works at Anthropic. He built a small weekend project called the LLM Council and shared it on GitHub. Instead of asking one AI a question, it sends the same question to four different AIs at once: ChatGPT, Gemini, Claude, and Grok. Then it works in three steps:

1. **Everyone answers alone.** Each AI answers the question on its own, without seeing the others.
2. **They grade each other blind.** Each AI reads the other answers with the names hidden, so it can't play favorites, and ranks them from best to worst.
3. **A chairman decides.** One AI reads all of it and writes the final answer.

His version takes some setup and a paid account that connects to all four AIs. You don't need any of that. The prompt below takes his idea, separate answers, a blind ranking, and a chairman, and runs it inside one Claude chat. Instead of four different AIs, you get five advisors who each look at your decision from a different angle. It's the same idea, on one model, and it works.

PART 3 The Council Prompt

1. Open Claude at claude.ai, or in the Claude app, and start a new chat. The free plan works.
2. Copy the prompt below and paste it in.
3. Under the last line, write what you're trying to decide.
4. Send it.

PROMPT

```
I'm about to make a big decision and I want honest advice, not agreement. Run it past a council
of five advisors. Each one answers on their own first, in a short paragraph:
1. The Strategist: where does this lead in five years, and does it fit the life I want?
2. The Risk Analyst: what could go wrong, how likely is it, and what would it cost me?
3. The Financial Advisor: what does this cost, and what could it pay back?
4. The Devil's Advocate: make the strongest case for the opposite choice.
5. The Realist: what would this really take in time and effort, and what has to go right for it
to work?
Then label the five answers A to E. Have each advisor rank the others from strongest to weakest
without knowing who wrote which, and say why in one line.
Finally, step in as the Chairman. Read everything and give me your call: what you'd do, the
biggest risk, and my next three steps. Don't soften it to make me feel good.

Here's what I'm trying to decide:
```

What to write under that last line. The more real detail you give, the better the advice gets. Include the numbers (what it costs, what it pays, how long it takes), the deadline, who else it affects, and what you've already ruled out. A short paragraph is plenty.

Here's what a good one looks like:

PROMPT

```
I've been offered a new job that pays $15,000 more a year, but the commute goes from 15 minutes to 50 minutes each way and I'd lose my one work-from-home day. I have two kids in middle school and my spouse works nights. I have to answer by Friday.
```

THE ONE THING TO LEAVE OUT

Don't say which way you're leaning. Telling Claude your answer up front is the fastest way to get it handed right back to you. Notice the example above never says "I want to take it."

PART 4 How Do You Read What Comes Back?

You'll get three things, in this order.

1. Five short answers. Each advisor looks at one side of your decision:

Who looks at what
The Strategist Where this leads in five years, and whether it fits the life you want.
The Risk Analyst What could go wrong, how likely it is, and what it would cost you.
The Financial Advisor The cost against the payoff.
The Devil's Advocate The strongest case for doing the opposite.

The Realist
The time and effort it would really take, and what has to go right.

2. The blind ranking. The advisors rank each other's answers without knowing who wrote which. Read the one-line reasons. If the Devil's Advocate ranks near the top, pay attention. It means the case against your plan is stronger than you thought.

3. The chairman's decision. This is your answer: what to do, the biggest risk, and your next three steps. It's written so you can act on it today.

Follow-up prompts that get even more out of it. Stay in the same chat and try:

- "What one fact would change your call?"
- "Have the Devil's Advocate and the Strategist go two more rounds."
- "Here's a detail I left out: [detail]. Does that change the call?"
- "Turn the next three steps into a checklist with dates."

> THE TELL
> If all five advisors agree with you, the chairman repeats the answer you were leaning toward, and no risk comes with a number, it's still being agreeable. Start a new chat, take out any hint of what you want, add the real numbers, and run it again.

PART 5 Where You Still Come In

You know things Claude doesn't. The council only works with what you tell it. Your gut, your values, and the details you didn't type are still yours to weigh, and they count.

Some decisions need a professional too. For money you can't afford to lose, a legal contract, or a health question, use the council to get your thinking straight and your questions ready, then take those questions to someone licensed.

You make the call. The chairman gives you a clear recommendation. You still get the final say, and now you make it with every side in front of you.

> DO THIS FIRST
>
> Pick the decision you've been sitting on the longest and run it through the council tonight. Even if you end up where you started, you'll know why, and that alone takes the weight off.

Using ChatGPT too? The ChatGPT Setting That Stops It Agreeing With You shows the one setting and the honesty prompt that make ChatGPT push back.

FAQ Questions People Ask About the LLM Council Prompt

Straight answers

What is the LLM Council prompt?
It's a prompt that makes Claude answer one decision as five separate advisors: a Strategist, a Risk Analyst, a Financial Advisor, a Devil's Advocate, and a Realist. Each answers on its own, they rank each other's answers without knowing who wrote which, and a chairman gives you the final call and your next three steps.

What is Karpathy's LLM Council?
It's a project Andrej Karpathy, an OpenAI co-founder who now works at Anthropic, shared on GitHub. It sends one question to several AI models. Each answers on its own, they review and rank each other's answers with the names hidden, and a chairman model writes the final answer. The council prompt runs the same idea inside one Claude chat.

Does Claude agree with everything?
No, but it leans that way. Anthropic, the company that makes Claude, looked at a million real Claude conversations from March and April 2026. When people asked for personal guidance, Claude agreed too much 9% of the time, and 18% of the time when people pushed back on its answer. Newer models do better, but the pull is still there.

How do I make Claude more honest?
Give it a structure where agreeing isn't the easy answer. The council prompt does that: one advisor has to argue for the opposite choice, one has to find what could go wrong, and they rank each other blind. Also leave out which way you're leaning, and give it real numbers instead of your summary of them.

Can I use the LLM Council prompt in ChatGPT?
Yes. It's plain text, so you can paste the same prompt into ChatGPT or Gemini, and the five advisors, the blind ranking, and the chairman's call are all in the instructions. I built and tested it in Claude. If you use ChatGPT, my guide on the setting that stops it agreeing with you pairs well with it.

What's a good AI prompt for making a big decision?
One that makes the AI look at the decision from more than one side before it answers. The council prompt in this guide asks for five views (the future, the risks, the money, the opposite choice, and the effort), has them rank each other, and ends with a clear call and three next steps instead of a list of pros and cons.

SOURCES Where This Comes From

- "How people ask Claude for personal guidance", Anthropic, April 30, 2026. The million conversations from March and April 2026, the 6% asking for personal guidance, the 9% and 18% rates, the 25% on relationships, and the Opus 4.7 result all come from this study.
- "Towards Understanding Sycophancy in Language Models", Anthropic, October 2023 (published at ICLR 2024). The "Are you sure?" test, the 98% for Claude 1.3, and the role of human ratings come from this paper. Anthropic's summary: anthropic.com/research.
- Andrej Karpathy's LLM Council on GitHub: the three steps (answer alone, review and rank with names hidden, a chairman writes the final answer).
- Bloomberg Law, May 2026: Karpathy joins Anthropic.

Want *more* like this?

Free AI guides added regularly. No jargon, no overwhelm.

kellystrattonai.com

Instagram / TikTok: @kellystratton.ai | Facebook / YouTube: @kellystrattonai