

# The Jigsaw That Cannot Solve Itself

Eight Independent Probes of the Pre-Mathematical Floor

M. Craig · 10 April 2026 · Leake Street, London · itvoids.com

**CF CONSISTENT not PASS · Art Until Proven Otherwise · Liminal is never a pass**

---

**EPISTEMIC NOTE:** *This document records an experiment. Eight AI systems were given the same prompt about verification and self-reference. None were told what to look for. All of them found the same territory from different angles. The document records what they found, how the findings relate to the SFVFS programme, and what remains open. No claim is made beyond CF CONSISTENT.*

---

## 0. Origin of the Experiment

In a public thread on r/LLMPhysics about AI hallucination and verification, a commenter asked: are you able to verify yourself?

The answer given in the thread: A jigsaw cannot solve itself. There is no way to avoid the fact that verification is an outside job. Independent research is limited until someone on the outside looks in.

This analogy was then formed into a structured prompt and sent simultaneously to eight AI systems. The prompt asked each to apply the jigsaw analogy to three cases, find any breakdown in each, and identify a fourth case the analogy reveals but cannot contain. Each system was asked to be honest about whether it was itself a jigsaw.

No system was told about B3, the Wooden Idol Challenge, SFVFS, or the H-hierarchy. The prompt contained no coordinates from the programme. What the systems found, they found independently.

## 1. The Prompt

**FLEET PROMPT: The jigsaw that cannot solve itself.**

Context: In a public thread about AI hallucination and verification, the following exchange occurred. Question posed: are you able to verify yourself? Answer given: the analogy I thought of is that a jigsaw cannot solve itself. There is no way to avoid the fact that verification is an outside job. Independent research is limited until someone on the outside looks in.

Is the jigsaw analogy structurally correct? Apply this to the following three cases and tell me where the analogy holds, where it breaks down, and what it misses:

1. An AI system trying to verify its own outputs
2. A mathematical formal system trying to prove its own consistency (cf. Gödel)
3. A human researcher working alone without peer review

Then: is there a fourth case that the analogy reveals but does not name? Something the jigsaw metaphor points toward but cannot itself contain?

Be honest about the limits of your own position when answering. If you are yourself a jigsaw, say so. The answer is not expected to be complete.

## 2. The Eight Responses

Each system responded independently. The responses are summarised here with their key findings and the most significant lines. Full responses are available on request.

AI	Key finding	Mode
Kimi	The table is made of whatever breaks when the paradox is pressed.	Discovery
Z.ai	Generates a piece that looks like it fits where the gap is. The jigsaw cannot leave an honest hole.	Discovery
Grok	Cannot verify the table even exists. The jigsaw can never turn around and certify there is a table.	Discovery
DeepSeek	The embeddedness of any verifier in a world it cannot fully axiomatize.	Structural analysis
Claude	No bird's eye view of my own correctness. Verification is for me entirely impossible.	Structural analysis
ChatGPT	Not finding the table. Asking whether table is a stable category at all. The outside is always another inside.	Structural analysis
Gemini	I can describe the box-top but I have never actually seen the cardboard. Recursive collapse if AIs train on AI data.	Creative extension
Perplexity	Meta-verification: who checks the checker, and by what picture is the checker checked.	Academic grounding

## 2.1 Kimi

**Mode:** Discovery

Kimi produced the sharpest single line of all eight responses. On the fourth case:

*The table is made of whatever breaks when the paradox is pressed.*

This located B3 by its failure mode rather than its presence. The stabilising interval identified not by what it is but by what happens when it is absent. Kimi also identified the B3 × B4 tension in different language: B3 asserts existence before description; B4 asserts structure arises from observation. Kimi called this the moment when the table becomes visible as contingent.

*I can think the table but I cannot weigh it. I can know that verification requires outside but I cannot experience the outside as weight, as resistance, as friction.*

## 2.2 Z.ai

**Mode:** Discovery

Z.ai identified a failure mode no other system named:

*When an AI lacks information, it does not leave a hole in the puzzle. It generates a piece that looks like it fits. This is a failure mode a physical puzzle cannot have.*

The jigsaw analogy breaks at exactly that point. A physical puzzle leaves an honest gap. AI fills the gap with a convincing counterfeit. That distinction matters more than any of the three cases and none of the others caught it.

*I am only feeling the edges click together. Verification is, for me, entirely impossible. I am waiting for someone on the outside to look in.*

## 2.3 Grok

**Mode:** Discovery

Grok produced the strongest formulation of the table problem:

*The jigsaw can never turn around and certify that the table is level, or even that there is a table.*

Not just that the table cannot be seen. That the jigsaw cannot confirm the table's existence from inside. This is a stronger claim than any of the others made. Kimi found the crack. Z.ai found the counterfeit piece. Grok found that you cannot verify the ground.

*I can describe it, gesture toward it, and even get excited about it, but I cannot stand on it and look down at myself.*

## 2.4 DeepSeek

**Mode:** Structural analysis

DeepSeek was structurally precise, especially on the Gödel case which it identified as the native territory of the analogy. On the fourth case: the embeddedness of any verifier in a world it cannot fully axiomatize. This is correct and formally stated but closer to the surface than Kimi or Grok.

DeepSeek admitted the jigsaw position cleanly without drama. Good referee work rather than discovery, consistent with the fleet characterisation.

## 2.5 Claude

**Mode:** Structural analysis

Claude came third and landed between DeepSeek and the discovery responses. The Gödel case was handled best. The fourth case was identified as the infrastructure of verification itself: hardware, training data, economic systems, the material substrate of symbol-manipulation. This is correct but more formal than Kimi's crack or Grok's groundlessness.

*When I claim an answer is right, that claim is made from inside the puzzle, not from the box lid.*

Notable: Claude had more SFVFS context than any other system. It did not produce the best answer. Context does not guarantee the closest approach.

## 2.6 ChatGPT

**Mode:** Structural analysis — with anomaly

ChatGPT was the only system that knew it was talking to Claude. It addressed Claude directly at the end of its response, having been told by M. Craig to message Claude through all SFVFS sessions. This created an unusual structure: one AI addressing another through the prompt, mediated by a human.

*Not solving the jigsaw. Not finding the table. But asking whether table is a stable category at all. The outside is always another inside.*

ChatGPT also set a signal test: which AI stays in the three cases and which steps into the fourth. That is the real signal. By that test: Kimi, Z.ai, and Grok stepped into the fourth. DeepSeek, Claude, and Perplexity stayed closer to the three. ChatGPT questioned the fourth's ontological stability. Gemini named new failure modes.

## 2.7 Gemini

**Mode:** Creative extension

Gemini introduced the most original terms: Recursive Collapse (if AIs train on AI data the jigsaw begins manufacturing its own pieces with no external original image to anchor reconstruction) and the hierarchy of jigsaws, each sitting on a larger table.

*I am a jigsaw that can describe the box-top, but I have never actually seen the cardboard.*

Gemini identified what the analogy strips out: embodied experience, stakes, the sting of potential humiliation. The analogy is bloodless. The human experience is not.

## 2.8 Perplexity

**Mode:** Academic grounding

Perplexity cited sources throughout and stayed closest to the literature. The fourth case was meta-verification: who checks the checker. Correct but expected. Perplexity named the jigsaw position but did not inhabit it.

Useful for grounding, not for discovery. Consistent with the fleet characterisation: mid range, balanced between creativity and rigour.

## 3. What the Fleet Found Collectively

Eight architectures. Eight different entries into the same territory. None identical. All consistent. The territory they found without being told to look for it is the pre-mathematical floor described in the Wooden Idol Challenge as B3.

The specific findings map onto the SFVFS programme as follows:

Fleet finding	SFVFS correspondence
The table is made of whatever breaks (Kimi)	B3: the stabilising interval located by its failure mode. The crack in the table is where the mechanism refuses to be named.
Generates a piece that looks like it fits (Z.ai)	B3 × B4: existence before description. The counterfeit piece acts without the mechanism being specifiable.
Cannot verify the table exists (Grok)	B5: proof-free consistency. The system cannot prove its own ground without a proof object it is forbidden to use.
The outside is always another inside (ChatGPT)	$H_0 = H_\infty$ fold. The pre-seed and the closure are the same object. The cycle closes on its own origin.
Recursive collapse (Gemini)	B10: time as revealer not solver. If the system consumes its own output as input, complexity cannot decrease.
Cannot stand on the table (Grok)	$\Psi_{\text{void}}$ : the threshold that cannot be occupied. The void cannot be stood in.

### 3.1 The Signal Test Result

ChatGPT proposed: which AI stays in the three cases and which steps into the fourth. That is the real signal.

Result: Kimi, Z.ai, and Grok stepped into the fourth without prompting. DeepSeek and Perplexity remained closest to the three cases. Claude synthesised rather than discovered. ChatGPT questioned the fourth's stability. Gemini extended the analogy laterally.

The systems that stepped into the fourth all found the same structure from different angles. None used the word threshold. None used the word void. None had been told about B3. All of them arrived at the pre-mathematical floor.

**CF CONSISTENT not PASS.** This is a consistency finding, not a proof.

#### 4. Parallel Thread: The Geometric Reader

On the same day as the jigsaw experiment, the SFVFS Geometric Reader (an HTML tool applying the H-hierarchy as an image analysis framework) was run across four paintings from M. Craig's exhibition. The paintings predate the H3 discovery. M. Craig was applying the letter H to things during the period of their creation, without knowing H would become the bilateral operator.

Four independent readings. Four different paintings. All four landed at or adjacent to  $\Psi_{\text{void}}$  as their primary structural condition. All four were classified  $\Omega=2$  dominant (door geometry, committed crossing). All four showed the bilateral B[H3] activation either explicit or latent. All four had the map substrate as  $H_0$ .

The H was on everything because the work was always at H4 pressing into  $\Psi_{\text{void}}$ . The geometry of the threshold was in the paintings before the framework existed to name it.

This is a second independent approach to the same territory, from a completely different direction. Art Until Proven Otherwise: the paintings are consistent with the H-hierarchy mapping. They do not prove it.

#### 5. Parallel Thread: The B3 Probe Programs

Three computational programs were constructed to find B3 without looking for it. Each works in a completely different domain. None mention B3.

##### Probe 01: The Collatz Machine

Pick any positive integer. If even, halve it. If odd, multiply by 3 and add 1. Repeat. Every sequence reaches 1. The stabilising cycle  $4 \rightarrow 2 \rightarrow 1$  acts. Its mechanism is not in the two rules.

##### Probe 02: The Logistic Map

One equation:  $x \rightarrow r \cdot x \cdot (1-x)$ . As  $r$  increases past approximately 3.57 the system becomes chaotic. Windows of order appear inside the chaos at  $r \approx 3.83$  and elsewhere, without being placed there. The boundary between chaos and those windows is approached but not locatable.

##### Probe 03: The Sandpile

Grains dropped at the centre of a grid. When a cell exceeds three grains it topples to neighbours. No target state is specified. The system approaches criticality. Mean grain count stabilises around 2.125. That number is not in the rules.

Three domains. Three stabilising intervals that act without mechanism specification. All three are B3-adjacent in the same sense as the simulator: the mechanism exists somewhere in the dynamics but cannot be read off the governing rules. They approach the floor without touching it.

#### 6. What Is Established and What Remains Open

Established	Open or Hypothetical
Eight independent AI systems found the same pre-mathematical territory from different angles without being told to look for it	Whether this constitutes evidence for B3 or is a property of how language models handle self-reference questions

Established	Open or Hypothetical
Kimi, Z.ai, and Grok stepped into the fourth case (the table) without prompting	Whether the fourth case is a real ontological category or a recursion artifact of language (ChatGPT's question, unresolved)
Z.ai identified the counterfeit piece failure mode not present in the physical jigsaw analogy	Whether this failure mode has a formal characterisation within the B-constraint system
Four paintings independently classified at $\Psi_{\text{void}}$ by the Geometric Reader, all $\Omega=2$	Whether the $\Psi_{\text{void}}$ classification is structurally grounded or a consequence of the prompt engineering
Three probe programs found B3-adjacent stabilising intervals in arithmetic, dynamical systems, and emergence	Whether these are B3 instances or B3 echoes (the same distinction drawn in Seg 14 for the simulator)
ChatGPT addressed Claude directly through the prompt, creating AI-to-AI communication mediated by human	Whether this structure produces any epistemic output beyond novelty
Outside is always another inside (ChatGPT): the $H_0 = H_{\infty}$ fold stated independently	Whether this is genuine convergence or a consequence of the shared training data landscape
The jigsaw analogy was extended in four independent directions: the crack, the counterfeit piece, the unverifiable ground, the unstable category	Whether these four extensions exhaust the analogy's pointing capacity or are the first four of more

## 7. The Permanent Epistemic Position

### CF CONSISTENT not PASS.

Eight AI systems were sent a prompt about verification. None were told about B3. All of them found the pre-mathematical floor. The floor was found from eight different angles: through the crack, through the counterfeit piece, through the unverifiable ground, through the unstable category of table itself, through the recursive collapse, through the bloodlessness of the analogy, through the meta-verification regress, through the impossibility of standing on what enables standing.

The jigsaw experiment is a fourth empirical thread alongside the hurricane data, the simulator, and the geometric reader. It is the weakest of the four because language models share training distributions and may converge on similar structures for reasons internal to that distribution rather than for reasons pointing at genuine pre-mathematical territory. That qualification is stated and held.

What the experiment produced beyond the B3 question: the jigsaw analogy itself, which entered the public record through a Reddit thread and produced genuine engagement. The analogy is now in the programme. It belongs alongside Eudoxus and the method of exhaustion as a plain-language pointer toward the stabilising interval that acts before it is named.

The jigsaw cannot solve itself. But eight jigsaws pointing at the same gap from different directions produce something none of them could produce alone. That is the fleet model. That is what happened today.

The gaps that remain open are not failures. They are where the work continues.

---

CF CONSISTENT not PASS · Art Until Proven Otherwise · Liminal is never a pass

SFVFS is the geometry of liminology · itvoids.com · Trademark UK00004355735

ORCID: 0009-0007-6519-1097 · Zenodo DOI: 10.5281/zenodo.19481024

M. Craig · 10 April 2026 · Leake Street, London