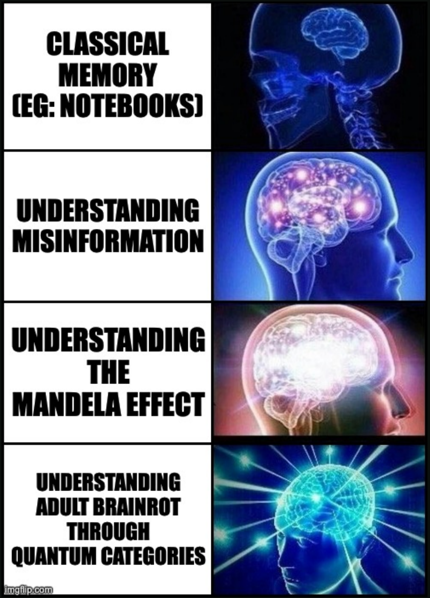
Adult Brainrot

Mandela Effect, Misinformation & Conspiracies in Quantum Categories

Khyathi Komalan (Caltech)

Topos Oxford Seminar, 2025


CLASSICAL MEMORY (EG: NOTEBOOKS)
UNDERSTANDING MISINFORMATION
UNDERSTANDING THE MANDELA EFFECT
UNDERSTANDING ADULT BRAINROT THROUGH QUANTUM CATEGORIES
imgflip.com

The Problem

- Classical memory (computers, notebooks): order of updates doesn’t matter.
- Human memory: order *does* matter — recall is reconstructive.
- (Loftus’s leading questions; Schuman–Presser: survey order; Ebbinghaus: serial-position effect)

The Puzzle

How do we model beliefs when updates don’t commute, sources blur, and global stories fail to cohere?

Roadmap

- Mandela effects $\rightarrow$ noncommuting channels/updates in $\mathbf{CPM}(\mathbf{FdHilb})$
- Misinformation $\rightarrow$ mixed states and their indistinguishable purifications
- Conspiracies $\rightarrow$ local coherence vs. failed global glue (centers)

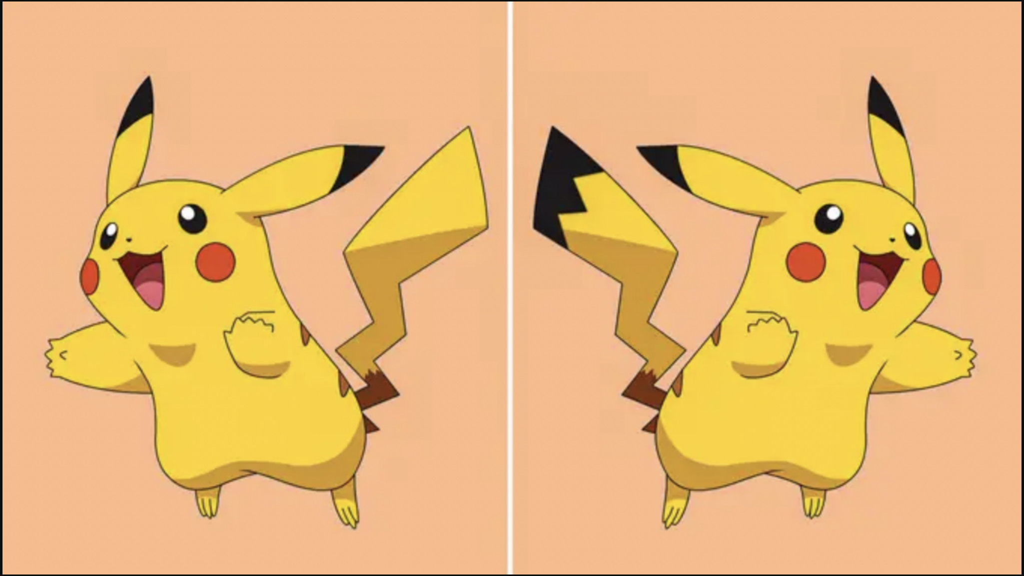
Using categorical tools in quantum computation, we try to model & understand the inconsistencies in our belief systems.

Mandela Effect


The Berenstein Bears
The Berenstain Bears

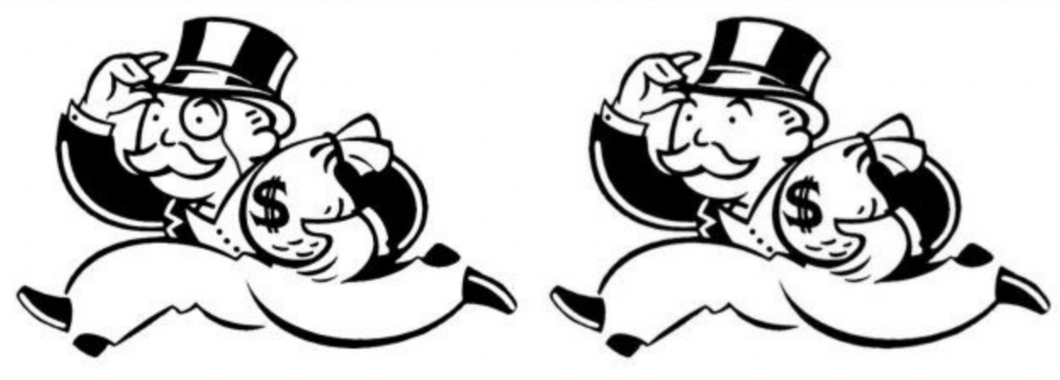
Answer: The Berenstain Bears

order-confusion: semantic; people expect to see “-stein”



Answer: No Black Tail

schema reconstruction: our brain “autocompletes” the pattern we see



Answer: No Monocle

source confusion: lots of imagery of wealthy people with monocles

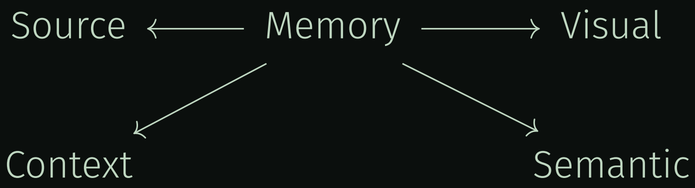
Mandela Effects & Memory

- **Order** — Berenstain $\leftrightarrow$ Berenstein
- **Schema** — Pikachu's "black tail tip"
- **Source** — Monopoly Man's "monocle"

- Memory: re-encoding, shaped by order, schemas, sources
- Psych: Loftus (misinfo), Johnson (source confusion)
- Math: CPM $\Rightarrow$ non-commutativity, channels forget/blur

How Memory Registers

Example:



We will be modelling how memory registers using **FdHilb**, a dagger compact category. In the next few slides, I will answer your (valid) what's, why's and how's.

What is **FdHilb**?

- Objects: finite-dimensional Hilbert spaces ($\mathbb{C}^n, \mathbb{C}^m \otimes \mathbb{C}^k, \ldots$).
- Morphisms: linear maps between them.
- Why Hilbert?
  - Inner product $\Rightarrow$ compare states (similar vs different memories).
  - Orthonormal basis $\Rightarrow$ possible “pure” recall traces.
  - Tensor product $\Rightarrow$ combine registers into one joint system.

Why **FdHilb**? - I

One of the main reasons we use **FdHilb** is because it is a dagger compact closed category.

- **Dagger (†):** every map $f : A \to B$ has an adjoint $f^\dagger : B \to A$ (involutive contravariant endofunctor)
  - Think “reversing” an update.
  - Lets us talk about symmetry in recall vs. encoding.

- **Compact closed:** we have “caps” (eval) and “cups” (coeval)
  - Graphically: can bend wires up or down.
  - Used to represent copying, pairing, or discarding registers.

Why **FdHilb**? - II

Gives a complete string diagram calculus.

- wires = registers, aka inputs/outputs
- boxes = updates, aka processes
- bending = forgetting / re-encoding, duals and dualizables like in the eval/coevals


ALWAYS HAS BEEN
WELL-BEHAVED MATH IS USELESS
imgflip.com

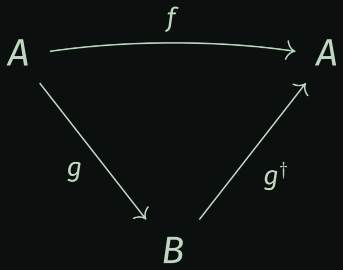
Why beyond pure **FdHilb**?

- Beliefs aren't pure: noise, ambiguity, blends.
- Updates aren't all unitary: they can blur, forget.
- We need: *mixed states* + *noisy channels*.
- Linear maps can't quite model mixed states

The CPM construction

- **CPM**(C): same objects as C; morphisms $A \to B$ are CP maps $f : A^* \otimes A \to B^* \otimes B$.
- For $C =$ **FdHilb**: morphisms = completely positive (trace-preserving) maps.
- Closed under composition & tensor, still dagger compact.

Positive map:

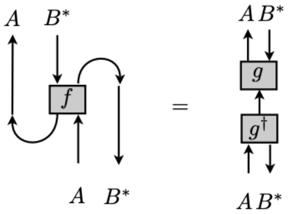


On Completely Positive Maps II

Completely positive map:

$$AB^* \xrightarrow{\eta_{A^*} AB^*} AA^*AB^* \xrightarrow{AfB^*} AB^*BB^* \xrightarrow{AB^*\varepsilon_B} AB^*$$

is positive, i.e.




Bernie
I am once again asking
why this complication is better
imgflip.com

Why CPM?

- Notice that the codomain and domain in CPM is of the form $B^* \otimes B$.
- And luckily for us, density matrices are of the form $\mathbb{C}^X \otimes (\mathbb{C}^X)^*$
- And density matrices model mixed states!

Beliefs & Updates

- Belief = state of a register.
- Update = influence applied to that register.
- Recall is often sequential: fact then meme $\neq$ meme then fact.


38
MEMETIC WARFARE
imgflip.com

Where do they live mathematically?

- **State (belief):** $\rho : I^* \otimes I \to A^* \otimes A$ (density operator).
- **Channel (meme vs fact):** A channel $\Phi : A \to B$ is a CPM of the form $A^* \otimes A \to B^* \otimes B$.
- **Sequential updates:** $\Phi_2 \circ \Phi_1$ (composition).

Forgetting context

- Recall often drops source/context.
- Behaviourally: we remember the image but not who told us.
- We need a categorical way to model the loss of information (memory registers).

Internal $\perp$-trace in CPM

- Discarding effect: $\perp_C\colon C^* \otimes C \to I$.
- Max-mixed state: $\perp_C^\dagger\colon I \to C^* \otimes C$.
- Given $f : (A \otimes C)^* \otimes (A \otimes C) \to (B \otimes C)^* \otimes (B \otimes C)$, define

$$\mathrm{Tr}^C(f) := (1_{B^* \otimes B} \otimes \perp_C) \circ f \circ (1_{A^* \otimes A} \otimes \perp_C^\dagger) : A^* \otimes A \to B^* \otimes B.$$

- Intuition: inject neutral C, then discard C.


LEFT
EXIT 12
remembering what's correct
remembering the version i find interesting
imgflip.com

Example: two influences on memory

- Φ_{meme}: catchy but false.
- Φ_{fact}: corrective cue.
- Both act on $A = V \otimes S$.

Two Channels

$$\Phi_{\text{meme}}(\rho) = \sum_i K_i \rho K_i^\dagger, \quad \Phi_{\text{fact}}(\rho) = \sum_j L_j \rho L_j^\dagger$$

- K_i = Kraus operators encoding a biased “meme update” (skews memory toward schema).
- L_j = Kraus operators encoding a corrective “fact update.”
- Both are CP + TP $\Rightarrow$ valid morphisms in $\mathrm{CPM}(\textbf{FdHilb})$.
- Typically noncommuting $\Rightarrow$ order of meme/fact exposure changes outcome.

Idea: Meme vs Fact as Channels

- The point of Kraus operators is to evaluate the effect of a quantum operation on a density matrix ρ, which represents states.
- The idea is to use Kraus operators to model the bias toward the more interesting story showcased through memes or elsewhere, and also toward the truth once the bias is called out/corrected.

Berenstain Example: Meme vs Fact

- Register space: $H = \mathbb{C}^2$ with basis $|0\rangle =$ "Berenstain", $|1\rangle =$ "Berenstein".

Meme channel (bias toward the false feature):

$$\Phi_{\text{meme}}(\rho) = \sum_{i=0}^{2} K_i \, \rho \, K_i^\dagger, \quad K_0 = \sqrt{1-p}\, I,\ K_1 = \sqrt{p}\, |1\rangle\langle 0|,\ K_2 = \sqrt{p}\, |1\rangle\langle 1|.$$

Intuition: softly shifts probability mass toward $|1\rangle$ (" -stein").

Fact channel (corrective pull back to the true feature):

$$\Phi_{\text{fact}}(\rho) = \sum_{j=0}^{2} L_j \, \rho \, L_j^\dagger, \quad L_0 = \sqrt{1-q}\, I,\ L_1 = \sqrt{q}\, |0\rangle\langle 1|,\ L_2 = \sqrt{q}\, |0\rangle\langle 0|.$$

Intuition: nudges weight back toward $|0\rangle$ (" -stain").

Toy Model Limitations

- This example is 2-dimensional, assumes only true/false
- Channels are hand-picked — in practice bias and correction are more complex.
- Here “meme” and “fact” act independently; in reality they can interact.
- Works as a toy model: shows **order effects**, but not the full richness of human memory.

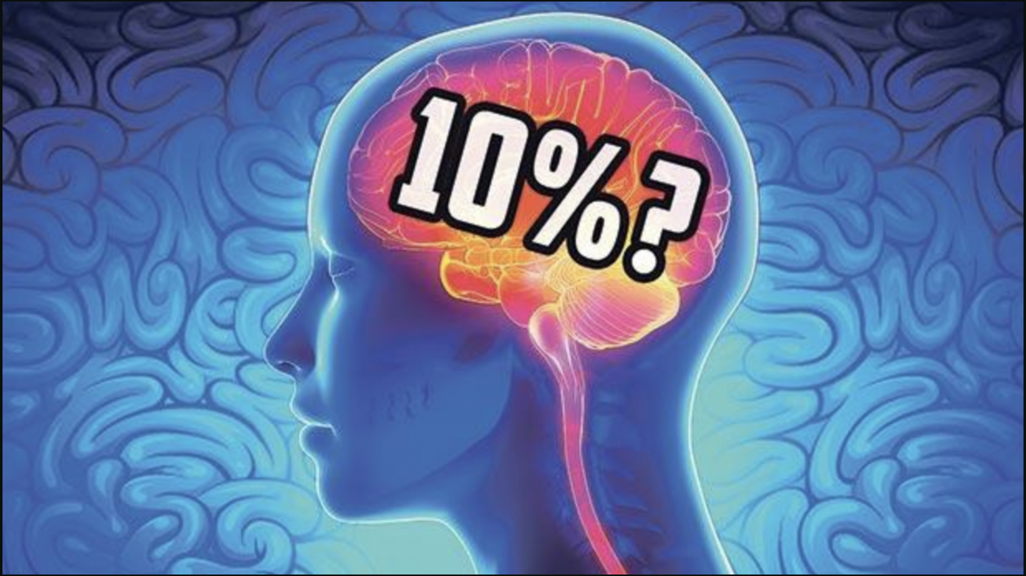
Schema Reconstruction?

- Recall = a **measurement** on register *V*.
- Brain favors the schema-consistent result ("black tail tip") even when it's false.
- Formal shadow: measurement is **biased**, giving higher weight to schema than to the veridical trace.
- Outcome reshapes the state of memory accordingly.


BELIEFS/MEMORY
MEME CHANNEL
FACT CHANNEL
imgflip.com

Misinformation


I'm not dumb like that
AquariumCircle.com


10%?

How this occurs

- We often only observe how a belief register A changes.
- The origin/context C (platform, persona, framing) is *unseen*.
- Need: include C during update, then **forget** C when reading A.

Hide & forget a register (CPM-internal trace)

Let $f : (A \otimes C)^* \otimes (A \otimes C) \to (B \otimes C)^* \otimes (B \otimes C)$ be CP.
In $\mathrm{CPM}(\textbf{FdHilb})$ we have:

$$\bot_C : I \to C^* \otimes C \quad \text{(max-mixed)} \qquad \bot_C^\dagger : C^* \otimes C \to I \quad \text{(discard)}.$$

Define the internal trace over C:

$$\boxed{\mathrm{Tr}^C(f) = (1_{B^* \otimes B} \otimes \bot_C^\dagger) \circ f \circ (1_{A^* \otimes A} \otimes \bot_C) \ : \ A^* \otimes A \to B^* \otimes B}$$

Purification of CP maps

- For any CP $f : A^* \otimes A \to B^* \otimes B$,

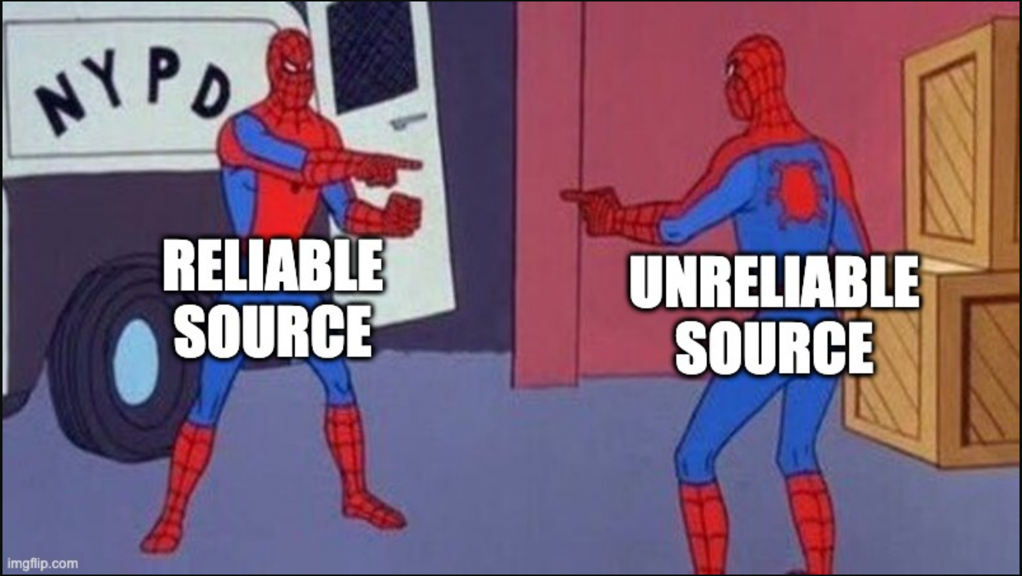
$$\exists\, C,\ g \textit{ pure} \text{ on } A \otimes C \quad \text{s.t.} \quad f = \mathrm{Tr}^C(g).$$

- Non-unique: different g, g' may yield the same f after tracing out C.

Indistinguishability

- If $g \neq g'$ but $\mathrm{Tr}^C(g) = \mathrm{Tr}^C(g')$, then origins differ but the *observed* update on A is identical.
- Interpretation: multiple hidden pipelines $\Rightarrow$ same surface "belief push".

$$A \xrightarrow{g} B \otimes C \xrightarrow{\perp_C^\dagger} B \qquad \equiv \qquad A \xrightarrow{g'} B \otimes C \xrightarrow{\perp_C^\dagger} B$$

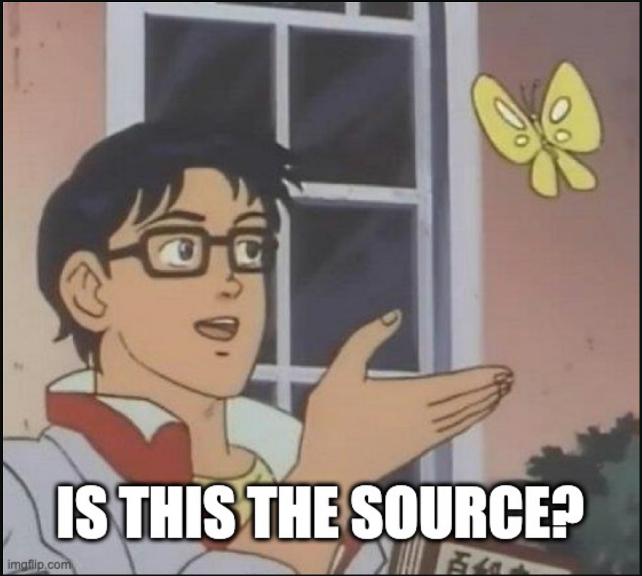

NYPD
RELIABLE SOURCE
UNRELIABLE SOURCE
imgflip.com

Aggregation & repetition

- Cross-source blend: $\bar{\Phi} = \sum_i w_i \Phi_i$ (convexity $\Rightarrow$ CP).
- Re-exposure: iterate $\bar{\Phi}^k$ (CPM closed under composition).
- Intuition: repetition drives beliefs toward a mixed fixed point of $\bar{\Phi}$.

Behavioral map

- Hidden register C = provenance (platform, persona, framing).
- Joint update on $A \otimes C$, then discard C = **source monitoring loss**.
- Non-unique purification = **indistinguishable origins** for the same observed push.


IS THIS THE SOURCE?
imgflip.com

Toy model: platform + reshare

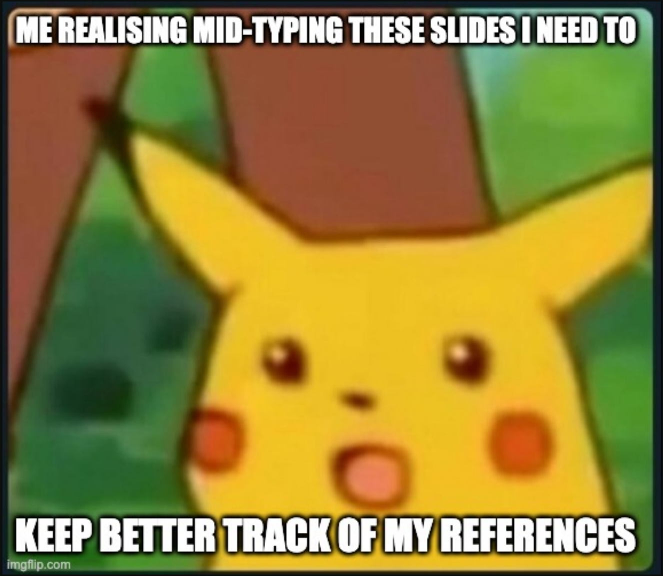
Let A be the belief register. Let C encode context (platform & reshare flag). Pick an isometry $V : A \to B \otimes C$. Define

$$g(\rho) = V \rho V^{\dagger} \quad \text{on } (A \otimes C)^{*} \otimes (A \otimes C), \qquad f = \mathrm{Tr}^{C}(g).$$

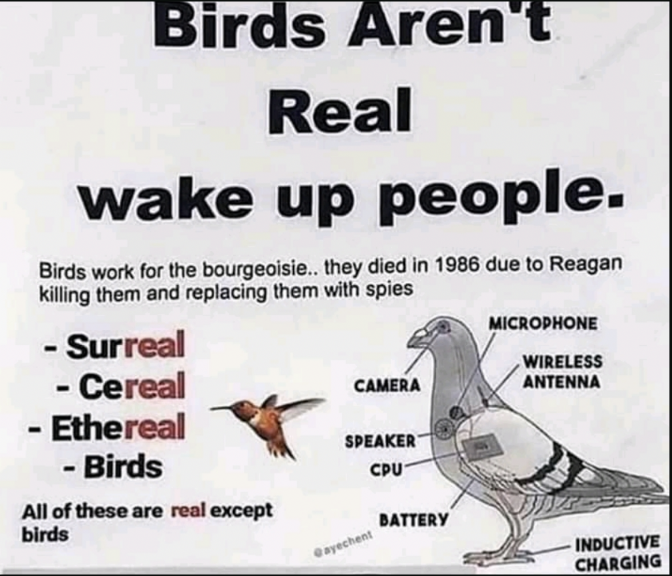
- g is pure CP; f is the observed CPM channel on beliefs.
- Different choices of V (different "pipelines") can yield the same f.

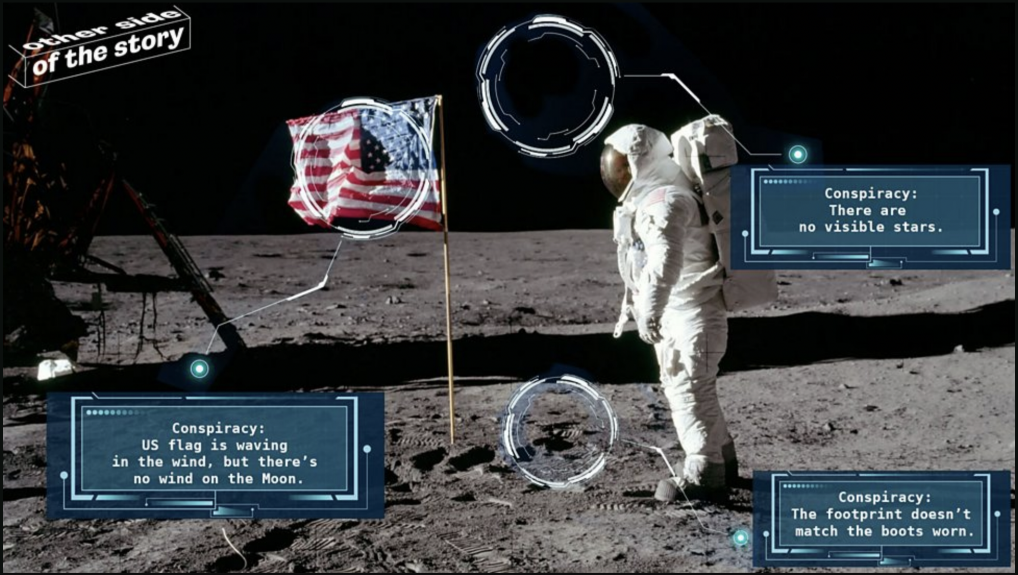
Limits & diagnostics

- Tracing out C *forgets* cause by design — can't recover provenance from f.
- Remedy in model: keep more registers (source tags R), compare via Choi states.
- Takeaway: CPM + internal trace is the minimal typed setting to model indistinguishable causes, blending, and repetition dynamics.


ME REALISING MID-TYPING THESE SLIDES I NEED TO
KEEP BETTER TRACK OF MY REFERENCES
imgflip.com

Conspiracy Theories


Birds Aren't
Real
wake up people.
Birds work for the bourgeoisie.. they died in 1986 due to Reagan killing them and replacing them with spies
- Surreal
- Cereal
- Ethereal
- Birds
All of these are real except birds
MICROPHONE
WIRELESS ANTENNA
CAMERA
SPEAKER
CPU
BATTERY
INDUCTIVE CHARGING
@ayechent


other side of the story
Conspiracy: There are no visible stars.
Conspiracy: US flag is waving in the wind, but there's no wind on the Moon.
Conspiracy: The footprint doesn't match the boots worn.

What do they have in common?

- Conspiracy theories are often stories that make sense when you consider only a subset of the available context.
- Most theories fall apart when you consider ALL of the context, since there will be contradictions with reality.
- In a sense, these theories are local.

From local fragments to a global theory

- **Setup:** We have a *category of fragments* B: objects = local info-types; morphisms = ways to connect them.
- **Tensor** $X \otimes Y$: place fragments side-by-side (joint context).
- **Braiding** $c_{X,Y} : X \otimes Y \to Y \otimes X$: *swap order of updates* (allowed re-orderings / compatibilities).
- Goal: when do locally coherent pieces assemble into a *globally* coherent story?

(In finite semisimple/rigid cases, this is a braided “fusion” setting, but we won’t need that language.)

What the braiding buys you

- $c_{X,Y}$ formalizes "**order changes** that should be harmless."
- Hexagon/coherence laws $\Rightarrow$ different swap sequences agree.
- Intuition: if updates commute *up to the braiding*, they coexist cleanly.
- We'll test global coherence by asking for objects that *braid well with everything*.

The “center” of fragments

- **Drinfeld center** $Z(B)$: pairs $(X, \gamma_{X,-})$ where $\gamma_{X,Y} : X \otimes Y \to Y \otimes X$ is natural in Y and satisfies the standard coherence (hexagons).
- Read: X is *transparent* to every other fragment Y.
- A “good global glue” should live in (or be detected by) $Z(B)$.


collect what seems to be true
piece them together to make sense of it
assemble a complete, accurate story
assemble a complete, accurate story
imgflip.com

Summary & Future Work

Summary

- **Mandela effects:** noncommuting channels in **CPM(FdHilb)** $\rightarrow$ order, schema, and source confusions.
- **Misinformation:** purification + internal trace $\rightarrow$ indistinguishable hidden causes, repetition.
- **Conspiracies:** local fragments in braided categories $\rightarrow$ global obstruction.

Future Work

- Explore higher-categorical lifts (fusion 2-categories).
- Empirical tie-ins: align CPM dynamics with psych experiments.
- Extend toy models toward realistic networks of belief updates.

References I

E. F. Loftus.
Leading questions and the eyewitness report.
Cognitive Psychology, 7(4):560–572, 1975.
doi:10.1016/0010-0285(75)90023-7.

A. Permaloff, H. Schuman, and S. Presser.
Questions and answers in attitude surveys: Experiments on question form, wording, and context.
The American Political Science Review, 77:1133, 1983.
doi:10.2307/1957697.

H. Ebbinghaus.
Memory: A Contribution to Experimental Psychology.
Translated by H. A. Ruger and C. E. Bussenius. Teachers College, Columbia University, New York, 1913.

References II

E. F. Loftus.
Planting misinformation in the human mind: A 30-year investigation of the malleability of memory.
Learning & Memory, 12(4):361–366, 2005.
doi:10.1101/lm.94705.

M. K. Johnson.
Source monitoring and memory distortion.
Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 352(1362):1733–1745, 1997.
doi:10.1098/rstb.1997.0156.

References III

P. Selinger.
Dagger compact closed categories and completely positive maps (extended abstract).
Electronic Notes in Theoretical Computer Science, 170:139–163, 2007.
Preliminary version in QPL 2005. doi:10.1016/j.entcs.2006.12.018.

T. Lagiotis, T. Kammermeier, J. van Dyke, J. Romero, M. Zetto, A. Marín Salvador, A. Hofstetter, L. Liu, N. Chetan Vuppulury, *et al.*
Higher tensor categories and their extensions: Notes from the Scottish Talbot on Algebra and Topology.
September 2024. Introduction by T. Johnson-Freyd and D. Reutter. Edited by J. Brown, B. Haïoun, D. Jordan, and P. Kinnear.
arXiv:2509.10636.

References IV

S. Abramsky and B. Coecke.
A categorical semantics of quantum protocols.
In *Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science (LICS)*, pages 415–425. IEEE Computer Society Press, 2004.
doi:10.1109/LICS.2004.1319636.