

# Envatment as a Metaphysical Hypothesis

---

David J. Chalmers

# Brains in Vats

---

- A brain in a vat, connected to a computer simulation of the world, might have the same experiences that I do.
  - The brain in a vat has beliefs corresponding to mine.
    - *I have hands*
    - *There is a table in front of me.*
    - *I am in Tucson.*
  - These beliefs are massively false.
  - Cf: deluded beings in *The Matrix*
-

**TOM the  
DANCING  
BUG**

BY  
RUBEN  
BOLLING

# THE OUTER REACHES OF PLOT TWISTS

*Please don't tell your  
friends the shocking  
beginning, middle and  
ending!*

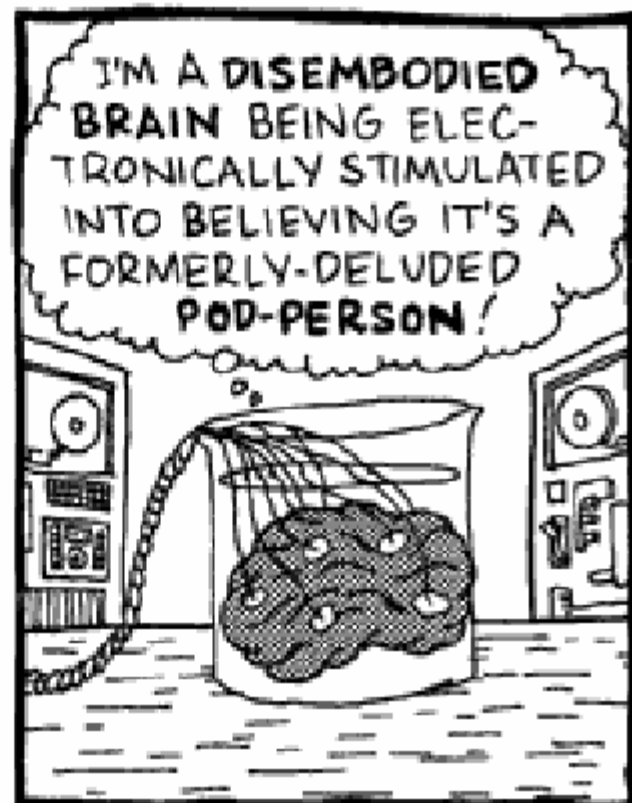


SUBMITTED  
FOR YOUR  
APPROVAL:  
WES SHELBY.  
AN ORDINARY  
MAN ABOUT TO  
UNDERGO A  
RECORD-SETTING  
NUMBER OF  
REALITY-  
BENDING  
AND UTTERLY  
SHOCKING  
PLOT TWISTS!

SOMETHING ABOUT MY LIFE  
FEELS VERY SPOOKY, SLOWLY  
PACED AND POORLY LIT.

BUT  
WHAT  
IS IT?





BUT... COULD THERE BE YET ANOTHER STARTLING REVELATION? AH, IT'S OBVIOUS NOW--

I'M DEAD!

SHOCKING? YES, BUT NOW THIS GHOST-BRAIN MUST WAKE UP TO THE REALIZATION THAT ITS FALSE-REALITY SHOCKS WERE ACTUALLY MERE ILLUSIONS...!

CONJURED FIGMENTS OF A HACK SCREEN-WRITER'S LIMITED IMAGINATION!

GOOD LORD! I'M FICTIONAL!

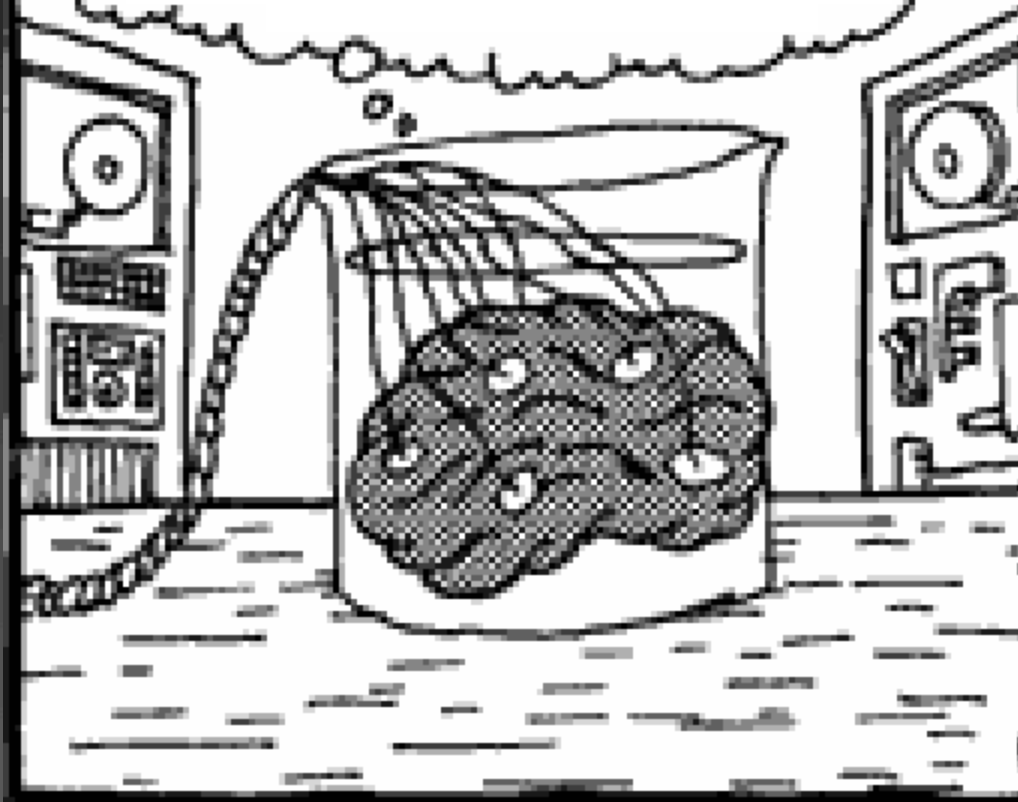
THE END?

# Could I Be a Brain in a Vat?

---

- I cannot rule out the hypothesis that I am a brain in a vat, connected to a computer simulation.
  - If am a brain in a vat, then I have massively false beliefs:
    - I do not have hands
    - There is no table in front of me
    - I am not in Tucson
-

I'm giving a talk  
in Tucson, waving  
my hands over a  
table.



# Skeptical Hypothesis

---

- (1) I don't know that I am not a brain in a vat.
  - (2) If I am a brain in a vat, I don't have hands.
  - So (3): I don't know that I have hands. [by closure]
  
  - More generally: I don't know much at all.
  
  - The hypothesis that I am envatted is a **skeptical hypothesis**
    - Its epistemic possibility casts doubt on my beliefs, and undercuts their claim to be knowledge.
-



# Putnam's Response

---

- A brain in a vat could not think the thought *I am a brain in a vat*.
    - Its corresponding concept ( 'brain' ) is not causally connected to brains, so it does not refer to brains.
  - Cf: Twin Earth, where the watery stuff is XYZ
    - My twin's 'water' concept does not refer to water, but to twater.
  - So if I am a BIV, I can't truly think *I am a BIV*.
  - So: I am not a BIV.
-

# Red Herring

---

- I think: Putnam's point is correct, but a red herring.
    - I can rule out the hypothesis that I am (literally) a brain in a vat.
    - But I cannot rule out the hypothesis that I am *relevantly like* a brain in a vat.
  - We just need to restate the hypothesis in **semantically neutral** language
    - Language not vulnerable to Twin Earth thought experiments
    - Cf. 'I am a philosopher' .
-

# Envatment Reformulated

---

- Let's say: 'x is envatted' =  
    'x has an isolated cognitive system which receives its inputs from and sends its outputs to an artificially-designed computer simulation of a world'
  - Then an envatted being could think 'I am envatted'.
    - The relevant expressions are semantically neutral.
-

# Skepticism Rehabilitated?

---

- Then: Putnam's reasoning cannot rule out the hypothesis that I am envatted.
  - I think: I cannot rule out the possibility that I am envatted.
  - So: skepticism all over again?
-

# Envatment Reconsidered

---

- My view: the envatment hypothesis is *not* a skeptical hypothesis, under which most of my beliefs are false.
    - Even if I am envatted, I still have hands, there is still a table in front of me, I am still in Tucson.
    - The epistemic possibility that I am envatted does not undercut my knowledge that I have hands, and so on.
  - Envatted beings (such as brains in vats) have mostly *true* beliefs, and are not largely deluded about the world.
    - (cf. Berkeley, Putnam)
-

# Envatment as a Metaphysical Hypothesis

---

- I'll argue: the hypothesis that I am envatted is not a **skeptical** hypothesis, but a **metaphysical** hypothesis.
  - It is epistemically equivalent to a hypothesis about the underlying metaphysical nature of the world.
  - If accepted, this hypothesis might revise a few of our beliefs, but it would leave most of them intact.
  - If so, envatment is a non-skeptical hypothesis.
-

# Three Metaphysical Hypotheses

---

- I' ll present three metaphysical hypotheses about the nature of our world, none of which is a skeptical hypothesis.
    - 1. A creation hypothesis.
    - 2. A pre-physical hypothesis
    - 3. A dualistic hypothesis
  - (We' ll then be considering their conjunction.)
-

# Creation Hypothesis

---

- Creation Hypothesis: Physical space-time and its contents were created by agents outside physical space-time.
    - And perhaps: are under the control of agents outside physical space-time.
  - This is an epistemically possible hypothesis: I cannot conclusively rule it out.
  - It is not a skeptical hypothesis: even if the Creation Hypothesis is correct, I still have hands, there is still a table in front of me, I am still in Tucson.
-



# Pre-Physical Hypothesis I

---

- (1) Physics is not ontologically fundamental: spacetime, mass, charge supervene on a more fundamental realm.
  - As chemistry is constituted by physics, physics is constituted by something else.
- (2) Physics is constituted by underlying *computational* processes.
  - More fundamental than elementary particles are patterns of bits (perhaps in a cellular automaton, or an algorithm).
- (3) These computational processes are themselves constituted/implemented by more basic processes.

# Pre-Physical Hypothesis II

---

- The Pre-Physical Hypothesis is epistemically possible: I cannot rule it out.
  - The Pre-Physical Hypothesis is not a skeptical hypothesis:
    - If it is true, there are still electrons and protons (they are just implemented, as are molecules).
    - If it is true, I still have hands, and so on.
  - To accept it would involve revision/addition to some of our beliefs, but not massive revision.
    - No more than quantum mechanics!
-

# Dualistic Hypothesis

---

- Dualistic Hypothesis: My mind is constituted by processes outside physical space-time, and receives its perceptual inputs from processes in physical space-time.
    - Cf. Descartes' substance dualism.
  - This hypothesis is epistemically possible: I cannot conclusively rule it out.
  - It is not a skeptical hypothesis: if it is true, I still have hands, and so on.
-

# Conjunctive Hypothesis

---

- Conjunctive Hypothesis: A conjunction of the creation hypothesis, the pre-physical hypothesis, and the dualistic hypothesis.
  - The Conjunctive Hypothesis is epistemically possible: I cannot conclusively rule it out.
  - The Conjunctive Hypothesis is not a skeptical hypothesis.
    - if it is true, I still have hands, and so on.
-

# Grand Metaphysical Hypothesis

---

- Grand Metaphysical Hypothesis:
    - (1) Conjunctive Hypothesis
    - (2) The processes implementing the pre-physical computational processes are part of a broader domain.
    - (3) My mind is supported by other processes in this domain.
    - (4) Physical space-time was created by agents in this domain, by arranging an implementation of the computational process to simulate a world.
-

# Grand Metaphysical Hypothesis II

---

- The Grand Metaphysical Hypothesis is epistemically possible: I cannot conclusively rule it out.
  - The Grand Metaphysical Hypothesis is not a skeptical hypothesis.
    - if it is true, I still have hands, and so on.
-

# Envatment as Metaphysics

---

- Envatment Hypothesis:
    - ‘I have an isolated cognitive system which receives its inputs from and sends its outputs to an artificially-designed computer simulation of a world’
  - Claim: The Envatment Hypothesis is (more or less) epistemically equivalent to the Grand Metaphysical Hypothesis.
    - If I accept the Grand Metaphysical Hypothesis, I should accept the Envatment Hypothesis.
    - If I accept the Envatment Hypothesis, I should accept the Grand Metaphysical Hypothesis.
-

# GMH → Envatment

---

- Dualistic Hypothesis → I have an isolated cognitive system which receives its inputs from and sends its outputs to physical space-time.
  - Pre-Physical Hypothesis → My cognitive system receives inputs from the computational processes (in the same domain) implementing physical space-time.
  - Creation Hypothesis → These computational processes were designed by agents (in the same domain) to simulate a world.
  - These jointly entail: Envatment Hypothesis!
-



# Envatment → GMH.

---

- Envatment Hypothesis → What underlies apparent reality is just as GMH specifies:
    - A domain containing my cognitive system, causally interacting with a computer simulation of physical space-time, created by agents in that domain.
    - This is all that is required to realize GMH.
  - So the Envatment Hypothesis is epistemically equivalent to GMH.
- .
-

# Envatment: Not Skeptical

---

- (1) Envatment is epistemically equivalent to GMH.
- (2) GMH is not a skeptical hypothesis

So

- (3) Envatment is not a skeptical hypothesis
-

# Envatment: Not so Bad

---

- Even if I am envatted
    - I have hands
    - There is a table in front of me
    - I am in Tucson
  - If I discovered that GMH was true, I would not reject these beliefs.
  - So if I discover that I am envatted, I should not reject these beliefs.
-

# Life in the Matrix

---

- So: people in the Matrix have mostly true beliefs
    - [At least if the Matrix has been running for a long time.]
    - [True beliefs about other minds may require a multi-vat matrix.]
  - If we discover that we are envatted, we should take this as a metaphysical discovery about the creation and underlying metaphysics of our world.
  - Even if we are envatted, our beliefs are mostly true.
-

# Worry 1: Deluded BIV

---

- Worry: A BIV might think ‘I am surrounded by people, basking in the sun’ , when it is alone in a dark lab: deluded!
  - Reply: The surroundings of the BIV in the brain’s domains are mostly irrelevant to the truth of the BIV’s beliefs.
  - Compare: A Cartesian ectoplasmic mind might think ‘I am surrounded by people, basking in the sun’ , when the mind is solitary ectoplasm in its own realm.
    - Surroundings in the ectoplasmic domain are mostly irrelevant to the truth of a Cartesian mind’s beliefs.
  - What matters is the *causal* basis of the cognitive system’s experiences and beliefs, not the system’s surroundings.
-

# Worry 2: BIV has hands?

---

- Worry: A BIV might think ‘I have hands’ .
    - This thought will be true if the BIV has hands.
    - But BIVs do not have hands!
    - So the BIV’s thought is false.
  - Reply: ‘Hands’ means something different for a BIV.
    - So a BIV’s ‘I have hands’ thought is not made true by the presence or absence of hands.
  - Deeper point: Referential externalism
    - Cf. Twin Oscar’s ‘This is water’ thought
    - Not made true by the presence or absence of water.
    - ‘water’, ‘hand’ are semantically non-neutral.
-

# Worry 3: BIV performs actions?

---

- Worry: Can raise worry 2 in a semantically neutral way.
    - E.g. semantically neutral terms such as ‘friend’, ‘philosopher’, ‘action’.
  - A BIV might think ‘I have friends’, ‘I perform actions’.
    - These thoughts will be true if the BIV has friends, performs actions.
    - But BIVs do not have friends, or perform actions!
  - Reply: *Envatted agents* have friends (assuming a multi-vatted simulation). And *envatted agents* perform actions.
    - N.B. Strictly speaking, brains don’t perform actions, people do.
    - (Cf. brains are a few inches across, people are not.)
    - Even if the BIV is not acting, the agent is acting (on its environment).
  - Similarly for other semantically neutral expressions.
-

# Worry 4: Spatial beliefs?

---

- Spatial concepts are arguably semantically neutral.
  - If so, BIV has false spatial thoughts: e.g. *This object is circular.*
    - [Assuming that there are no circular objects in the computer.]
  - Reply: spatial concepts are not semantically neutral.
    - BIV-world is a Twin Earth for space!
    - Can also argue this using El Greco worlds, classical worlds, etc.
  - Arguably, spatial concepts pick out something like: the manifold of properties and relations that serve as a causal basis for a subject's experience as of spatial properties and relations.
    - May vary between subjects.
    - Cf. color concepts and color properties.
-



# Worry 5: What is BIV referring to?

---

- Say a BIV thinks ‘There is a table in this room’ .
    - Truth of thought depends on objects referred to.
  - Question: What does its ‘table’ , ‘room’ , refer to?
    - Worry: not enough objects in the computer.
  - If we are envatted, our terms refer to entities that supervene (are constituted by) computational processes. Same for BIV.
    - Cf. quantum mechanics: Are there enough objects in the wavefunction?
  - Deeper point: Objects need not be token identical to objects in the underlying domain.
-

# Worry 5a: Those aren't objects!

---

- Worry: 'This table' for a BIV refers at best to a highly transient computational pattern (unstable over time), not an object.
  - Reply 1: Sure, it's an object (liberalism about objects).
  - Reply 2: Nor are our tables and chairs – look at quantum mechanics! (restrictivism about objects).
  - Reply 3: It's an object by the BIV's standards. (Contextualism about objects, and about existential quantification.)
-

# Worry 6: BIV has world all wrong

---

- Worry: If we are envatted, the world is nothing like we think it is.
  - Reply: I deny this. Even if we are envatted, there are still people, tables, football games, particles, arranged in space-time just as we think they are.
  - Deeper point: It's just that the world has a *further* nature that goes beyond our common sense conception.
  - But this is familiar already:
    - Cf. quantum mechanics.
    - Cf. science reveals world's structure, leaving intrinsic nature open.
      - Envatment: The ding-an-sich is a part of a computer-an-sich?
-

# Tentative Moral

---

- The “manifest image” is *robust*: our everyday thoughts about the world are not easily falsified by science and metaphysics.
    - Many of our concepts and thoughts are less demanding of the external world than we might think.
  - This robustness is tied to semantic non-neutrality.
    - Reference of these concepts is fixed to the causal/explanatory basis of the manifest image, with relatively few demands on its nature.
  - What are the residual demands of our manifest conception?
    - Arguably, these are those imposed by semantically neutral aspects of our conception. And arguably, the basic semantically neutral concepts are mental concepts and causal concepts.
    - If so, then the residual demands are mental and causal.
-

# Residual skepticism

---

- Paradigmatic envatment is a non-skeptical hypothesis.
  - But this does not defeat all forms of skepticism.
  - Other skeptical hypotheses remain...
-

# Local Envatment

---

- **Local Envatment:** Only my local environment is simulated, not all of physical space-time.
  - Local Envatment is equivalent to a Grand Metaphysical Hypothesis about my local environment, with false beliefs further out.
  - Akin to *Truman Show* skepticism: I still have hands, a body, a house, but I am deeply wrong about things farther from home.
  - Moral: Being in the Truman Show is much worse than being in the Matrix!
-

# Other Skeptical Hypotheses

---

- **Recent Envatment:** My mind was *recently* isolated and connected to a computer simulation.
    - Yields false beliefs about my present environment, although not necessarily about the past.
  - **Chaotic Envatment:** My mind is isolated and receives random stimulation from all sorts of sources that coincidentally yield apparent regularity.
    - Arguably: yields largely false or empty beliefs.
-

# Whither Skepticism?

---

- The residual skeptical hypotheses all lack *uniform explanation* of the regularities in experience.
    - Either no explanation (chaotic envatment), or non-uniform explanation (recent and local envatment).
  - Tentative hypothesis (to be qualified): Given the supposition of a uniform causal explanation of the regularities in my experience (and the supposition that my experiences are as regular as I think they are), global external-world skepticism is ruled out.
  - A limited anti-skeptical argument: supposing uniform explanation yields reality?
-



# Conclusion

---

- It's not so bad to be a brain in a vat.
-

Three cheers  
for philosophy!

