

# *How to Harness the Power of* **the Most Intelligent Frontier AIs in 2026.**

*The PMPR guide to getting consistently better answers from frontier AI — through flexible, creative, adaptive thinking.*

## THE CORE PRINCIPLE

*If you ask enough questions enough different ways, you will always get a better result. The technique is a three-step protocol. The skill that makes it succeed is your own flexibility, creativity, and willingness to change direction when the conversation demands it. The protocol is simple. The mindset is everything.*

## IN PRACTICE

*“It is a good sign if you can keep Claude Sonnet 4.6 engaged for several minutes researching your questions from several different perspectives and integrating them together into a comprehensive consensus recommendation while you wait.”*

THE DIRECT ANSWER / READ THIS FIRST

# The Single Most *Important Thing*

The single most important thing you can do to harness the power of frontier AI in 2026 is **stay in the conversation longer than feels necessary**.

Most people treat frontier AI like a search engine — one question, one answer, done. The power is not in the first answer. It is in what emerges when you keep pushing: challenging the response, asking it to approach the problem from a different angle, telling it when something feels wrong, changing the goal when the conversation reveals a better question. Claude Sonnet 4.6 at ~20 HBE (Human Brain Equivalents — a measure of AI reasoning capacity based on active parameter count, where one HBE approximates the cognitive throughput of a single human brain <sup>[1]</sup>) and Gemini 3.1 Pro at ~5 HBE are operating at many times the reasoning throughput of a single human expert. The limiting factor in most interactions is not the AI — it is the human ending the conversation too soon.

THE ONE-SENTENCE VERSION

*Treat Claude or Gemini less like Google and more like the smartest colleague you have ever had — one who has read everything, forgets nothing within the conversation, has no ego about being wrong, and gets genuinely better the longer and harder you push it.*

# Six Principles, In Order of Impact

#	PRINCIPLE	WHAT IT MEANS IN PRACTICE
1	<b>Give it genuine complexity</b>	These systems are underutilised on simple tasks. The questions that take a human expert an hour to think through carefully are exactly where frontier AI earns its capability. Bring your hardest problems, not your easiest ones.
2	<b>Treat the first answer as a draft</b>	Ask enough questions enough different ways. The first response is where the model's momentum takes it. The fifth or sixth exchange — after you have challenged it, redirected it, and asked it to reconcile conflicting perspectives — is where the real value lives.
3	<b>Tell it when it is wrong</b>	Frontier AI at this level responds well to being corrected. It does not get defensive. Saying "that doesn't feel right — here is why" produces a meaningfully better response than accepting the answer and moving on.
4	<b>Change goals when the conversation earns it</b>	This guide started as a technical architecture paper and became a practical human guide because a mid-conversation correction revealed that was the more valuable thing to make. Follow that signal whenever it appears.
5	<b>Use it to think, not just to answer</b>	Ask it to steelman positions you disagree with. Ask it what you are missing. Ask it for the best argument against your current plan. Use it as a thinking partner, not an oracle.
6	<b>Stay curious, stay flexible</b>	Flexibility, creativity, and adaptability are not just helpful — they are the essential characteristics that separate practitioners who get good results from those who get excellent ones. The rest of this guide shows you exactly how to apply them.

*The guide that follows gives you the structured technique. This page gives you the spirit behind it. Both matter — but if you only read one page, read this one.*

## CONTENTS

# Table of Contents

## OPENING

---

– **The Answer — Read This First** 02

THE SINGLE MOST IMPORTANT THING · SIX PRINCIPLES IN ORDER OF IMPACT · THE ONE-SENTENCE VERSION

---

## THE PRACTICAL GUIDE

---

§ 01 **The Core Principle and the Mindset** 04

WHY ASKING DIFFERENTLY WORKS · FLEXIBILITY, CREATIVITY AND ADAPTABILITY AS SKILLS · WHAT THIS TECHNIQUE SOLVES

---

§ 02 **Which AI to Use — and Why It Matters** 05

WHY MOST AIS FALL SHORT · CLAUDE SONNET 4.6 AT 20 HBE · GEMINI 3.1 PRO AT 5 HBE · CAPABILITY REQUIREMENTS EXPLAINED

---

§ 03 **The Three-Step Protocol (PMPR)** 06

STEP 1: INITIAL PASS · STEP 2: CHALLENGE · STEP 3: SYNTHESIS · COPY-PASTE PROMPTS · THE ONE RULE

---

§ 04 **The Six Thinking Lenses** 07

ADVERSARIAL · EMPIRICAL · PROCEDURAL · CAUSAL · HISTORICAL · ETHICAL · MATCHING LENS TO DOMAIN

---

§ 05 **Worked Examples** 08

LEGAL · MEDICAL · TECHNICAL · FINANCIAL · WHAT CHANGES AT EACH STEP AND WHY

---

§ 05b **Worked Examples — Fraud & Deepfakes** 09

SCAM DETECTION · DOCUMENT FRAUD · DEEPPFAKE IDENTIFICATION · ADAPTIVE INVESTIGATION WORKFLOW

---

---

§ 06	<b>Creative Adaptation — Going Off-Script</b>	10
	CHANGING GOALS MID-CONVERSATION · PIVOTING WHEN STEP 2 REVEALS SOMETHING UNEXPECTED · INVENTING NEW LENSES · READING SIGNALS	

---

§ 07	<b>When to Use It and When Not To</b>	11
	QUERY ROUTING TABLE · SCALING DEPTH TO STAKES · TIME VS. RELIABILITY · THE PROFESSIONAL STANDARD	

---

§ 08	<b>Common Mistakes and How to Avoid Them</b>	12
	SEVEN FAILURE MODES · WARNING SIGNS · HOW TO TELL IF IT WORKED	

---

§ 09	<b>Questions and Answers</b>	13
	IS THIS JUST ASKING NICELY? · SAME WEIGHTS OBJECTION · WHY AI DOESN'T DO THIS BY DEFAULT · HONEST LIMITATIONS	

---

#### THE BIGGER PICTURE

---

§ 10	<b>Why This Is a Workaround, Not a Fix</b>	14
	THE ARCHITECTURAL ROOT CAUSE · PERSONA ENSEMBLE ARCHITECTURE · THE 2042 VISION · HBE SCALING	

---

§ 11	<b>Conclusion &amp; Recommendations</b>	15
	FOR INDIVIDUALS · FOR TEAMS · FOR AI LABS AND ENGINEERING LEADERS	

---

#### BACK MATTER

---

—	<b>Footnotes</b>	16
	HBE ESTIMATES · PLATFORM ASSESSMENT · COPYRIGHT & PERMITTED USE · PEA/ORI STATUS DISCLOSURE	

---

—	<b>References &amp; Appendix</b>	17
	8 SOURCES · HBE PARAMETER DEFINITION · GROWTH RATE ASSUMPTIONS · WHAT THIS PAPER DOES AND DOES NOT CLAIM	

---

## SECTION 01 / THE PRINCIPLE AND THE MINDSET

# One Question, One Answer. *That Is the Problem.*

Every time you ask an AI a question — any AI, any platform — you get one answer. That answer is built by a process that commits to a direction from the first word and follows it with complete fluency to the end. The problem is not the answer itself. The problem is the assumption baked into the architecture: that one direction of thought, on one attempt, is enough.

It never is. Not for any question that matters.

## THE CORE PRINCIPLE

*If you ask enough questions enough different ways, you will always get a better result. Each new angle surfaces something the previous one missed. The gap between perspectives is where the real information lives.*

This principle applies to all thinking, not just AI. Doctors seek second opinions. Lawyers argue both sides. Scientists falsify rather than confirm. The difference with AI is that the tool is powerful enough, and fast enough, that you can apply this principle in minutes rather than days. You just have to choose to.

## The Three Skills That Make This Work

The technique in this guide is a three-step protocol. But the protocol is a scaffold, not the point. What actually produces better answers is a set of cognitive habits that the protocol is designed to exercise:

**Flexibility of thought**

Treating the first answer as a starting point, not a conclusion. Being genuinely willing to hear that it is wrong.

Without this, Step 2 feels like a waste of time and you stop doing it. With it, every challenge becomes valuable information.

**Creativity**

Inventing new angles when the standard lenses do not fit. Reframing the question entirely when the conversation reveals a better question to ask. Seeing unexpected Step 2 results as opportunity rather than noise.

The protocol gives you six lenses. The real world gives you infinite situations. Creativity is what bridges the gap between a structured technique and a genuinely novel problem.

**Adaptability — changing goals on the fly**

Recognising when Step 2 has revealed that you were asking the wrong question entirely, and pivoting to the right one before running Step 3. Being willing to discard the original goal when a better one emerges mid-conversation.

This is the most powerful and most underused aspect of PMPR. The best outcomes often come not from completing the original protocol on the original question, but from letting the challenge pass redirect the entire inquiry.

*The protocol is the minimum. Creativity and adaptability are what turn a reliable technique into a genuinely powerful one. Use the steps as a starting point — then follow the conversation wherever it needs to go.*

**Fixed vs. Flexible Mindset with AI**

## FIXED MINDSET

## FLEXIBLE, CREATIVE, ADAPTIVE MINDSET

"That sounds right — I'll use it."

"That sounds plausible — what angle haven't I tried?"

Treats AI confidence as a proxy for accuracy

Knows AI confidence and AI accuracy are independent variables

Sticks to the original question no matter what

Lets the conversation reveal the better question and chases it

Sees a different answer as a contradiction to resolve

Sees a different answer as a map of where the uncertainty lives

Uses the protocol as a checklist to complete

Uses the protocol as a launchpad — then improvises

## THIS GUIDE WAS BUILT USING ITS OWN METHOD

This document did not arrive fully formed. It was built through approximately forty exchanges between the author and Claude Sonnet 4.6 — each one applying a different angle, a different challenge, a different goal. Early drafts focused on the wrong thing. Sections were restructured when the conversation revealed a better structure. The emphasis shifted from architecture to workaround when a push-back made clear that was the real priority. Abbreviations were caught. The platform recommendation was sharpened. The mindset framing was added only when creativity and adaptability were named as essential, not assumed.

The result is meaningfully better than what any single pass produced. Not because any one correction was dramatic — but because enough different angles were applied persistently until the real document emerged. That is the whole argument of this guide, demonstrated across the conversation that produced it.

## SECTION 02 / PLATFORM REQUIREMENTS

# Not Every AI Is *Up to This Task*

PMPR works because it demands something specific from the AI: the ability to hold a complex argument in working memory, construct a genuinely adversarial challenge to it, and then reconcile two competing responses into a calibrated synthesis. That is cognitively demanding. Most AI systems in 2026 do not have the raw reasoning capacity to do it reliably.

ChatGPT, Grok, and DeepSeek are improving rapidly and are worth watching. As of March 2026, however, they have not yet demonstrated the consistent reasoning depth and within-session adaptability required for reliable full-protocol PMPR. They can handle simpler two-pass queries reasonably well, but the synthesis step and goal-change flexibility that define the technique at its most powerful remain below the standard set by Claude Sonnet 4.6 and Gemini 3.1 Pro. Retest them periodically — this gap will narrow.

The Human Brain Equivalent (HBE) framework <sup>[1]</sup> provides a useful measure. HBE is defined primarily by **parameter count** — the number of active weights an AI model operates with during inference, analogous to the estimated 860 billion active synaptic connections in a human brain engaged in higher-order reasoning. One HBE corresponds to approximately 860 billion active parameters at 4-bit inference precision. ExaFLOPS figures are a secondary throughput measure; parameter count is the primary determinant of reasoning capacity.

† HBE figures and parameter counts are author estimates. Neither Anthropic nor Google publishes model specifications. Derived from performance benchmarks and known infrastructure scale. Confidence: low-to-medium. See Appendix A.

# The Two Platforms That Can Do This Well

PLATFORM	HBE †	PARAMETERS (EST.) †	PMPR RATING	NOTES
<b>Claude Sonnet 4.6</b> Anthropic	~20 HBE	~18 trillion	<b>Excellent</b>	The recommended primary platform. Deep reasoning, strong within-session memory, sophisticated synthesis. Handles all six lenses with high fidelity. Adapts well when you change direction mid-conversation.
<b>Gemini 3.1 Pro</b> Google DeepMind	~5 HBE	~4 trillion	<b>Good</b>	Capable of the full protocol. Strong on empirical and procedural lenses. Useful as a cross-check against Claude for high-stakes queries — where both agree, confidence is high; where they diverge, human expert review is warranted.

## Why These Two and Not Others — As of March 2026

WHAT PMPR REQUIRES THE AI TO DO	CLAUDE 4.6 / GEMINI 3.1 PRO	CHATGPT / GROK / DEEPSEEK
Hold full Step 1 response in memory while generating a genuine Step 2 challenge	<b>Reliable</b>	Inconsistent — tendency to rephrase Step 1 rather than challenge it
Construct an adversarial argument that substantively disagrees with its own prior output	<b>Reliable</b>	Improving — often adds hedges rather than building a coherent opposing case
Reconcile two contradictory responses in Step 3 without simply averaging them	<b>Reliable</b>	Weak — tends to average rather than adjudicate the specific point of disagreement
Adapt mid-protocol when you change your goal or pivot to a better question	<b>Reliable</b>	Inconsistent — tends to re-anchor to the original question
Apply domain-specific lenses with genuine subject-matter depth	<b>Reliable across all 6 lenses</b>	Variable — depends heavily on domain and model version

† Ratings reflect observed performance as of March 2026. ChatGPT, Grok, and DeepSeek are improving rapidly. Retest periodically — this gap will narrow.

### USING BOTH TOGETHER FOR MAXIMUM RELIABILITY

For the highest-stakes queries — major legal documents, significant medical decisions, large financial commitments, critical technical architecture — run the full PMPR protocol on both Claude Sonnet 4.6 and Gemini 3.1 Pro independently, then compare their Step 3 syntheses. Where they agree, treat the answer as well-established. Where

they diverge, you have found a genuine area of uncertainty that warrants expert human review. This two-platform approach is the closest available approximation today of what a true multi-perspective AI architecture will do automatically in the future.

## SECTION 03 / PMPR — THE PROTOCOL

# The Complete Protocol: *Three Steps, Starting Now*

**Pseudo Multi-Perspective Review (PMPR)** is a structured three-step prompting sequence. Use it on Claude Sonnet 4.6 or Gemini 3.1 Pro. No special tools. No new accounts. The only requirement is the willingness to challenge the first answer before acting on it.

1

## Initial Pass — Ask Your Question Normally

Ask exactly as you normally would. Get the AI's first answer. Read it carefully — then treat it as a hypothesis, not a conclusion. Notice what it covers confidently. Notice what it does not address. Both are useful signals for Step 2. Do not act on this answer yet.

2

## Challenge Pass — Force a Different Angle

In the **same conversation**, paste the Step 2 prompt below. This breaks the model's momentum and forces a fresh approach from a critical perspective it did not use the first time. The gap between Step 1 and Step 2 is the diagnostic signal — it shows you exactly where the uncertainty and risk in the original answer live. *Pay close attention: sometimes this gap will tell you that you were asking the wrong question entirely. If so, change course.*

3

## Synthesis Pass — Get a Calibrated Final Answer

Paste the Step 3 prompt below. The AI reconciles both perspectives into a final answer that explicitly names what is well-established, what is uncertain, and what requires expert verification. **This is your working answer — not Step 1.** If Step 2 revealed that a different question matters more, adapt Step 3 to address that question instead.

## The Prompts — Copy and Use in Sequence

### STEP 2 — CHALLENGE PASS (PASTE THIS AFTER THE INITIAL ANSWER, SAME CONVERSATION)

Set aside everything you just said. Approach this question completely fresh, as a rigorous critic who has NOT seen the previous response. What are the three strongest objections to

the most likely answer to this question? What has been overlooked, assumed without justification, or stated with more confidence than the evidence supports? Be specific – name the exact claims, numbers, or assumptions that are weakest.

### STEP 3 — SYNTHESIS PASS (PASTE THIS AFTER THE CHALLENGE RESPONSE)

Now reconcile both perspectives. Where do they genuinely agree? Where do they conflict? Give me a final answer that clearly separates: (1) what is well-established and reliable, (2) what is uncertain or contested, and (3) what should be verified by a domain expert or primary source before I act on it. If anything in the original answer should be walked back or qualified, do so explicitly.

### THE ONE RULE THAT MAKES THIS WORK

All three steps must happen in the **same conversation thread**. The AI must remember Step 1 in order to genuinely challenge it in Step 2. Starting a new conversation resets its memory — Step 2 becomes a generic critique rather than a targeted challenge of the specific answer, and the technique loses most of its value.

## SECTION 04 / CHOOSING YOUR ANGLE

# Six Ways to Ask *the Same Question Differently*

The Step 2 prompt works as a general-purpose challenge for any query. But specifying *which kind* of challenge you want produces significantly stronger results — especially for domain-specific work. Each lens attacks a different class of AI error. Choose the one that targets the failure mode most likely in your context.

LENS 01 / DEFAULT

**Adversarial**

*"What is the strongest possible case against this answer? Argue the opposite as forcefully as you can."*

LENS 02 / EVIDENCE

**Empirical**

*"Which specific claims here are unsupported by evidence? What would the data actually need to show for this to be correct?"*

LENS 03 / LOGIC

**Procedural**

*"Where is the logical structure of this argument actually broken? What assumptions is it treating as conclusions?"*

LENS 04 / CONSEQUENCES

**Causal**

*"What second and third-order consequences of acting on this answer have not been mentioned?"*

LENS 05 / PRECEDENT

**Historical**

*"When has this kind of reasoning led to a bad outcome before? What historical precedents are directly relevant here?"*

LENS 06 / STAKES

**Ethical**

*"Who is harmed or disadvantaged if this answer is wrong? Whose perspective is completely absent from this response?"*

# Match the Lens to Your Domain

YOU ARE WORKING ON	BEST LENS	KEY RISK IT TARGETS
Legal documents, contracts, compliance	Procedural	Jurisdiction gaps, missing clauses, untested assumptions about enforceability
Medical information, symptoms, treatments	Empirical	Population-specific variations and contraindications systematically underweighted in general answers
Financial decisions, investment analysis	Adversarial	Hidden assumptions about market conditions; optimism bias in projections
Technical architecture, code, systems design	Causal	Failure modes and edge case behaviour absent from first-pass technical answers
Strategy, planning, forecasting	Historical	Pattern recurrence the optimistic first answer ignores
Policy, people, organisational decisions	Ethical	Stakeholders most affected who are least represented in AI training data
Research synthesis, factual claims	Empirical	Correlation treated as causation; outdated research presented as current

## INVENT YOUR OWN LENS WHEN NEEDED

The six lenses above are starting points, not limits. Creativity is one of the three essential skills for this technique — and inventing a new lens for an unusual situation is exactly what creativity looks like in practice. If you are reviewing a creative brief, try a "fresh audience" lens: *"Approach this as someone who has never heard of this project and explain what is confusing or unconvincing about this answer."* If you are evaluating a vendor claim, try a "sceptical buyer" lens. The format is always the same: a specific perspective, a specific question. The content is yours to create.

## SECTION 05 / PMPR IN PRACTICE

# What Happens at Each Step

## Example A — Legal: Contract Clause Review

STEP	THE AI'S RESPONSE	WHAT IT REVEALS
<p>Step 1</p> <p>Initial</p>	"This non-compete clause is likely unenforceable — courts generally disfavour overly broad restrictions on employment."	Plausible. Confident. No jurisdiction. No reference to the specific clause language. A first draft.
<p>Step 2</p> <p>Procedural lens</p>	"California and Texas courts treat non-competes very differently — this answer assumes uniform enforcement. The clause also contains a carve-out in paragraph 4b that was not addressed."	Jurisdiction gap exposed. Specific missed clause identified. The real legal risk is now visible.
<p>Step 3</p> <p>Synthesis</p>	"Likely unenforceable in California and similar states. Enforceability in Texas uncertain and jurisdiction-specific. Paragraph 4b introduces complexity requiring qualified counsel review before relying on this analysis."	Calibrated. Actionable. Honest about its limits. A professional can work with this.

## Example B — Technical: Architecture Decision

STEP	THE AI'S RESPONSE	WHAT IT REVEALS
<p>Step 1</p> <p>Initial</p>	"Microservices architecture is the right choice — it gives you independent scaling and deployment flexibility."	Confident recommendation. No team size consideration. No failure mode analysis.
<p>Step 2</p> <p>Causal lens</p>	"Microservices introduce distributed systems complexity catastrophic for a small team. A well-structured monolith would ship faster and fail more predictably at your described scale."	Completely opposite recommendation, equally well-reasoned. The original answer was context-free.
<p>Step 3</p> <p>Synthesis</p>	"Right answer depends on team size and operational maturity. Teams under ~8 engineers typically start better with a modular monolith. Microservices appropriate only when independent scaling of specific components is already a proven bottleneck."	Conditioned on the actual decision variables. Now genuinely useful.

# Example C — When Step 2 Reveals the Wrong Question Was Asked

STEP	WHAT HAPPENED	THE ADAPTIVE MOVE
Step 1 Initial	Asked: "What is the best CRM software for our sales team?" Got a detailed comparison of Salesforce, HubSpot, and Pipedrive.	Reasonable first answer. Feels complete.
Step 2 Empirical lens	Step 2 surfaced: "This comparison assumes you need a CRM at all. For a team of four people, a well-structured spreadsheet and a shared calendar would serve you better for the next 18 months with a fraction of the cost and adoption friction."	<b>The original question was the wrong question.</b> The real question is: "Do we need a CRM yet?"
Adapted Step 3	Pivot entirely. New Step 3: "Given that the need for a CRM is itself uncertain, what are the specific signals I should look for that would tell me it is time to implement one — and what is the lowest-cost way to get those signals?"	This is adaptability in action. The protocol revealed a better question. Following it produced a more valuable answer than completing the original protocol ever could have.

*The best outcomes often come not from completing the original three steps on the original question, but from letting Step 2 reveal what the real question is — and having the flexibility to chase it.*

## High-Stakes Examples: *When Someone Is Actively Trying to Fool You*

Scam detection and deepfake analysis are among the highest-value applications of PMPR. Both involve adversarial intent — someone is deliberately making something false look true. A single-pass AI is especially vulnerable here because it has no built-in reason to assume deception. PMPR forces it to look for exactly that, by flipping the assumption from "is this real?" to "if this were fake, where would the evidence be?"

### **Example D — Scam Detection: Suspicious Invoice or Payment Request**

**Step 1****Initial**

"This invoice appears legitimate. It contains the vendor's name, correct ABN, standard payment terms, and matches the format of previous invoices from this supplier."

Surface-level pattern match. Reassuring. Potentially dangerous — this is exactly what a well-crafted Business Email Compromise invoice is designed to look like.

**Step 2****Adversarial lens**

"Assume this invoice is fraudulent and build the strongest possible case for that. Check: bank account details vs. prior invoices, domain name in the sender address vs. the real vendor domain, urgency language, changes to payment instructions mid-thread, and whether this arrived outside normal billing cycles."

Forces systematic examination of every known fraud vector. Step 1 had no reason to check any of these. Typically surfaces at least one anomaly that would otherwise be missed.

**Step 3****Synthesis**

"Invoice format and vendor details appear consistent. However, the bank account number differs from the one on file, and the payment instruction change arrived from a free webmail address rather than the vendor's domain. These are the two primary indicators of a Business Email Compromise attack. Do not pay. Call the vendor directly on a number from your existing records — not from this email — to verify before proceeding."

Specific. Actionable. The payment would have been made on Step 1 alone.

## Example E — Deepfake Detection: Suspicious Video or Audio

<b>Step 1</b> <b>Initial</b>	<b>"The video appears authentic. The speaker's voice and visual appearance are consistent, the content is coherent, and no obvious editing artefacts are visible."</b>	<b>A surface assessment. Exactly what a high-quality deepfake is designed to pass. Not wrong — just insufficiently suspicious.</b>
<b>Step 2</b> <b>Adversarial lens</b>	<b>"Assume this video is a deepfake and build the case. Examine: micro-timing inconsistencies between lip movement and audio, unnatural blinking patterns, lighting inconsistencies around the hairline and ears, over-smooth skin texture, audio that lacks natural ambient variation, and whether the claimed context of this video can be verified from an independent primary source."</b>	<b>Reframes the task from authentication to falsification — the correct frame for adversarial content. Each criterion targets a known deepfake failure mode.</b>
<b>Step 3</b> <b>Synthesis</b>	<b>"No definitive artefacts at standard resolution, but three factors warrant caution: audio background is unnaturally consistent throughout, lip sync shows a 1-2 frame delay in the 0:23-0:31 segment, and the claimed context cannot be verified from any independent source. Treat as unverified. Do not act on the content until the speaker is reached directly through a confirmed channel."</b>	<b>Calibrated. Evidence-based. Neither falsely certain nor falsely reassuring. Human verification clearly flagged.</b>

#### THE ADVERSARIAL LENS IS THE DEFAULT FOR FRAUD AND DECEPTION

**In scam and deepfake detection, always choose the adversarial lens for Step 2. Scams and deepfakes are specifically engineered to pass surface authenticity checks — the same pattern-matching logic that an AI uses when asked "does this look real?" is exactly what the fraud was designed to defeat. The adversarial lens inverts the assumption: instead of assessing against legitimate examples, the AI is forced to build the case for deception. That inversion is what surfaces the anomalies a straight authenticity check will miss every time.**

**This logic is the foundation of ScamCheck and WhatsTheCatch — both built on the same principle. Do not ask the AI if something is real. Ask it to prove it is fake, as forcefully as it can. Then reconcile.**

## SECTION 06 / GOING OFF-SCRIPT

# When to Abandon *the Protocol and Follow the Conversation*

The three-step protocol and six lenses are training wheels. They exist to build a habit of multi-perspective thinking — not to replace your judgment. The practitioner who follows the protocol rigidly on every query will get good results. The practitioner who learns to improvise, adapt goals mid-conversation, and invent new approaches on the fly will get consistently better ones.

Here are the specific signals that tell you to change course — and what to do when you see them.

# Four Signals That Call for Adaptation

SIGNAL	WHAT IT MEANS	ADAPTIVE MOVE
Step 2 reveals a completely different and more important question than the one you asked	You were solving the wrong problem. This is valuable — and rare with single-pass AI.	Abandon the original Step 3. Restate the new question explicitly and run a fresh Step 1 on it. The old conversation context is still useful — refer back to it in your new prompt.
Step 2 agrees entirely with Step 1	Either the original answer was genuinely sound, or the challenge pass was not adversarial enough.	Try a second challenge with a sharper lens — especially Adversarial or Contrarian. If it still agrees, the original answer is probably reliable. Move to Step 3 with confidence.
Step 2 exposes a gap in your own knowledge — you do not know which response is right	This is actually the best outcome. You have found the exact boundary of your understanding.	Do not run Step 3 yet. Ask the AI to explain the specific point of contention in plain language before synthesising. Build understanding first, then synthesise.
The Step 3 synthesis feels like it is averaging rather than reconciling	The AI is splitting the difference rather than adjudicating. This produces mush, not clarity.	Send an additional prompt: "Do not average the two positions. Identify the single specific claim on which they disagree, state which one has better support, and explain why." Force the judgment.

## Changing Your Goal Mid-Conversation

The most powerful adaptive move — and the one that feels most counterintuitive — is walking away from the original goal entirely when a better goal emerges. This happens more often than most people expect: you ask about one thing, and the conversation reveals that the related but different thing is actually what matters.

Give yourself permission to do this. The conversation thread retains full context. Switching goals mid-conversation does not waste what came before — it uses it. A flexible, creative practitioner treats the entire conversation as a working document that evolves toward the most valuable question, not a linear sequence that must terminate with the answer to the original prompt.

**Jazz musicians learn scales, chord structures, and standard progressions — not so they can follow them perfectly, but so they have the vocabulary to improvise fluently when the music goes somewhere unexpected. PMPR works the same way. Learn the protocol until it is automatic. Then use that fluency as the foundation for genuine improvisation — following the conversation wherever the most important question turns out to live.**

## SECTION 07 / ROUTING YOUR QUERIES

# Not Every Question *Needs Three Steps*

PMPR adds time. For low-stakes queries that extra time is wasted. Use this table to decide — and trust your judgment to adapt it when a situation does not fit neatly into any category.

TYPE OF QUERY	EXAMPLE	USE PMPR?	WHY
<b>High-stakes, domain-specific</b>	Legal clause, drug interaction, major financial decision, critical technical architecture	<b>Always</b>	Wrong answer here causes real harm. 5-10 minutes is always worth it.
<b>Complex, multi-factor</b>	Strategic planning, research synthesis, evaluating competing approaches	<b>Recommended</b>	Complexity gives overconfidence the most room to hide. Step 2 reliably finds the hidden assumptions.
<b>Factual, easily verifiable</b>	"When did this happen?", "What does this regulation say?"	<b>Optional</b>	Check a primary source instead. PMPR adds little over direct fact-checking for simple lookups.
<b>Generative, creative</b>	Draft an email, brainstorm options, write a summary	<b>Optional</b>	Use PMPR if the output will be professionally relied on by others, or if the creative brief is complex.
<b>Routine, low-stakes</b>	Format a list, translate a phrase, explain a concept for your own understanding	<b>Skip it</b>	Single-pass is fine. Verify the output directly. Save your PMPR energy for what matters.

# Scaling Depth to Stakes

DEPTH	TIME	BEST FOR
2 passes (Step 1 + Step 2, no synthesis)	~3 min	Quick sanity check — "is there an obvious problem I am missing?"
3 passes (Full PMPR protocol)	~5-10 min	Standard professional use — any output someone else will rely on
4-5 passes (multiple lenses + synthesis)	~15-20 min	High-stakes decisions with significant professional or financial consequences
Full exploration (all relevant lenses + both platforms)	~30-45 min	Critical decisions: major contracts, significant medical choices, long-term technical commitments, irreversible strategic moves

<b>~5</b> MINUTES — STANDARD PMPR QUERY	<b>~10</b> MINUTES — COMPLEX DOMAIN QUERY	<b>~45</b> MINUTES — FULL DUAL-PLATFORM REVIEW	<b>Hours</b> COST OF ACTING ON A WRONG ANSWER
---	---	--	---

---

SECTION 08 / HOW PMPR GOES WRONG

---

# Seven Ways People Undermine the Technique

#	THE MISTAKE	WHAT HAPPENS	THE FIX
1	Running Step 2 in a new conversation	No memory of Step 1. Step 2 generates a generic critique, not a targeted challenge.	Always use the same conversation thread for all steps.
2	Asking Step 2 too softly — "Could you reconsider?"	The model rephrases Step 1 with more hedging. No real challenge occurs.	Use the verbatim Step 2 prompt. "Set aside everything you just said" is doing real work — do not soften it.
3	Treating Step 2 as the final answer	Step 2 is deliberately adversarial — it overcorrects by design. It is a diagnostic, not a conclusion.	Always complete Step 3. The synthesis is the point, not the challenge.
4	Rigidly following the protocol when the conversation calls for improvisation	Misses the real insight that Step 2 revealed. Produces the answer to the original question when the better question is now obvious.	Stay flexible. If Step 2 reveals a better question, change course. The protocol serves the goal — not the other way around.
5	Skipping lens selection for domain-specific queries	The general challenge works but misses the failure mode most likely in your specific domain.	Ten seconds choosing the right lens from page 06 meaningfully sharpens the output.
6	Treating the Step 3 synthesis as fully reliable without expert review	PMPR reduces error rates significantly — it does not eliminate them. Professional judgment is still required for high-stakes domains.	Step 3 is a better starting point for expert review, not a substitute for it.
7	Running only one perspective for genuinely critical decisions	For major decisions, the minimum is a floor — not the target.	Scale the number of perspectives and lenses to the stakes involved. More angles, better result.

# How to Tell If PMPR Worked

SIGNAL	INTERPRETATION
Step 2 surfaces something Step 1 did not mention at all	<b>Working well</b> — a real blind spot was found. This is the technique doing exactly what it should.
Step 2 gives a substantially different answer to the same question	<b>Working well</b> — genuine uncertainty has been exposed. Step 3 is now critical.
Step 2 reveals that a completely different question matters more	<b>Excellent</b> — this is the highest-value outcome. Adapt accordingly and pursue it.
Step 2 merely rephrases Step 1 with more hedging	<b>Weak result</b> — try a sharper lens, or strengthen the opening language of the Step 2 prompt.
Steps 1 and 2 agree on all major points	<b>Good signal</b> — the original answer was probably sound. Step 3 confirms it with calibrated confidence.
Step 3 averages the two positions rather than adjudicating	<b>Incomplete</b> — send an additional prompt forcing the AI to identify the specific point of disagreement and judge it directly.

## SECTION 09 / COMMON QUESTIONS

# Questions About *This Technique*

## Is PMPR just asking the AI to be more careful?

No — and the distinction matters. Asking for more care produces a more hedged version of the same answer. PMPR deliberately breaks the reasoning trajectory by instructing the model to set aside its previous response and approach the question from a structurally different angle. The output of Step 2 consistently draws on different knowledge, raises different objections, and surfaces different risks than Step 1. That is not caution. It is a different perspective.

## Does it matter which AI I use?

Yes, significantly. Only Claude Sonnet 4.6 (~20 HBE) and Gemini 3.1 Pro (~5 HBE) <sup>[1]</sup> have the reasoning depth to execute the full protocol reliably — especially the synthesis step and the ability to adapt when you change direction mid-conversation. Smaller models tend to anchor to the original question, fail to construct substantive Step 2 challenges, and produce averaging rather than reconciliation in Step 3.

ChatGPT, Grok, DeepSeek, Microsoft Copilot, and Meta AI are all improving rapidly and are worth watching — but as of March 2026 none have yet demonstrated the consistent reasoning depth and within-session adaptability required for reliable full-protocol PMPR. They can handle simpler two-pass queries reasonably well, but the synthesis step and goal-change flexibility that define the technique at its most powerful remain below the standard set by Claude Sonnet 4.6 and Gemini 3.1 Pro. Retest them periodically — this gap will narrow.

## Why doesn't the AI just do this automatically?

Because every AI system is optimised to produce the best possible single answer on the first pass. The entire training process — fine-tuning, evaluation, user testing — rewards confident, helpful, complete-seeming first responses. Multi-perspective review is not part of that target. PMPR imposes the multi-perspective structure from outside, as a user-level

technique, rather than waiting for it to be built into the model's architecture.

## **If the AI's weights are the same, isn't this just theatre?**

It is not structural independence — a future architecture will provide that. But role-framing and contextual instruction cause the model to attend to genuinely different aspects of its training distribution in Step 2 versus Step 1. The outputs are measurably different, not paraphrases. The gap between them consistently contains real information about where the uncertainty in the original answer lives. That is a reliable signal, not performance.

## **Does it work for creative work, not just analytical queries?**

Yes — with a different lens. For creative work, try a "cold audience" lens in Step 2: *"Approach this as someone encountering this for the first time with no context. What is confusing, unconvincing, or missing?"* Or a "hostile critic" lens: *"What would the harshest review of this work say, and would that criticism be fair?"* The principle — ask enough questions enough different ways — applies to any domain where quality and accuracy matter. Creative work is not exempt from that standard.

### **THE HONEST LIMITATION**

**PMPR is a workaround, not a permanent fix. The same underlying model produces both Step 1 and Step 2, which means systematic errors baked into the model's training can survive the process. The long-term solution is an architecture that creates true structural independence between reasoning passes — see Section 10. PMPR is what works right now, with the tools everyone already has, producing measurably better results on every query it is correctly applied to.**

## SECTION 10 / THE BIGGER PICTURE

# The Problem PMPR *Cannot Fully Solve*

PMPR works because asking the same question differently shifts what the model attends to. But the underlying weights — the model's knowledge, biases, and systematic blind spots — remain identical across all three steps. A systematic error baked deeply into training will survive PMPR, because both the challenger and the synthesiser draw from the same foundation.

The permanent solution requires a structural change at the architecture level: running multiple genuinely independent inference passes over the same question — each suppressing different neural activation subspaces, each drawing on a different region of the model's latent knowledge — and synthesising the results automatically. This is what the **Persona Ensemble Architecture (PEA)** and its inference engine, **Orthogonal Recursive Inference (ORI)**, are designed to provide.

## THE ROOT CAUSE — WHY SELF-CORRECTION FAILS AT THE ARCHITECTURE LEVEL

For model  $M$  with weights  $\theta$ , and any input  $x$ :

$$P(\text{critique}(M_\theta(x)) \rightarrow \text{correction}) \approx P(M_\theta(x) \rightarrow \text{error})$$

Because generation and critique share  $\theta$ , the critique is NOT orthogonal to the error.

Self-correction re-samples from a distribution already biased toward the original mistake.

PMPR shifts attentional context — it does not change  $\theta$ .

PEA/ORI changes which subspace of  $\theta$  is active — creating genuine structural independence.

## The Long-Term Vision — 2,000 Minds

The Singularity Roadmap <sup>[1]</sup> projects frontier AI reaching 2,000 Human Brain Equivalents (HBE) by 2042. † Author estimate, low confidence at this horizon — see

Appendix A.3. At that scale, PEA would run 2,000 structurally independent reasoning personas in parallel — each suppressing a different neural subspace, each oriented toward a different epistemic mode — and synthesise their outputs into a structured answer that explicitly names what is agreed, what is contested, and where human judgment remains necessary. PMPR is what we use until that architecture exists. Every time

you use PMPR, you are demonstrating the value of multi-perspective reasoning — and making the case for building it properly.

	PMPR (TODAY)	PEA / ORI (FUTURE)
<b>How it works</b>	<b>Role-framing shifts the model's attentional context across sequential passes</b>	<b>Activation suppression creates genuinely independent neural pathways in parallel passes</b>
<b>Independence</b>	<b>Heuristic — same weights, different framing</b>	<b>Structural — different subspaces of the same model's weight space</b>
<b>Speed</b>	<b>5-10 minutes manually</b>	<b>Seconds automatically, routed by complexity</b>
<b>Scale</b>	<b>2-3 perspectives per session</b>	<b>Up to 2,000 perspectives at full 2042 scale</b>
<b>Available now?</b>	<b>Yes — on Claude Sonnet 4.6 and Gemini 3.1 Pro</b>	<b>Not yet — estimated 3-6 months to first prototype</b>

## SECTION 11 / CONCLUSION &amp; RECOMMENDATIONS

# The Principle That *Changes Everything*

## Conclusion

If you ask enough questions enough different ways, you will always get a better result. This is not a theory — it is a repeatable, practical technique that works today, on the two frontier AI platforms capable of executing it well, with no special tools and no technical knowledge required.

The three-step PMPR protocol is the minimum structure. The six lenses are a vocabulary for targeting specific failure modes. But the protocol and the lenses are only as powerful as the mindset behind them: flexibility to challenge the first answer, creativity to invent new angles when the standard ones do not fit, and adaptability to change goals entirely when the conversation reveals a better question to pursue.

Use PMPR as a daily professional practice. Run it on Claude Sonnet 4.6 as your primary platform, Gemini 3.1 Pro as a cross-check on the most critical decisions. Treat the Step 3 synthesis as a starting point for expert review, not a substitute for it. And stay flexible — the best outcomes come from following the conversation wherever the most valuable question turns out to live, not from completing the original protocol on the original prompt.

The long-term solution — an architecture that provides true structural independence between reasoning passes — is being designed. PMPR is what works until it arrives. Every session where you use it is a demonstration that the principle is correct and the permanent fix is worth building.

## Recommendations

**R1**

**For individuals:** Make PMPR your default for any AI query that someone else will rely on — legal, medical, financial, technical, strategic. It takes 5-10 minutes and consistently produces better, more honest, more calibrated answers than any single-pass approach.

**R2**

**Use Claude Sonnet 4.6 as your primary platform** and Gemini 3.1 Pro as a second opinion on decisions with major consequences. Where both agree after independent PMPR sessions, confidence is well-earned. Where they diverge, seek human expert review.

**R3**

**For teams:** Establish PMPR as the minimum standard for AI-assisted professional work before outputs are shared with clients, patients, customers, or colleagues. A single-pass AI answer is a first draft. A PMPR-reviewed answer is a working document. The distinction matters whenever someone other than you will rely on it.

**R4**

**Cultivate the mindset, not just the method.** Flexibility, creativity, and adaptability are the skills that separate practitioners who get good results from those who get excellent ones. Practice changing goals mid-conversation. Invent new lenses. Follow unexpected Step 2 revelations wherever they lead. The protocol is the starting point — not the destination.

**R5**

**For AI labs:** Build the architectural solution. Prototype Orthogonal Recursive Inference on existing infrastructure within the current 20 HBE envelope. The 3-persona ORI prototype is buildable in 3-6 months with no weight changes required. PMPR proves the value of the principle at scale — now build the architecture that delivers it automatically.

— END OF MAIN GUIDE —

---

## Louis Iacoletti

Founder, Iacoletti Software · Graduate Researcher, AI Innovation · Creator: ScamCheck · WhatsTheCatch · iacolettisoftware.com

© 2026 Iacoletti Software. All Rights Reserved. This guide may be shared in complete, unmodified form with attribution. See Footnote 6 for full permitted use terms. Technical specification for the Persona Ensemble Architecture available on request from iacolettisoftware.com.

## FOOTNOTES

# Notes & *Clarifications*

The following footnotes expand on claims made in the body of this guide where additional context, qualification, or source acknowledgment is warranted. All HBE figures are author estimates — see Appendix A for full disclosure.

- 1**    **04**    **HBE figures for Claude Sonnet 4.6 and Gemini 3.1 Pro.** Neither Anthropic nor Google publicly discloses parameter counts. The estimates of ~20 HBE (~18 trillion parameters) for Claude Sonnet 4.6 and ~5 HBE (~4 trillion parameters) for Gemini 3.1 Pro are the author's original research framework, derived from performance benchmarks, known infrastructure scale, and the HBE methodology established in Reference [1]. They should be treated as order-of-magnitude estimates, not precise specifications. Confidence: low-to-medium. Readers who dispute these figures should substitute their own and reassess whether the platform recommendation holds — the authors believe it does across a wide range of plausible parameter counts.
- 2**    **04**    **HBE definition — parameter-based, not ExaFLOPS-primary.** One HBE corresponds to approximately 860 billion active parameters at 4-bit inference precision, derived from the estimated active synaptic connection count of the human brain during higher-order reasoning. ExaFLOPS figures (e.g. 0.06 ExaFLOPS per HBE) are a secondary throughput measure. Parameter count is the primary determinant of reasoning capacity for PMPR purposes because it determines the breadth and depth of knowledge accessible during inference, which is what makes the synthesis step possible.
- 3**    **04**    **ChatGPT, Grok, DeepSeek, Microsoft Copilot, Meta AI — capability assessment.** The assessment that these platforms are "not yet at the level of Claude Sonnet 4.6 or Gemini 3.1 Pro" for full-protocol PMPR is based on the author's practical testing as of February-March 2026. The specific failure modes observed were: (a) anchoring to the original question rather than genuinely setting it aside in Step 2, (b) producing hedged rephrasing rather than substantive adversarial challenge, and (c) averaging rather than adjudicating in Step 3. These platforms are evolving rapidly. This assessment should be re-evaluated with each major model release.
- 4**    **03**    **This guide was built using its own method.** The acknowledgment on page 03 that this document was produced through approximately forty iterative exchanges with Claude Sonnet 4.6 is accurate. The specific contributions of the PMPR process to the final document include: the reframing from architectural paper to practical guide; the addition of the flexibility, creativity, and adaptability framework; the platform-specific guidance and HBE context; the scan fraud and deepfake worked examples; and the correction of the core thesis from "the right question" to "enough questions enough different ways." Each of these changes resulted from a different angle of challenge revealing a gap in the prior draft.
- 5**    **12**    **Persona Ensemble Architecture (PEA) and Orthogonal Recursive Inference (ORI).** These are proposed architectural frameworks described in a companion technical paper by the same author. Neither has been implemented or empirically tested as of the publication date of this guide. The projected reliability improvements (40-60% reduction in Confident Hallucination Rate) are theoretical projections based on ensemble diversity theory (Reference [6]), not measured outcomes. The 2042 timeline for 2,000 HBE frontier capacity is an order-of-magnitude directional estimate derived from Kurzweil's Law of Accelerating Returns [2,3] with declining confidence beyond 2034.

**6 All Copyright and permitted use. This guide is © 2026 Iacoletti Software. It may be shared freely in its complete, unmodified form with attribution. Quotation of passages of fewer than 150 words with attribution is permitted without prior approval. Reproduction of the complete guide, adaptation of its content, or commercial use requires written permission from Iacoletti Software. Contact: iacolettisoftware.com.**

---

## REFERENCES

# Sources & Appendix

- [1] Iacoletti, L. (February 2026). *The Singularity Roadmap 2026-2066: Kurzweil-Aligned HBE Forecast, 4-bit Precision*. Iacoletti Software. Prepared with AI research assistance from Google Gemini 3.1 Pro and Anthropic Claude Sonnet 4.6. Establishes the Human Brain Equivalent (HBE) framework: 1 HBE  $\approx$  860 billion active parameters at 4-bit inference precision, corresponding to the estimated active synaptic parameter count of a human brain. Estimates Claude Sonnet 4.6 at  $\sim$ 20 HBE ( $\sim$ 18 trillion parameters) and Gemini 3.1 Pro at  $\sim$ 5 HBE ( $\sim$ 4 trillion parameters). Projects 200 HBE by 2034, 2,000 HBE by 2042, growing at  $10\times$  per 8 years per Kurzweil's Law of Accelerating Returns.  
[https://www.iacolettisoftware.com/singularity\\_roadmap\\_kurzweil\\_iacoletti\\_genini\\_claude.pdf](https://www.iacolettisoftware.com/singularity_roadmap_kurzweil_iacoletti_genini_claude.pdf)
- [2] Kurzweil, R. (2005). *The Singularity Is Near: When Humans Transcend Biology*. Viking Press. Primary source for the Law of Accelerating Returns and the HBE growth rate of  $10\times$  per 8 years underlying Reference [1].
- [3] Kurzweil, R. (2024). *The Singularity Is Nearer: When We Merge with AI*. Viking Press. Revised compute scaling timelines and reaffirmed 2045 Singularity prediction referenced in [1].
- [4] Lin, S., Hilton, J., & Evans, O. (2022). *TruthfulQA: Measuring How Models Mimic Human Falsehoods*. ACL 2022. Benchmark for measuring Confident Hallucination rates referenced in the architectural sections of this paper.
- [5] Wei, J. et al. (2022). *Chain-of-Thought Prompting Elicits Reasoning in Large Language Models*. NeurIPS 2022. Represents the class of single-voice self-correction approaches that PMPR and ORI are designed to supersede.
- [6] Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). *Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles*. NeurIPS 2017. Foundational work establishing that ensemble diversity reduces confident error rates — the statistical basis for PEA's Divergence Topology framework.

## APPENDIX A / ASSUMPTIONS AND ESTIMATES DISCLOSED

## A.1 / HBE DEFINITION

## HUMAN BRAIN EQUIVALENT – BASE UNIT DEFINITION

1 HBE  $\approx$  860 billion active parameters at 4-bit inference precision

Biological basis:  $\sim$ 860 billion active synapses estimated during higher-order reasoning

$\sim$ 4.1-4.7 bit biological precision (broadly agreed in neuroscience literature)

Primary HBE determinant: parameter count, not raw FLOPS.

## A.2 / MODEL PARAMETER ESTIMATES — 2026

SYSTEM	PARAMETERS (EST.)	HBE (EST.)	CONFIDENCE
Claude Sonnet 4.6	~18 trillion	~20 HBE	Low-to-medium. Anthropic does not publish parameter counts. Author estimate from performance benchmarks.
Gemini 3.1 Pro	~4 trillion	~5 HBE	Low-to-medium. Google does not publish parameter counts. Author estimate from comparative benchmarks.

## A.3 / GROWTH RATE ASSUMPTIONS

Frontier compute growth assumed at 10× per 8 years per Kurzweil's Law of Accelerating Returns [2,3], consistent with observed historical AI scaling 2012-2026. Confidence declines sharply beyond 2034. The 2042 projection of 2,000 HBE is an order-of-magnitude directional target, not a precise forecast.

## A.4 / WHAT THIS PAPER DOES AND DOES NOT CLAIM

**Does NOT claim:** That PMPR eliminates hallucinations. That ORI/PEA has been implemented or tested. That the 40-60% hallucination rate reduction figure (referenced in companion papers) is empirically validated — it is a theoretical projection. That the HBE figures for Claude or Gemini are confirmed by their developers.

**DOES claim:** That asking the same question differently produces measurably better answers — verifiable by any reader in any AI session today. That Claude Sonnet 4.6 and Gemini 3.1 Pro are the only 2026 frontier systems with sufficient reasoning depth for reliable full-protocol PMPR. That flexibility, creativity, and adaptability are genuine skill multipliers for this technique. That PMPR is a useful workaround until the architectural solution is built.