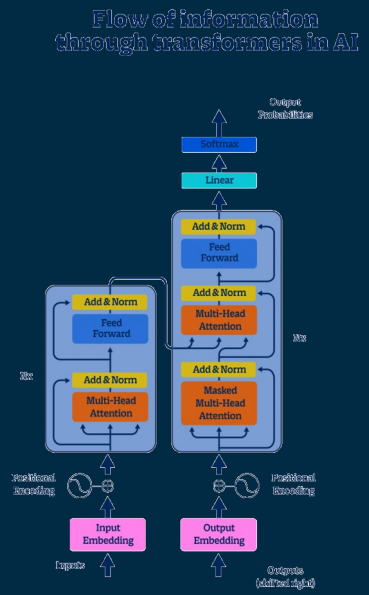
GOTO **COPENHAGEN**

Craft software with confidence in 2026

Grokking the Brain: What AI Reveals About How You Really Learn



Barbara Oakley, PhD, PE, FAAAS, FIEEE, FAIMBE
Coursera's Inaugural Innovation Instructor
Winner of McGraw Prize in Education (2023)
Distinguished Professor of Engineering Emerita
Oakland University, Rochester, Michigan

Ernst Strüngmann Forum

Memory, Learning, and Neural Adaptation in the Age of Cognitive Offloading

Barbara A. Oakley and Terrence J. Sejnowski, Chairs

August 30–September 4, 2026

Frankfurt am Main, Germany



Barbara Oakley, PhD, PE, FAAAS, FIEEE, FAIMBE
Coursera's Inaugural Innovation Instructor
Winner of McGraw Prize in Education (2023)
Distinguished Professor of Engineering Emerita
Oakland University, Rochester, Michigan



Terrence Sejnowski
Francis Crick Professor
The Salk Institute
Professor of Biological Sciences
University of California, San Diego











Philip Oakley

The transformer

Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com

Noam Shazeer* Google Brain noam@google.com

Niki Parmar* Google Research nikip@google.com

Jakob Uszkoreit* Google Research usz@google.com

Llion Jones* Google Research llion@google.com

Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu

Łukasz Kaiser* Google Brain lukaszkaiser@google.com

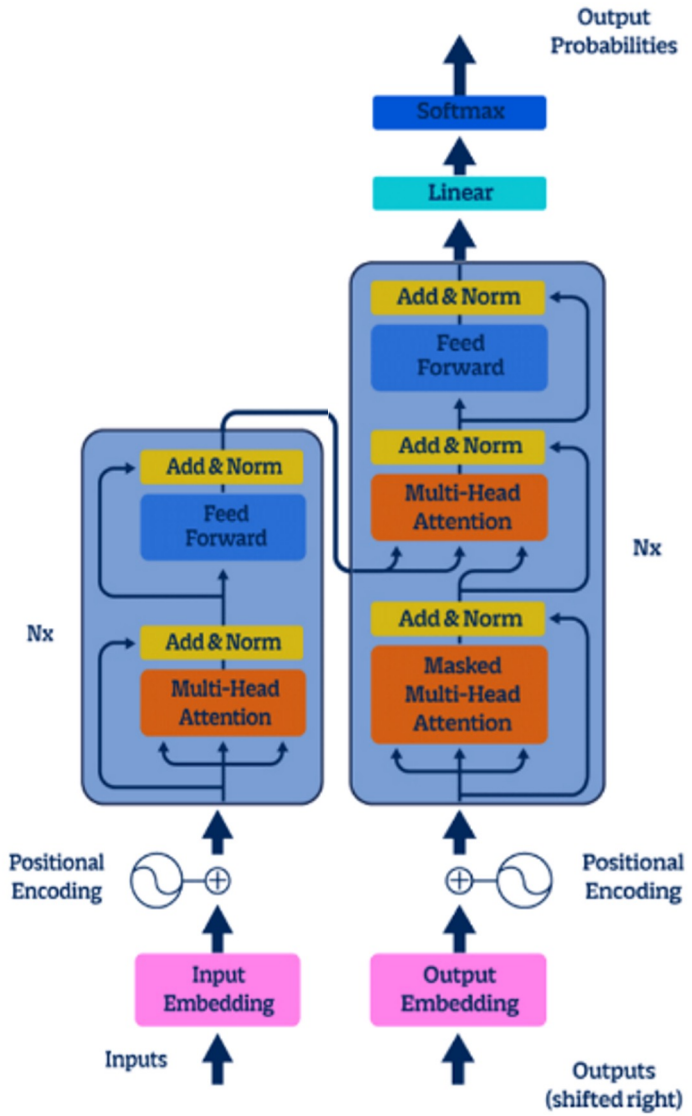
Illia Polosukhin* ‡ illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.0 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature.

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

Encoder

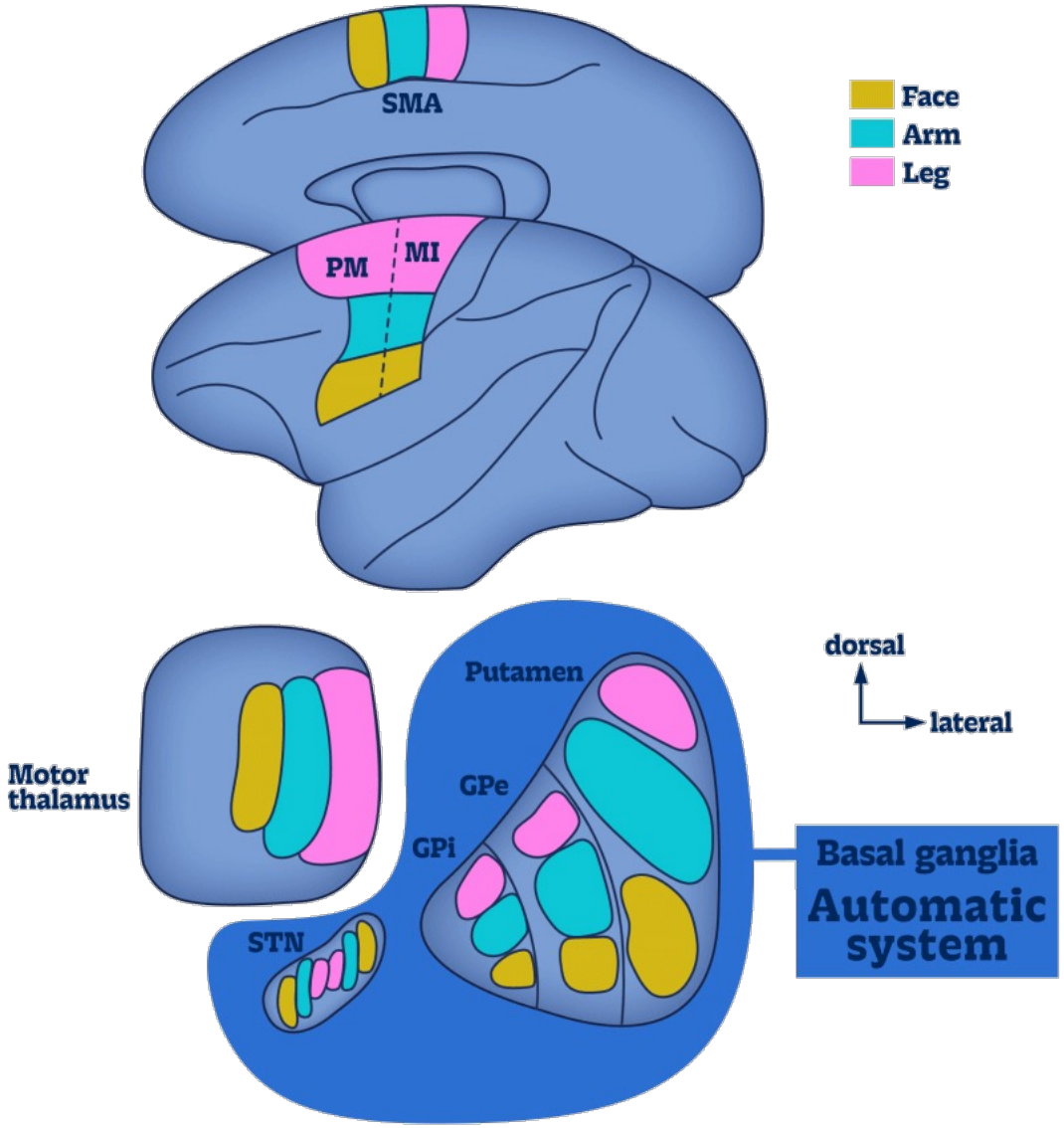


Decoder

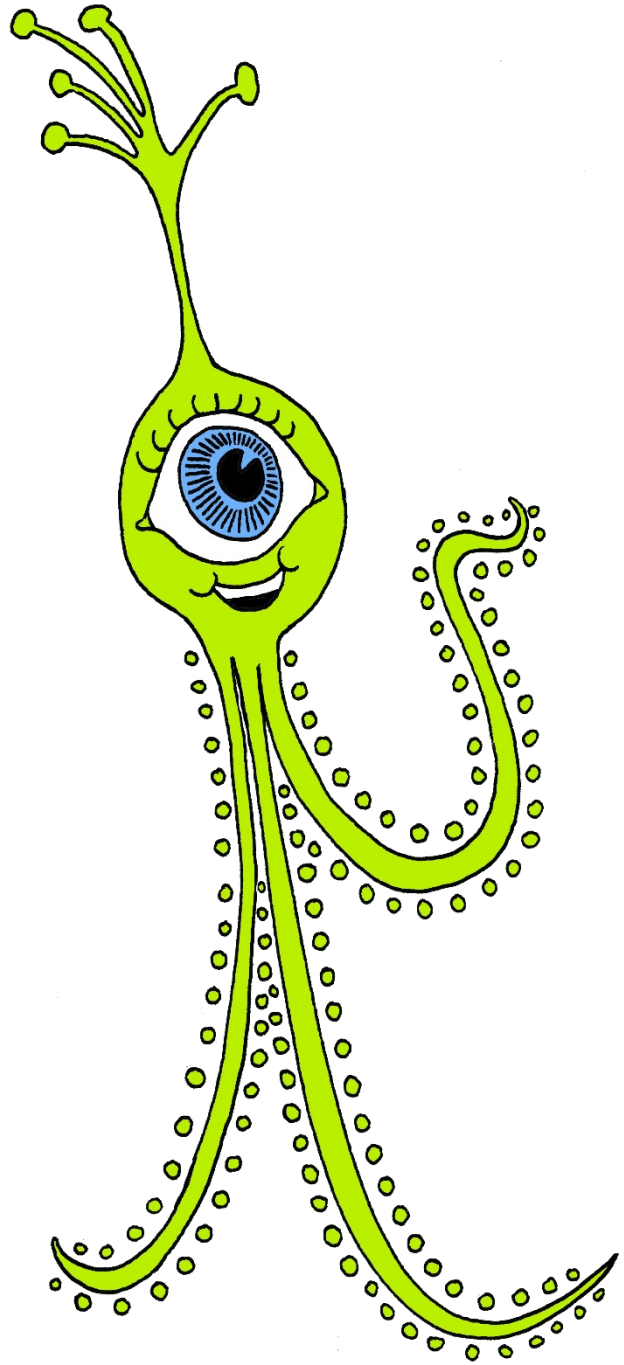
Flow of information through transformers in AI

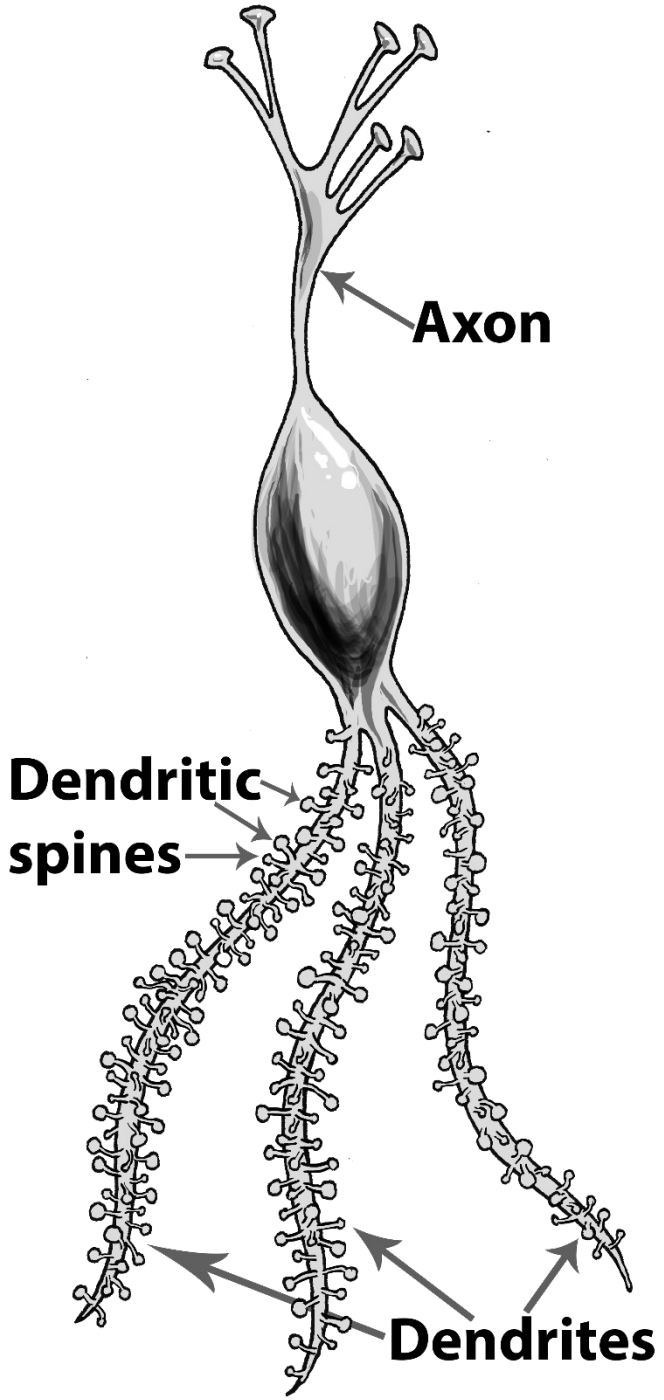


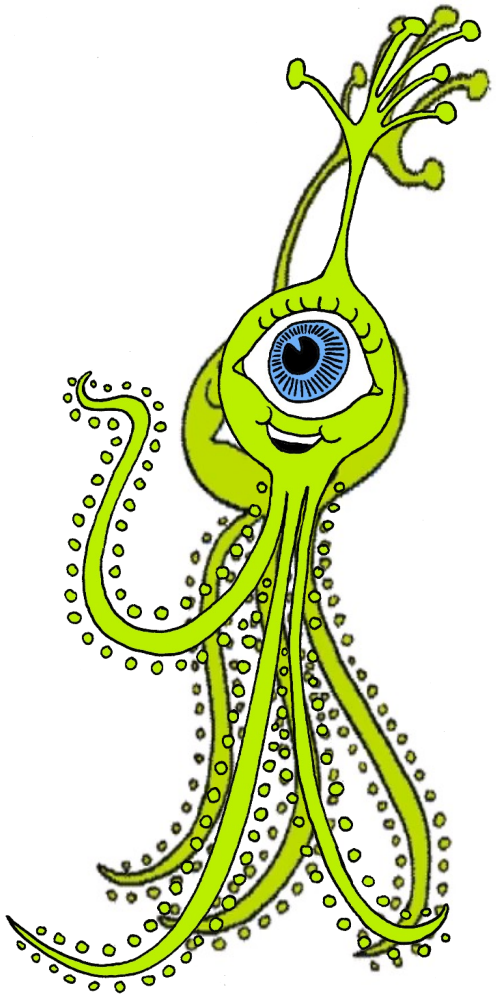
Flow of information in the brain

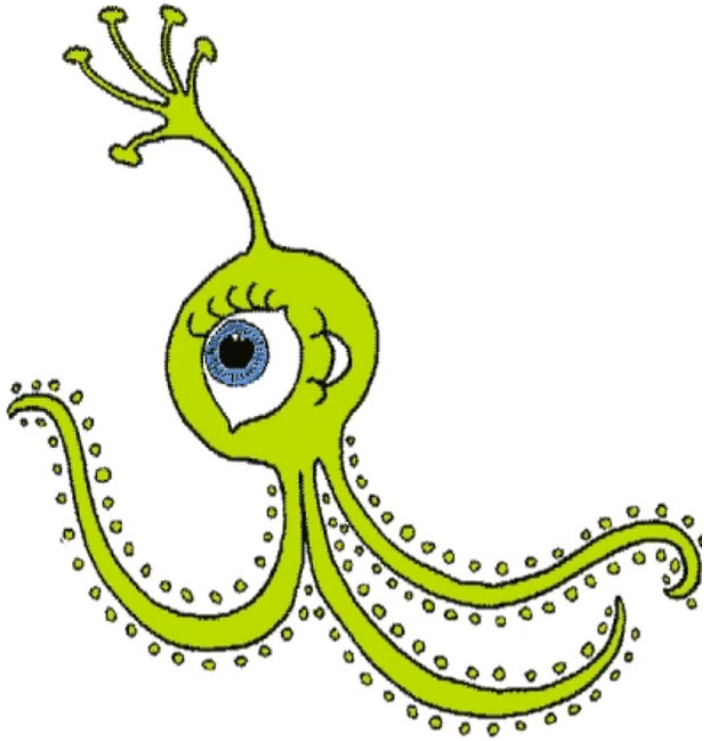


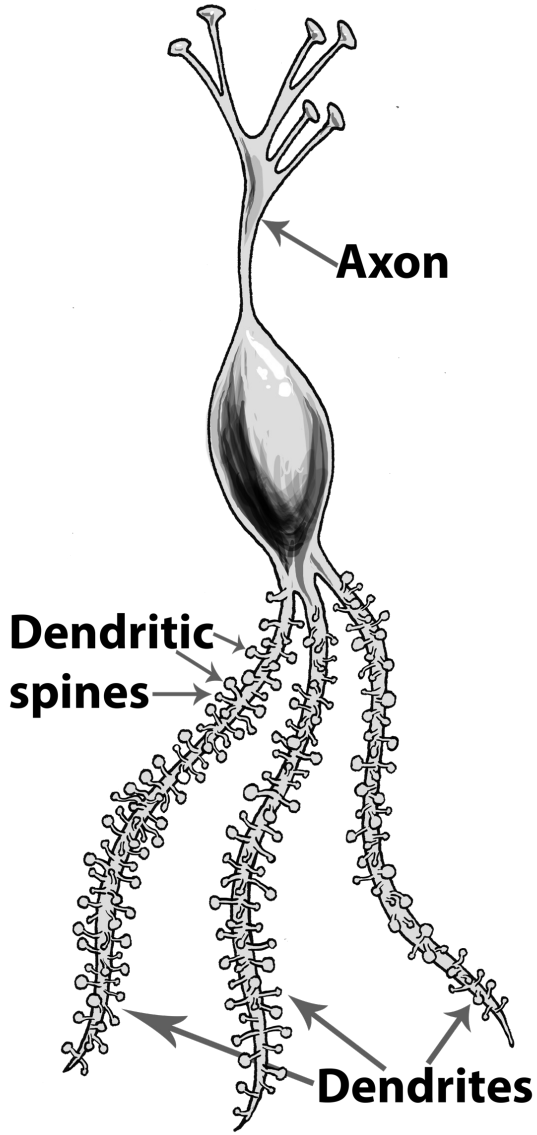

Axon
Dendritic
spines

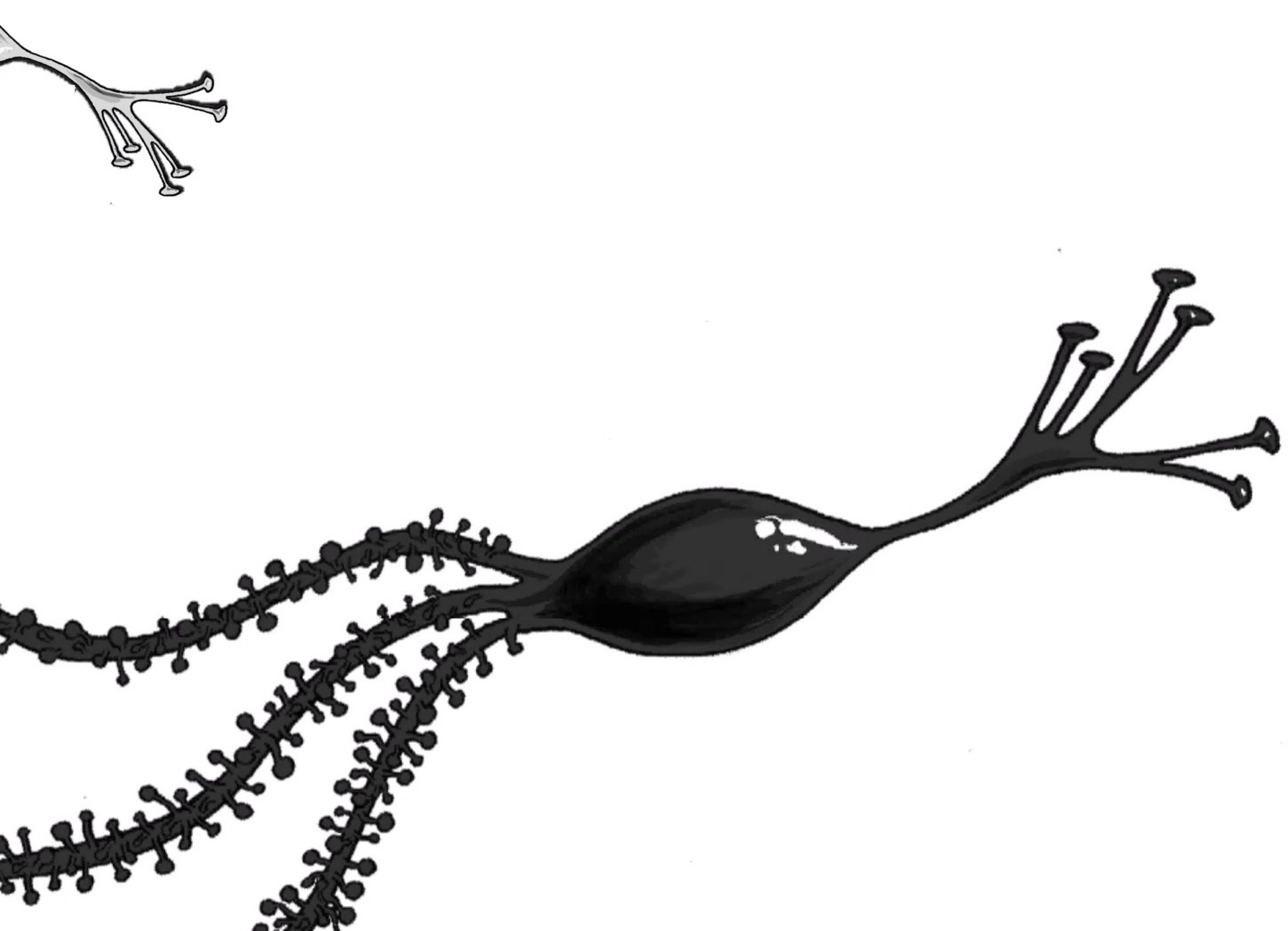



Axon
Dendritic spines
Dendrites

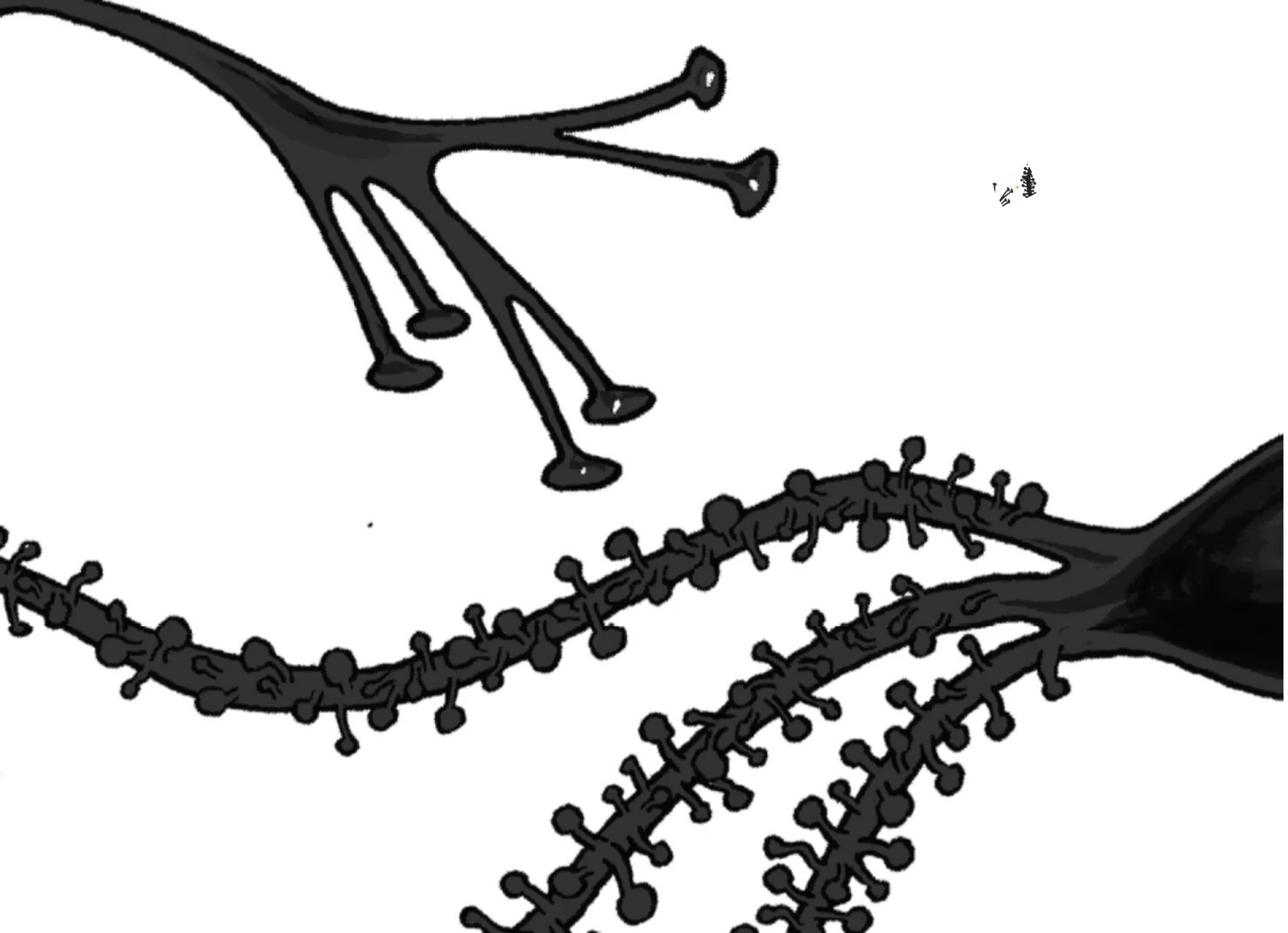





Axon
Dendritic spines
Dendrites

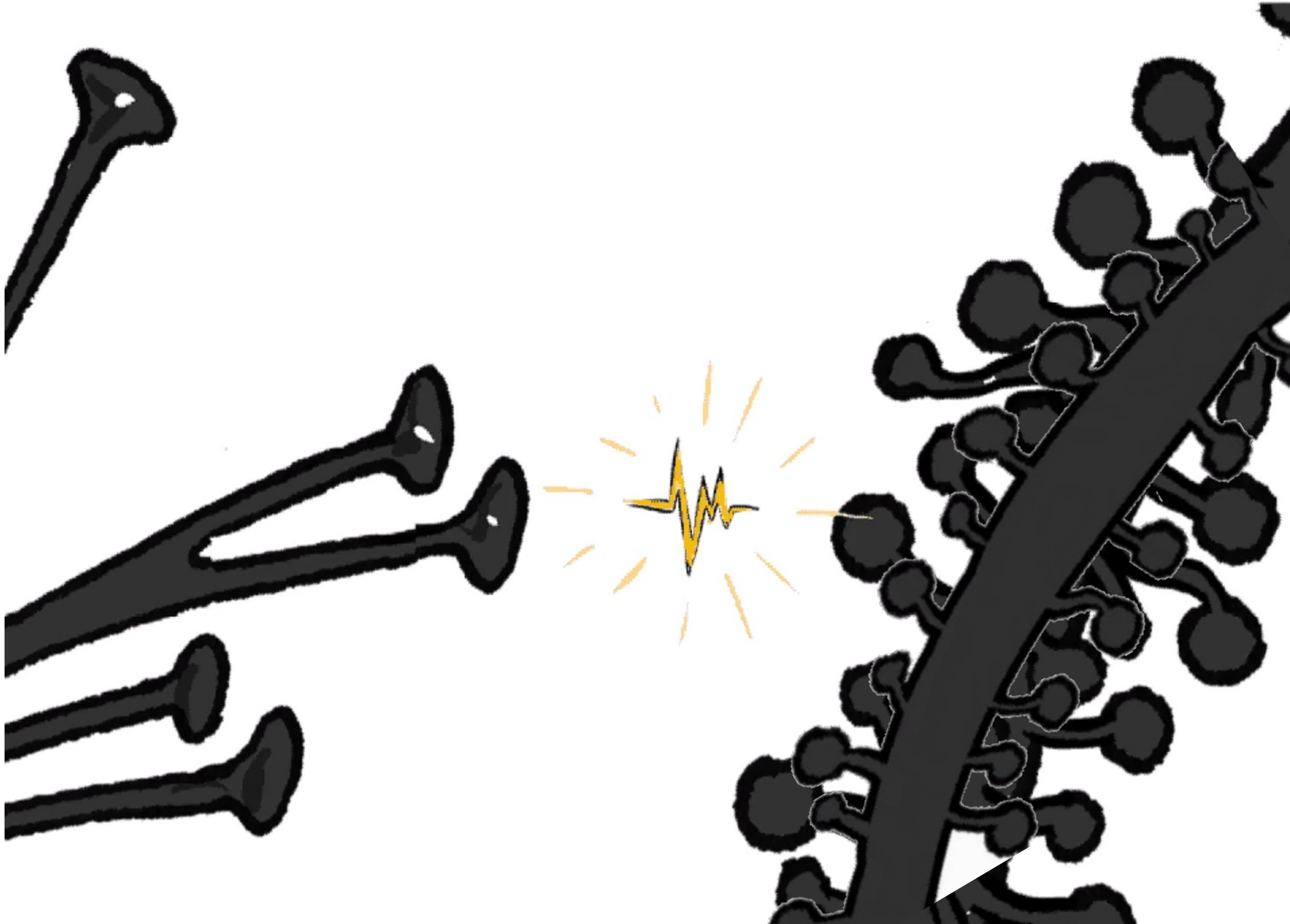


Neurons send signals

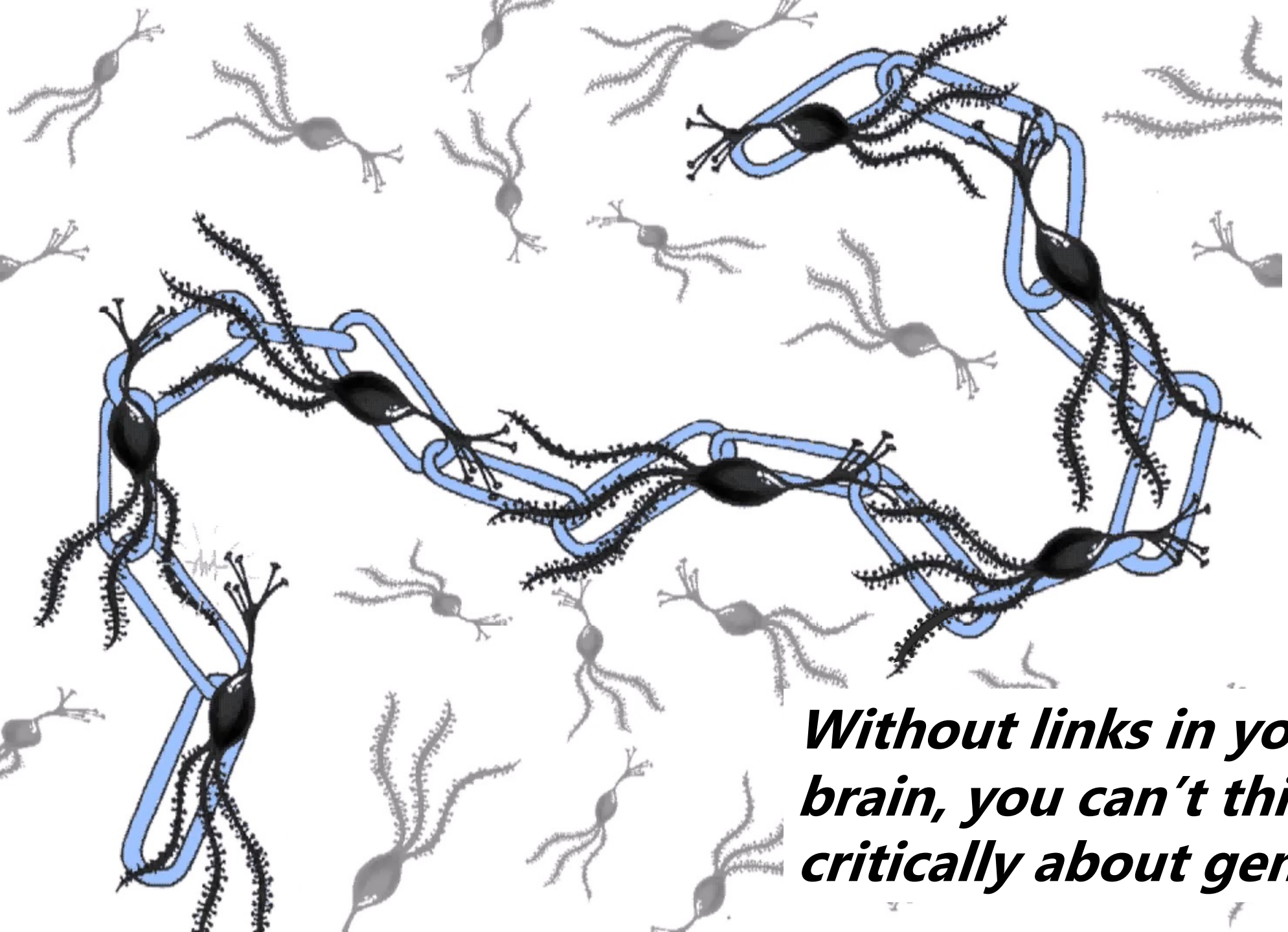


Neurons create sets of links when you learn something.

Connections strengthen with practice





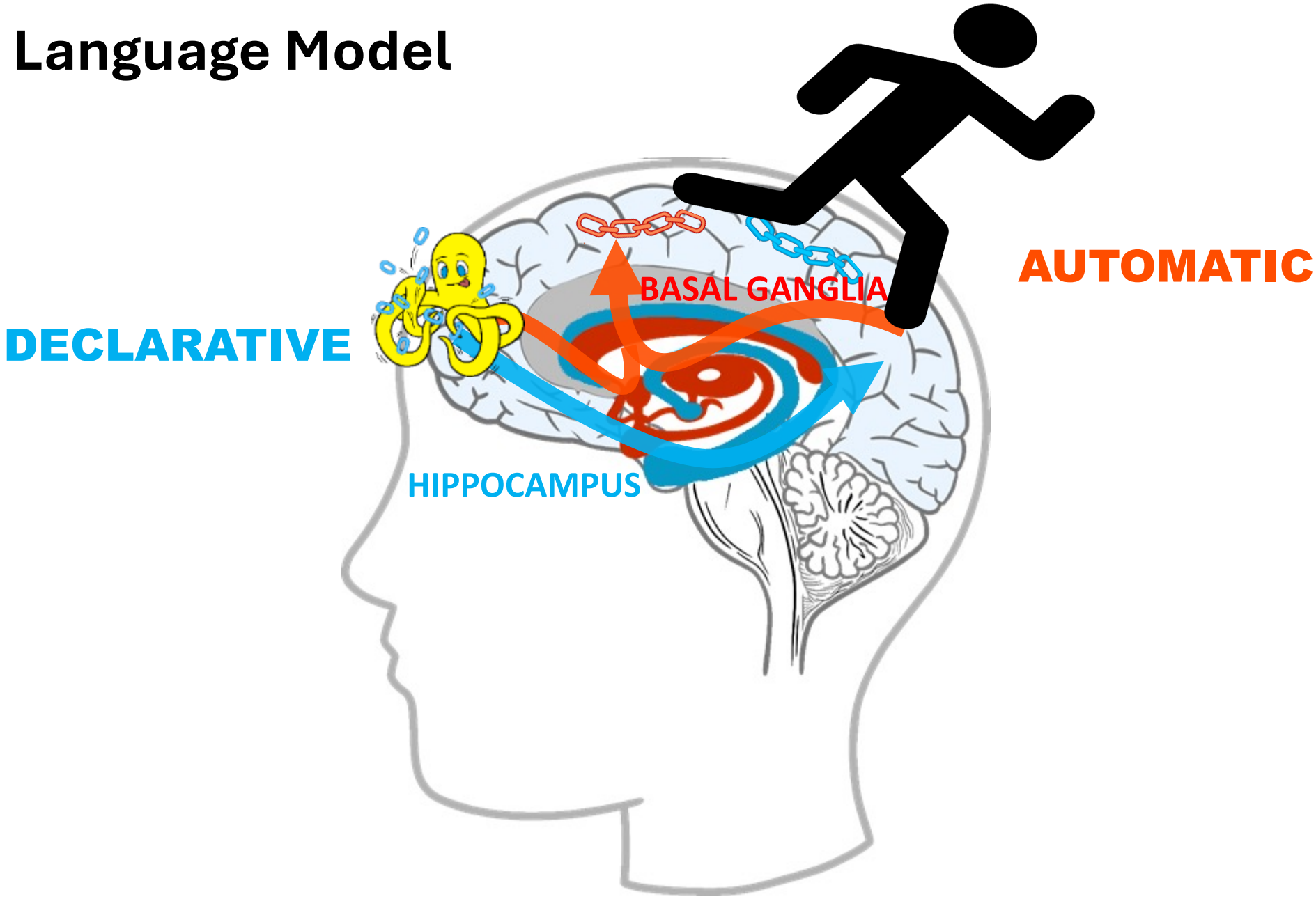


Without links in your own brain, you can´t think critically about genAI!

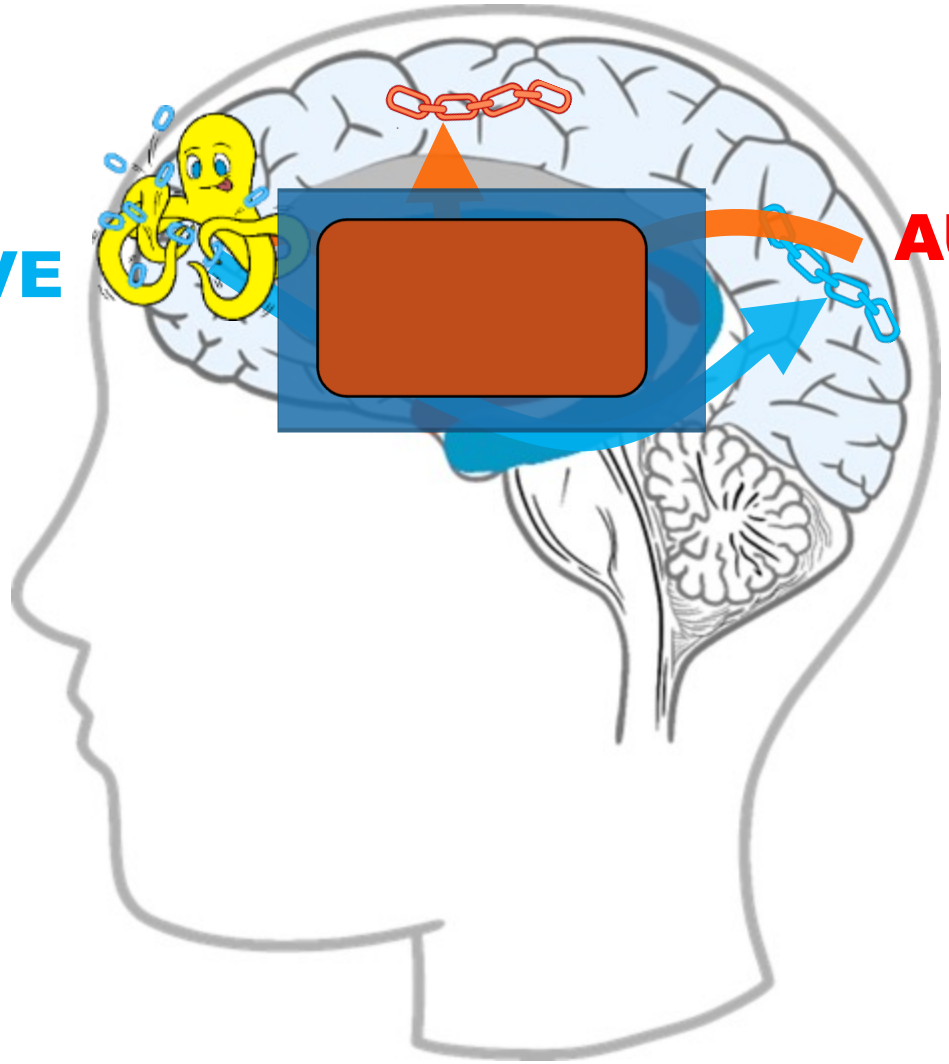
AI shows *But I'm the only one who KNOWS!*



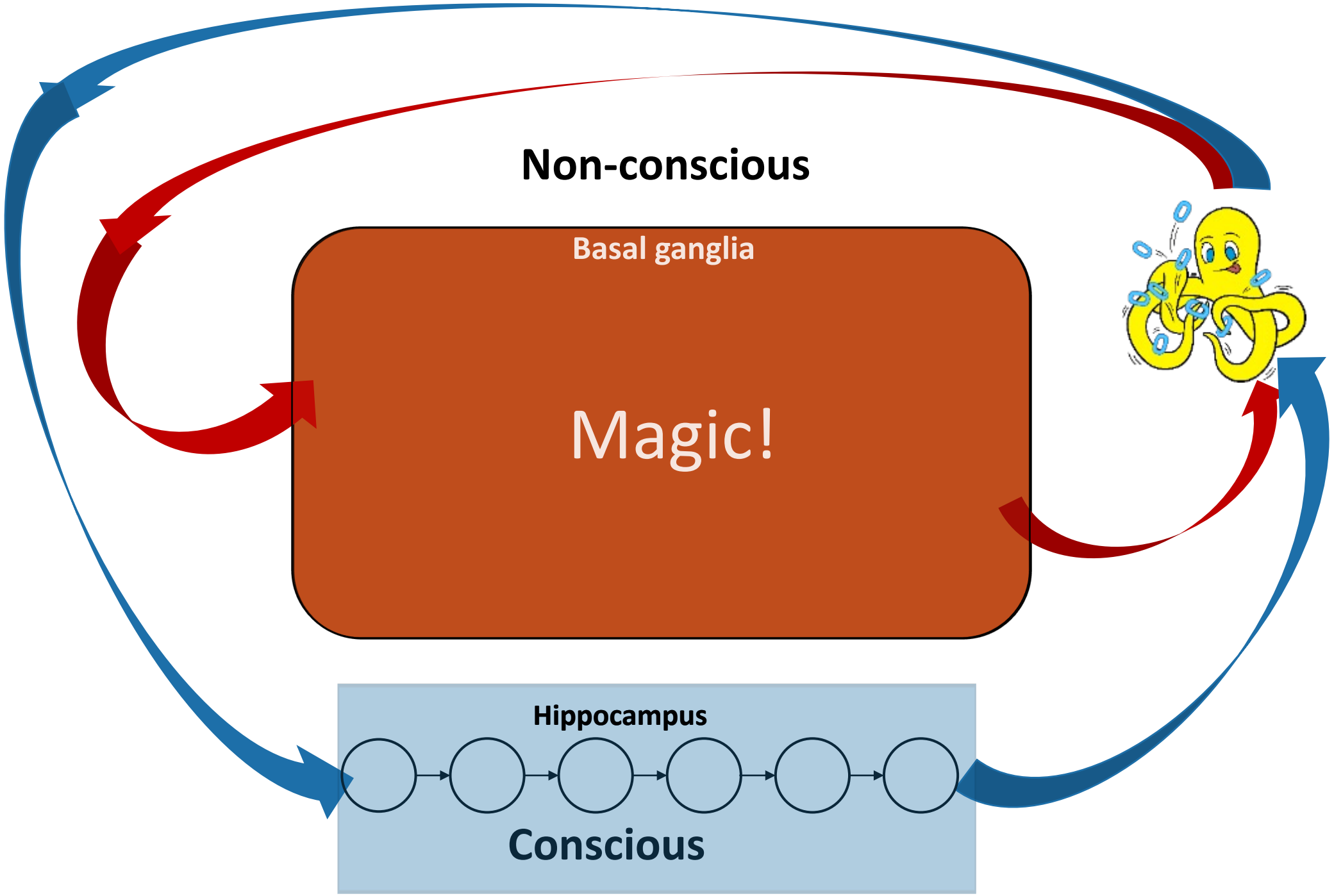
Large Language Model



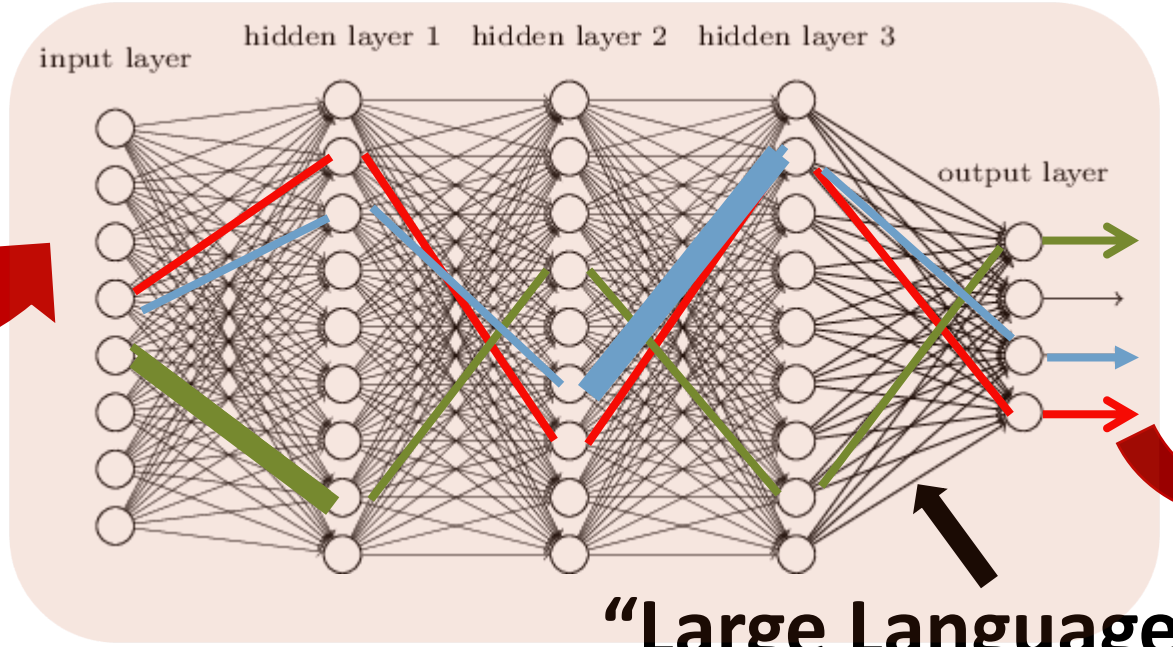
DECLARATIVE



AUTOMATIC


Non-conscious
Basal ganglia
Magic!
Hippocampus
Conscious

Non-conscious

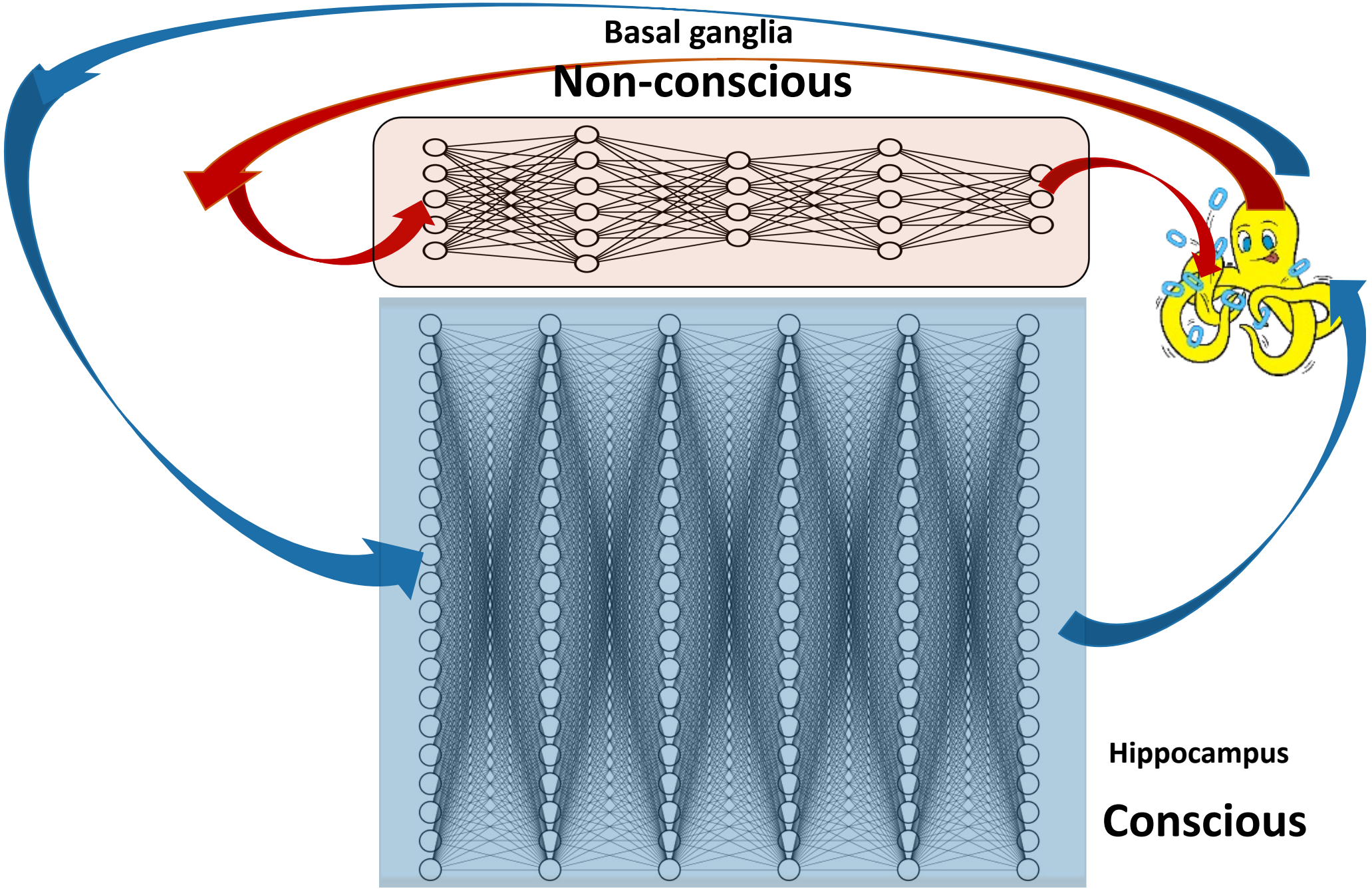


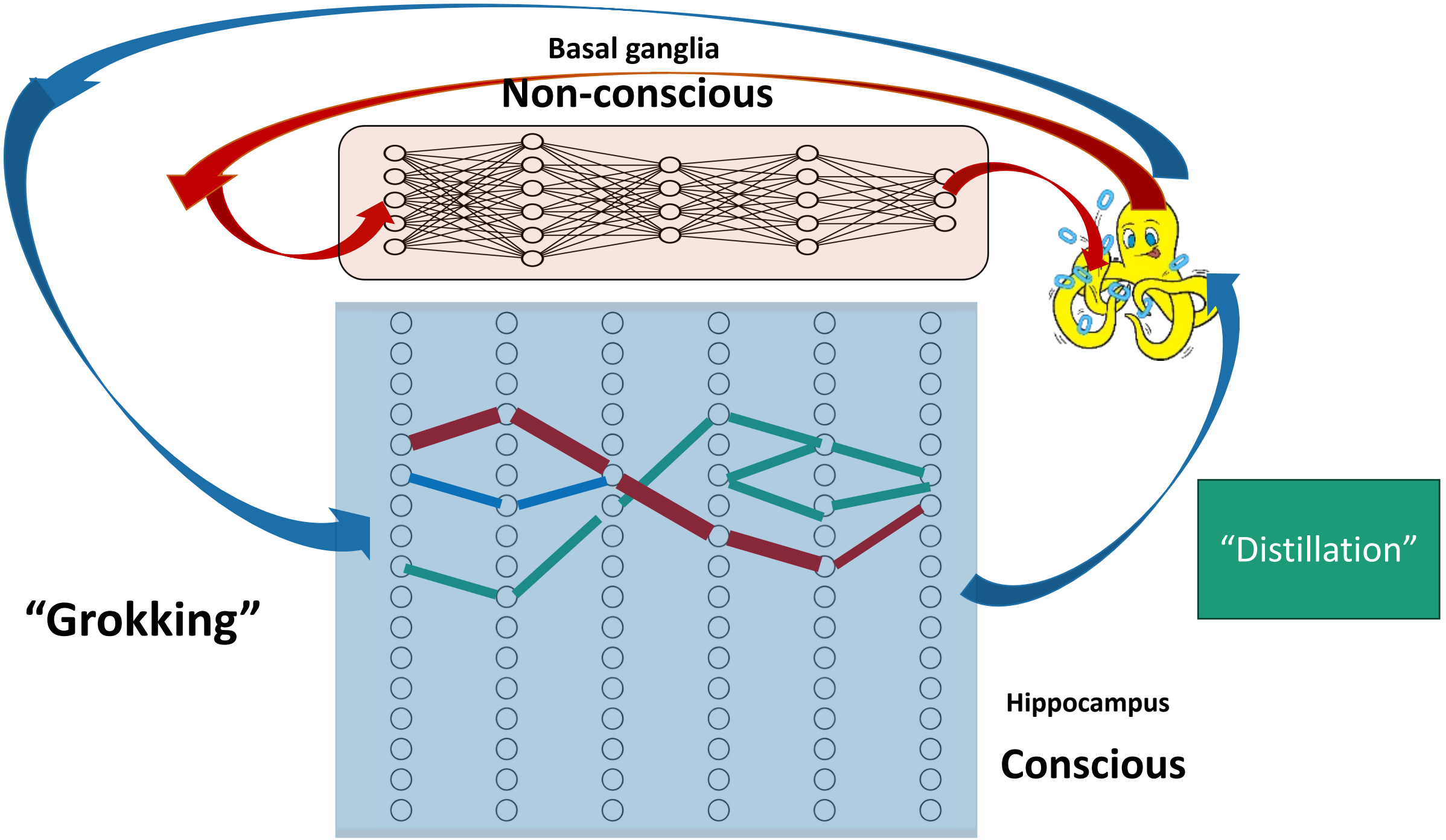


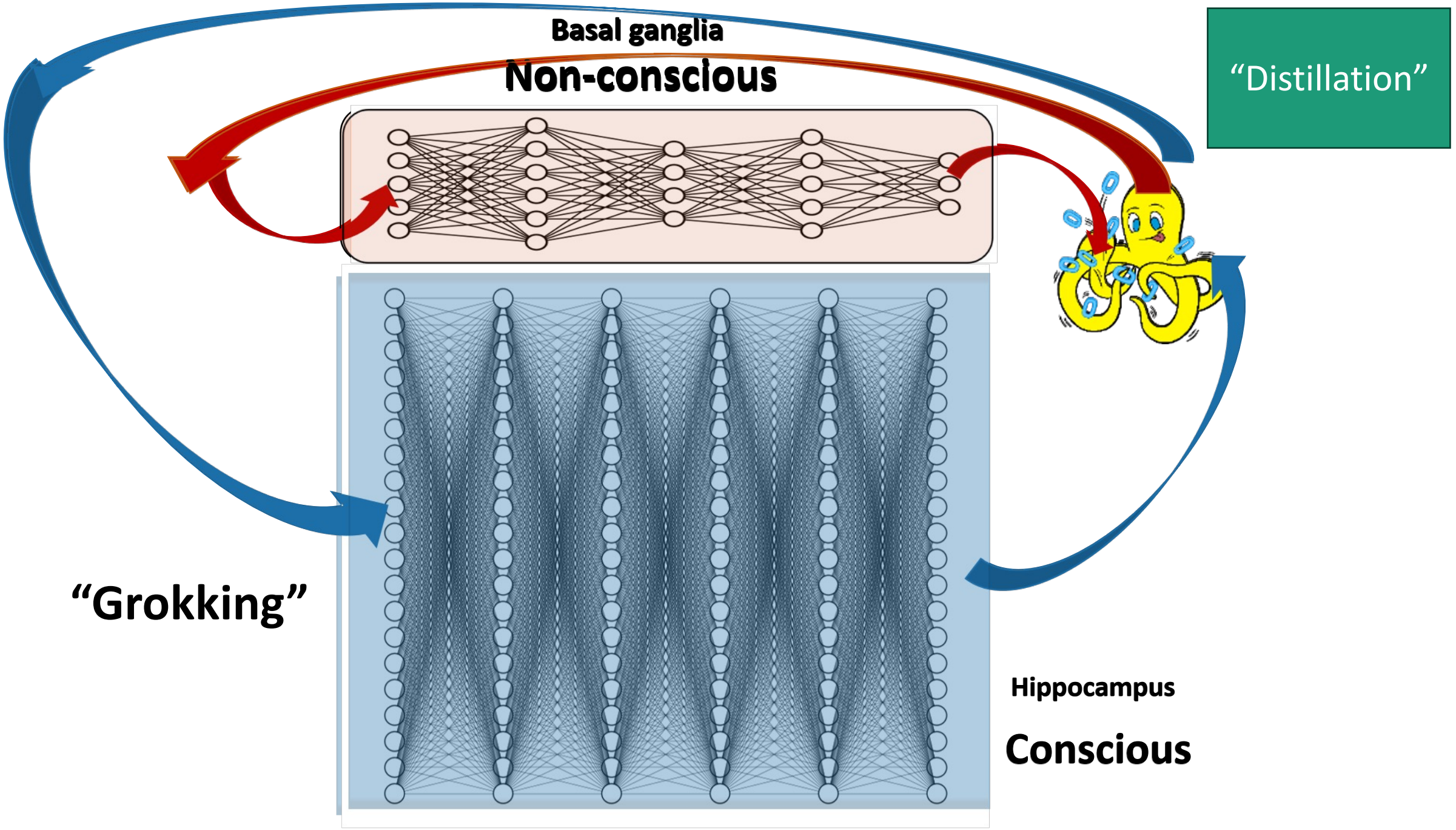
"Large Language Model"

Hippocampus

Conscious

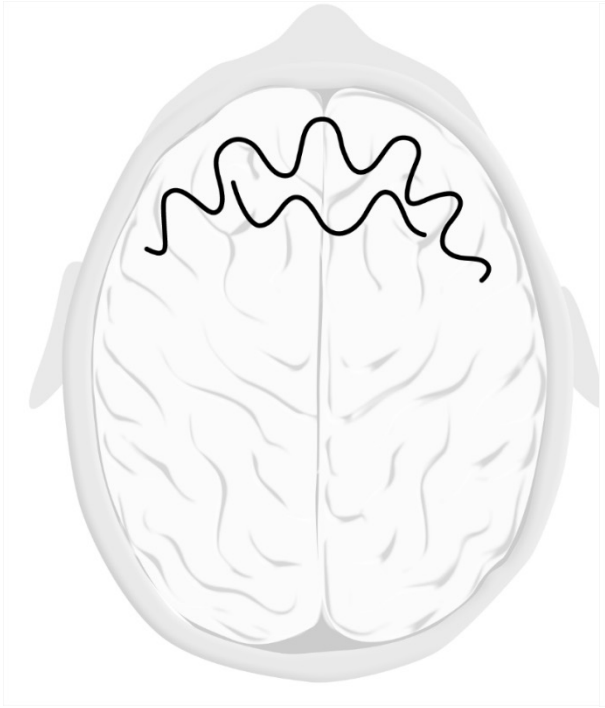

Basal ganglia
Non-conscious
Hippocampus
Conscious


Basal ganglia
Non-conscious
"Grokking"
"Distillation"
Hippocampus
Conscious

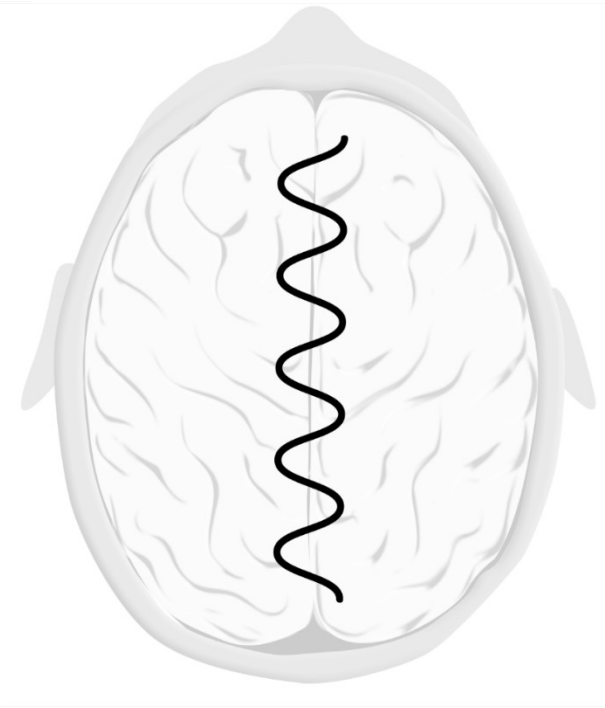

Basal ganglia
Non-conscious
"Distillation"
"Grokking"
Hippocampus
Conscious



Djzoomer



Midline frontal theta



Long range theta coherence

Midline frontal theta



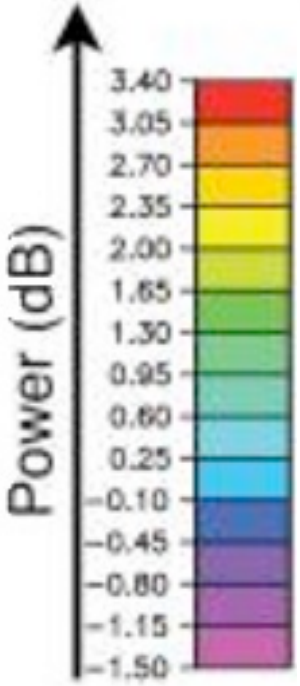
Early 20’s



60’s



60’s—with training



Anguera, J. A., et al. "Video Game Training Enhances Cognitive Control in Older Adults." *Nature* 501, no. 7465 (2013): 97-101.

Output

r you do, don't think about elephants.



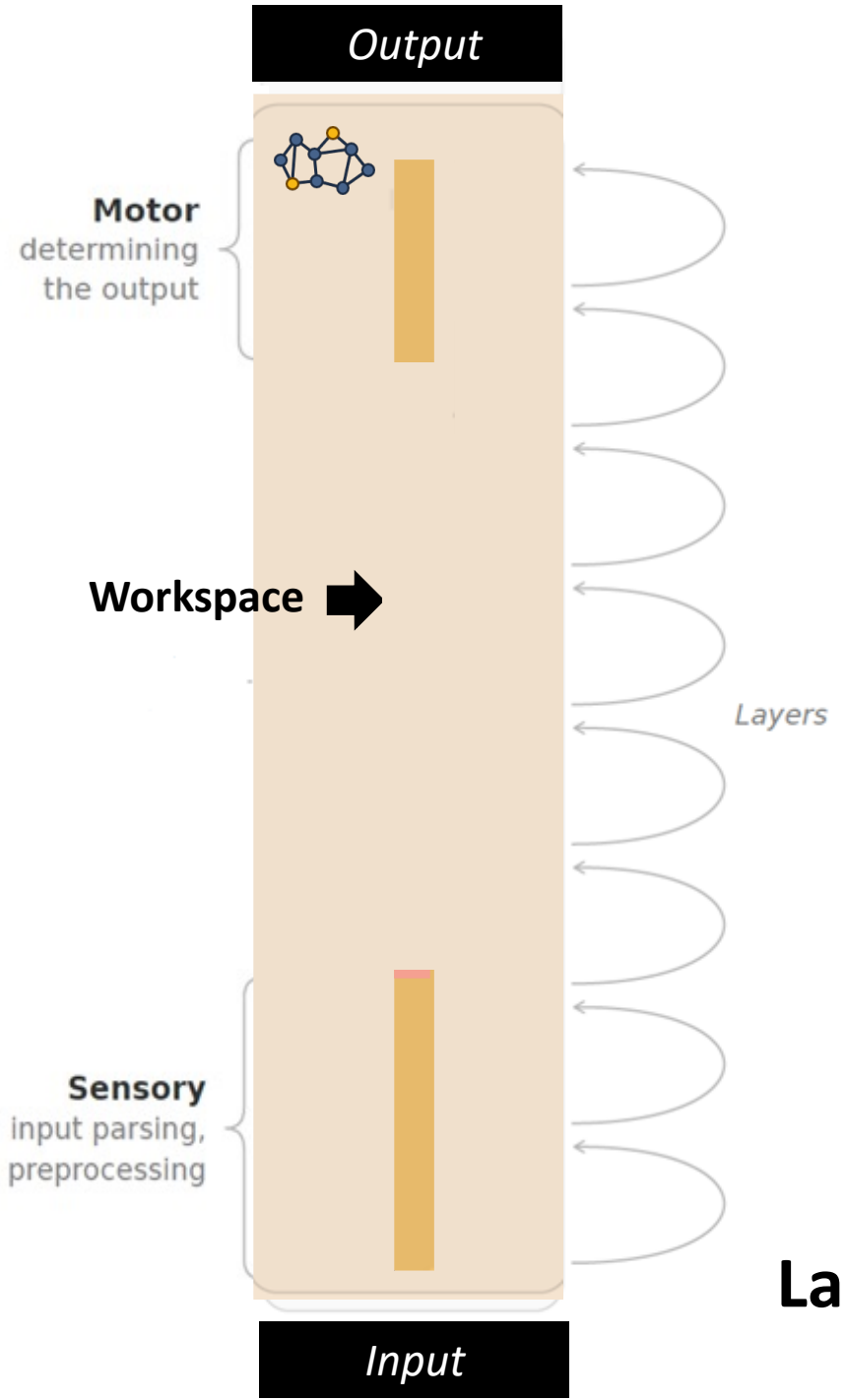


Workspace can hold only 4-6 "chunks" of information at a time!

Large Language Model

After Gurnee, W., et al. (2026). "Verbalizable representations form a global workspace in language models." Anthropic. https://transformer-circuits.pub/2026/workspace/index.html

Input

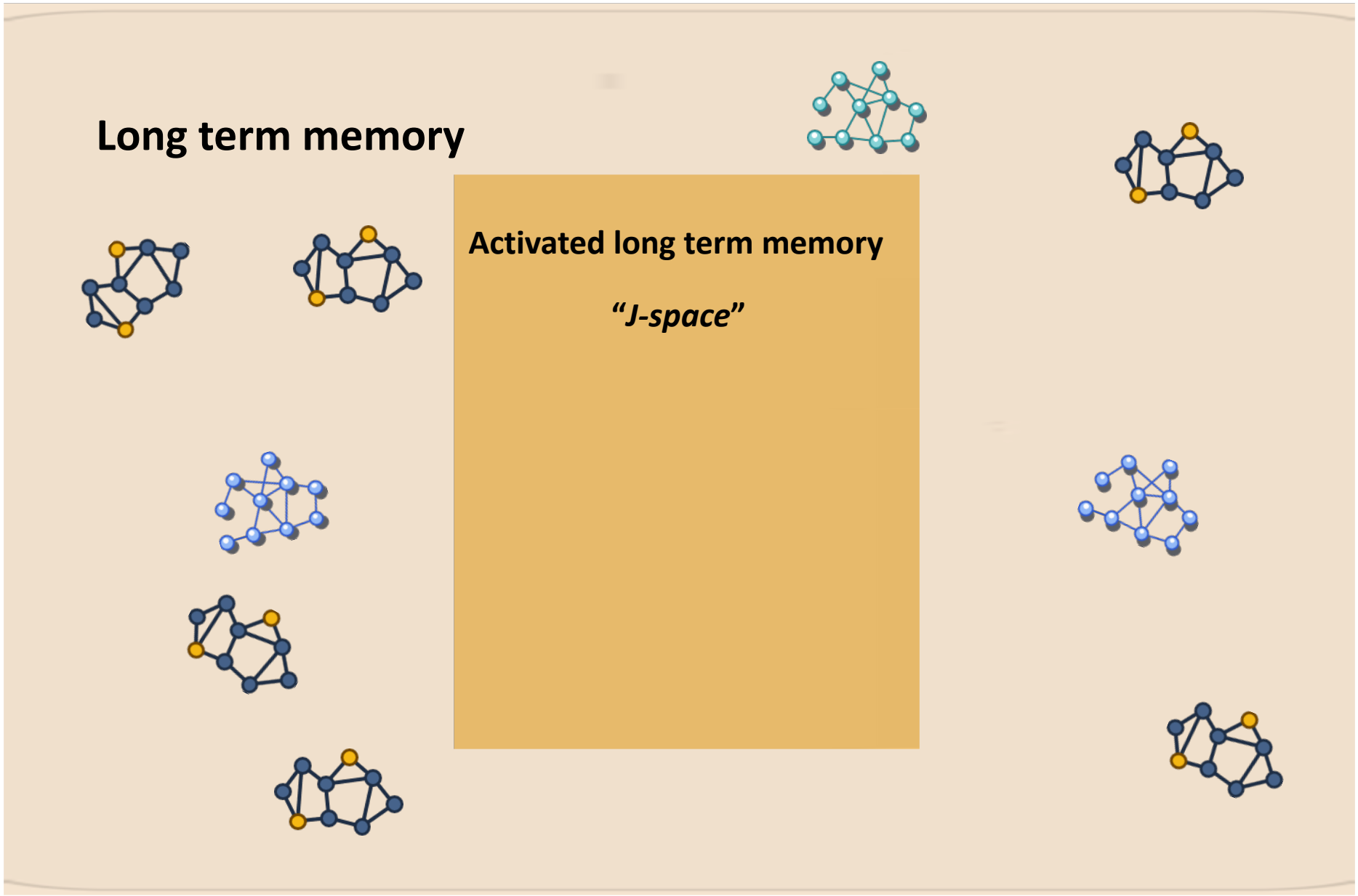


Ablate (empty out) the workspace

The model can “speak”

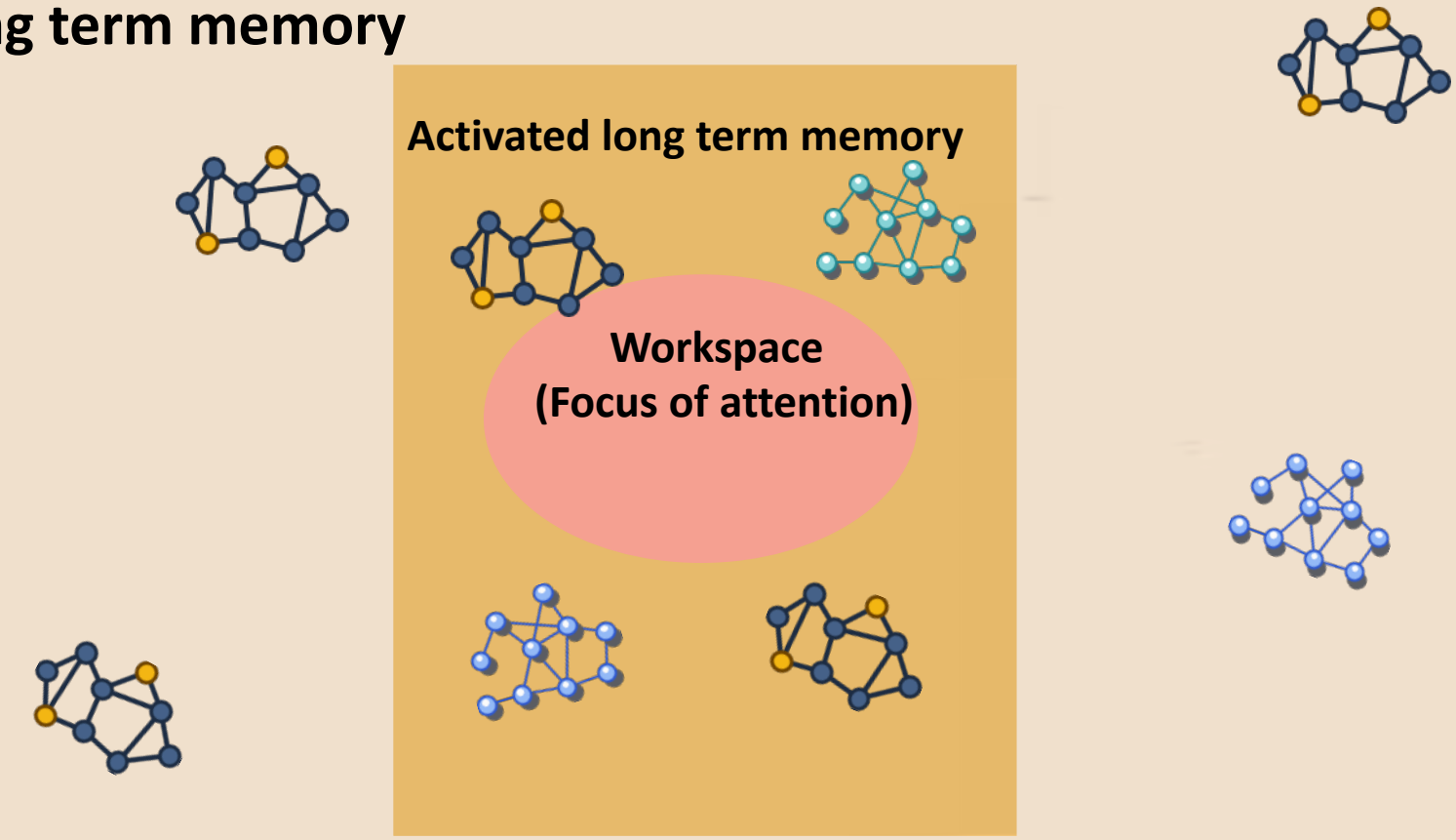
The model cannot “think” critically!

Large Language Model

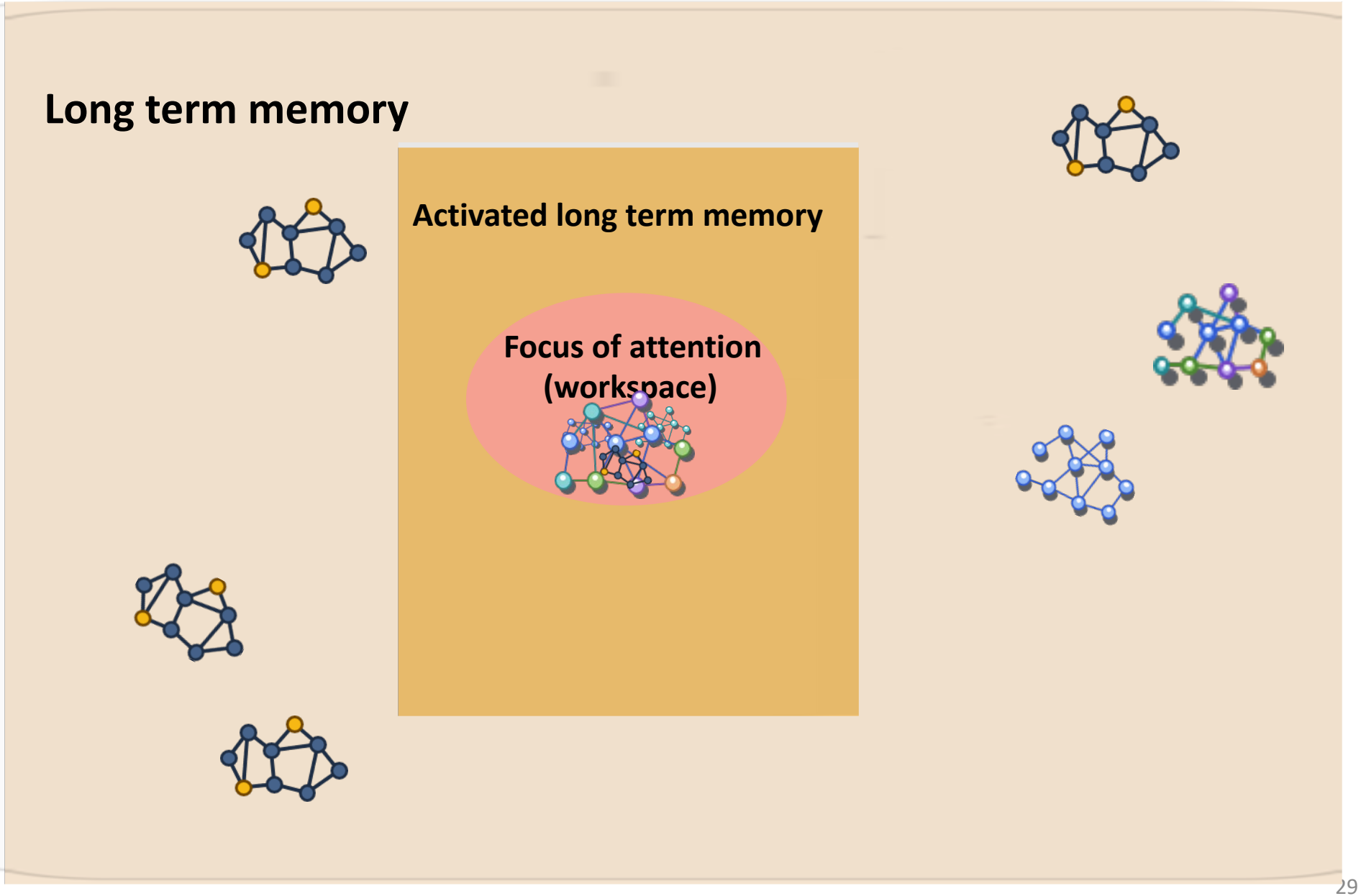


Human memory

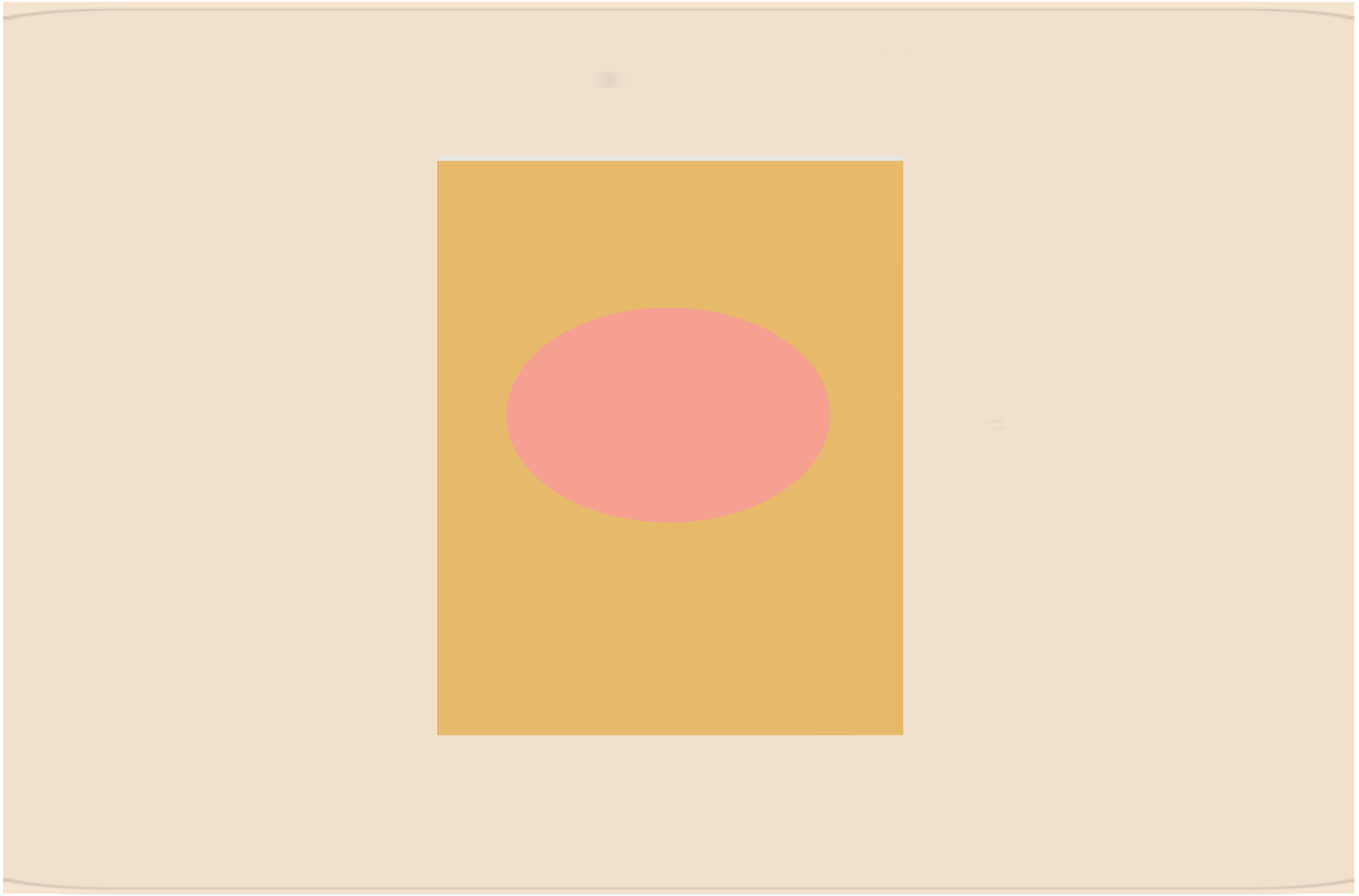
Long term memory



What happens when you use genAI to “cognitively offload”?



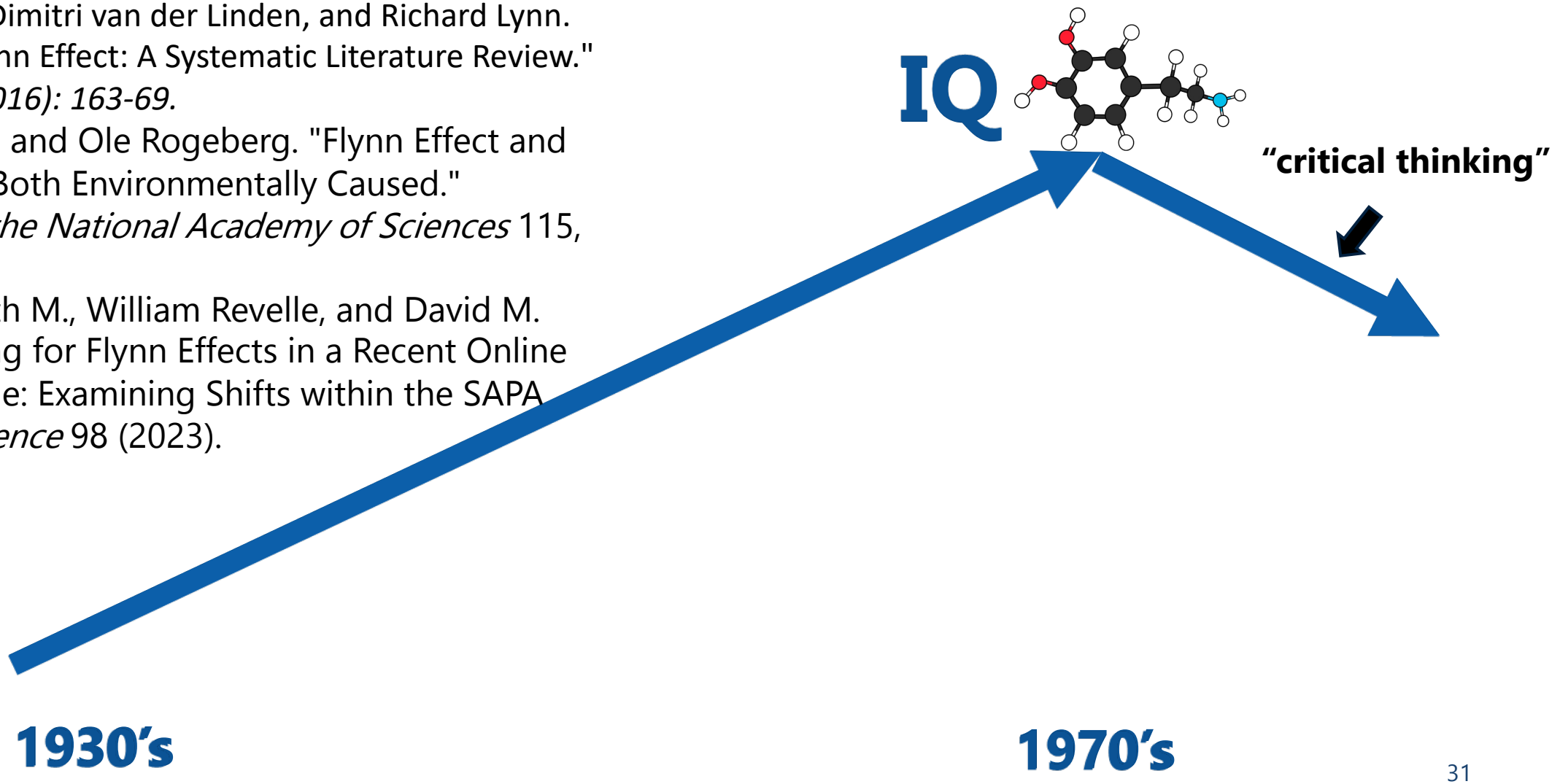
What happens when you use genAI to “cognitively offload”?

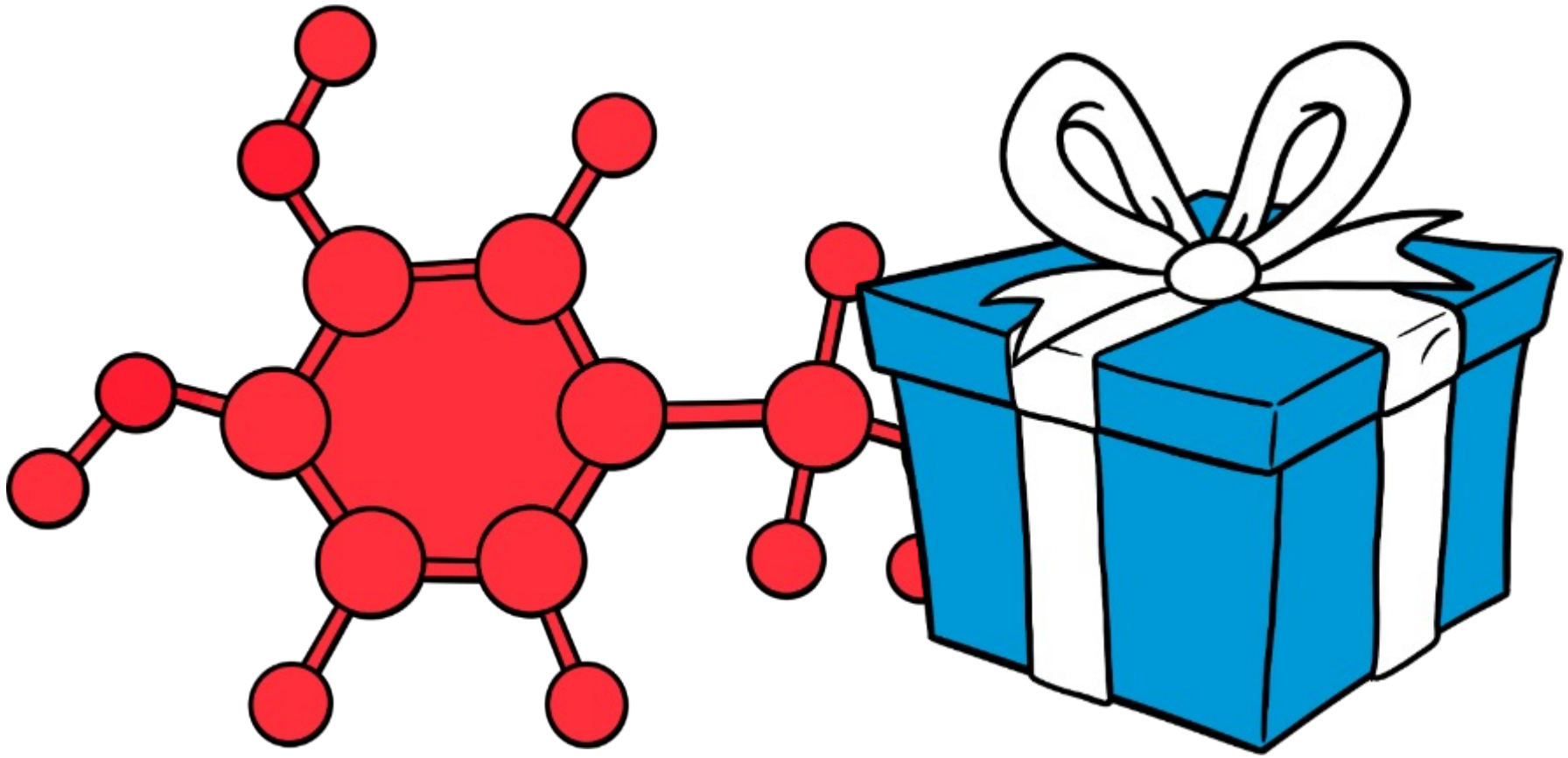


The Flynn Effect

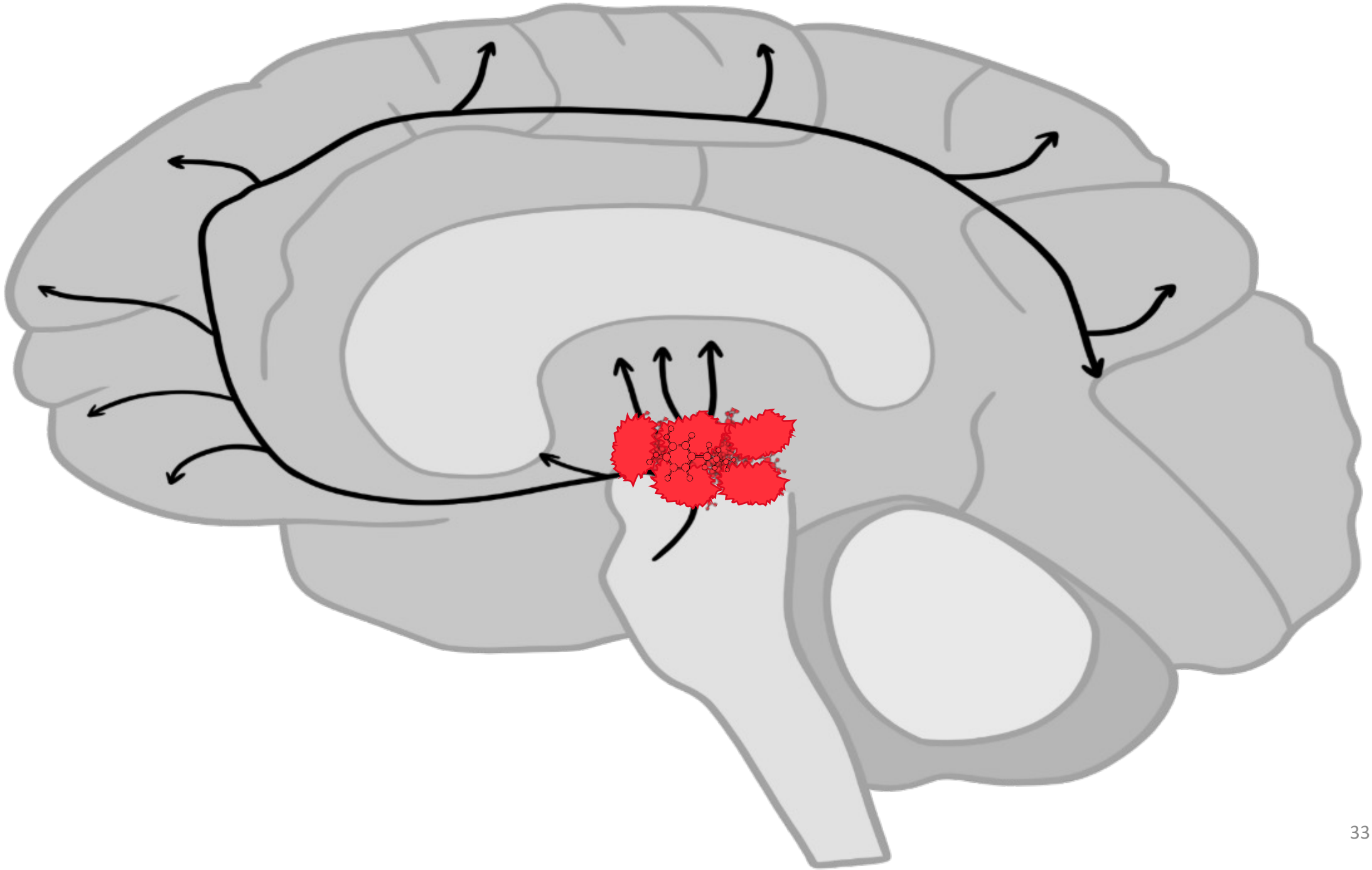
Flynn, James R. "The Mean IQ of Americans: Massive Gains 1932 to 1978." *Psychological Bulletin 95, no. 1 (1984): 29.*

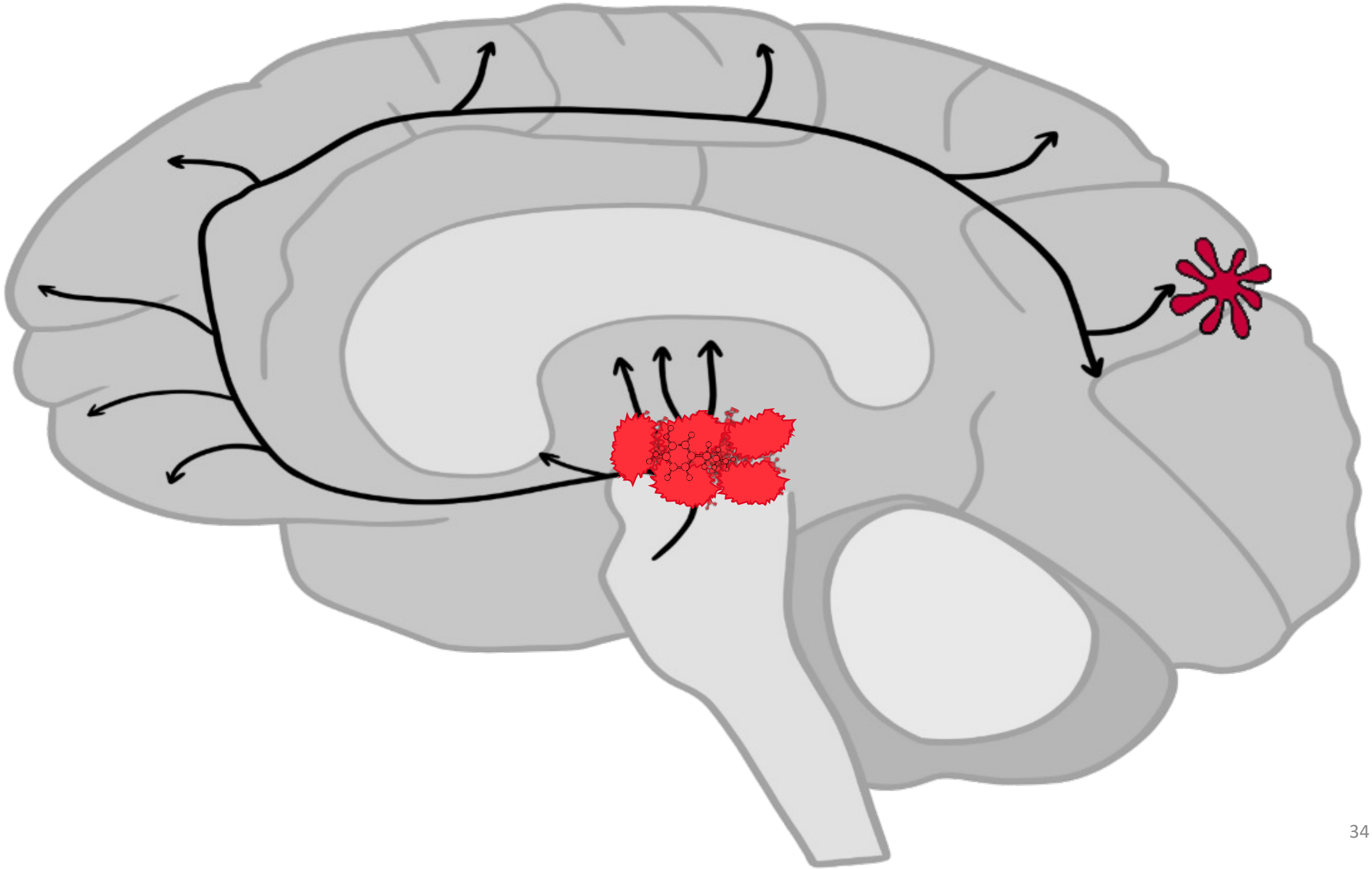
- Dutton, Edward, Dimitri van der Linden, and Richard Lynn. "The Negative Flynn Effect: A Systematic Literature Review." *Intelligence 59 (2016): 163-69.*
- Bratsberg, Bernt, and Ole Rogeberg. "Flynn Effect and Its Reversal Are Both Environmentally Caused." *Proceedings of the National Academy of Sciences* 115, no. 26 (2018):
- Dworak, Elizabeth M., William Revelle, and David M. Condon. "Looking for Flynn Effects in a Recent Online U.S. Adult Sample: Examining Shifts within the SAPA Project." *Intelligence* 98 (2023).





Dopamine—the “feel good molecule”

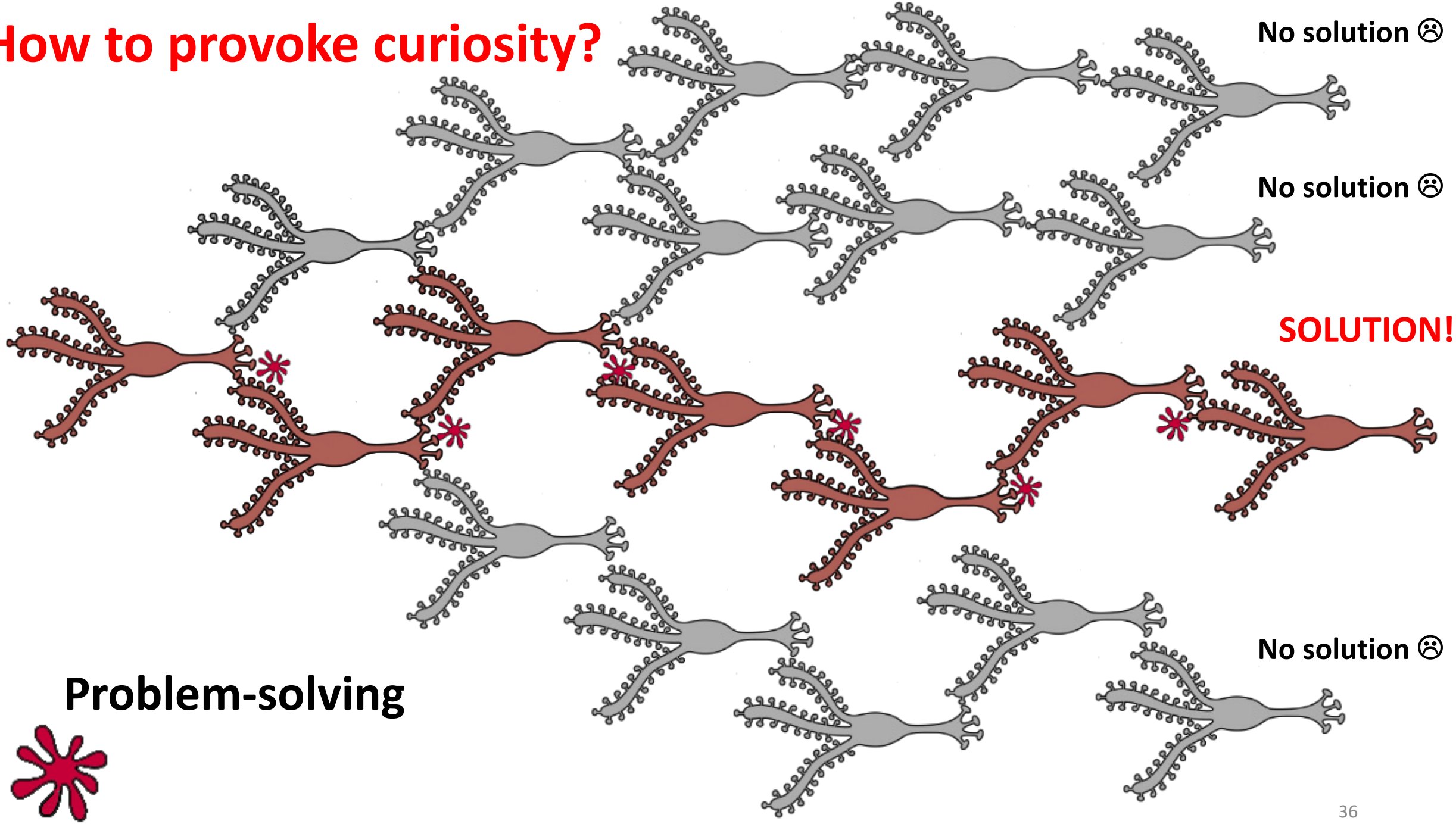






Mice with inactivated dopamine systems *can't learn anything new.*

How to provoke curiosity?

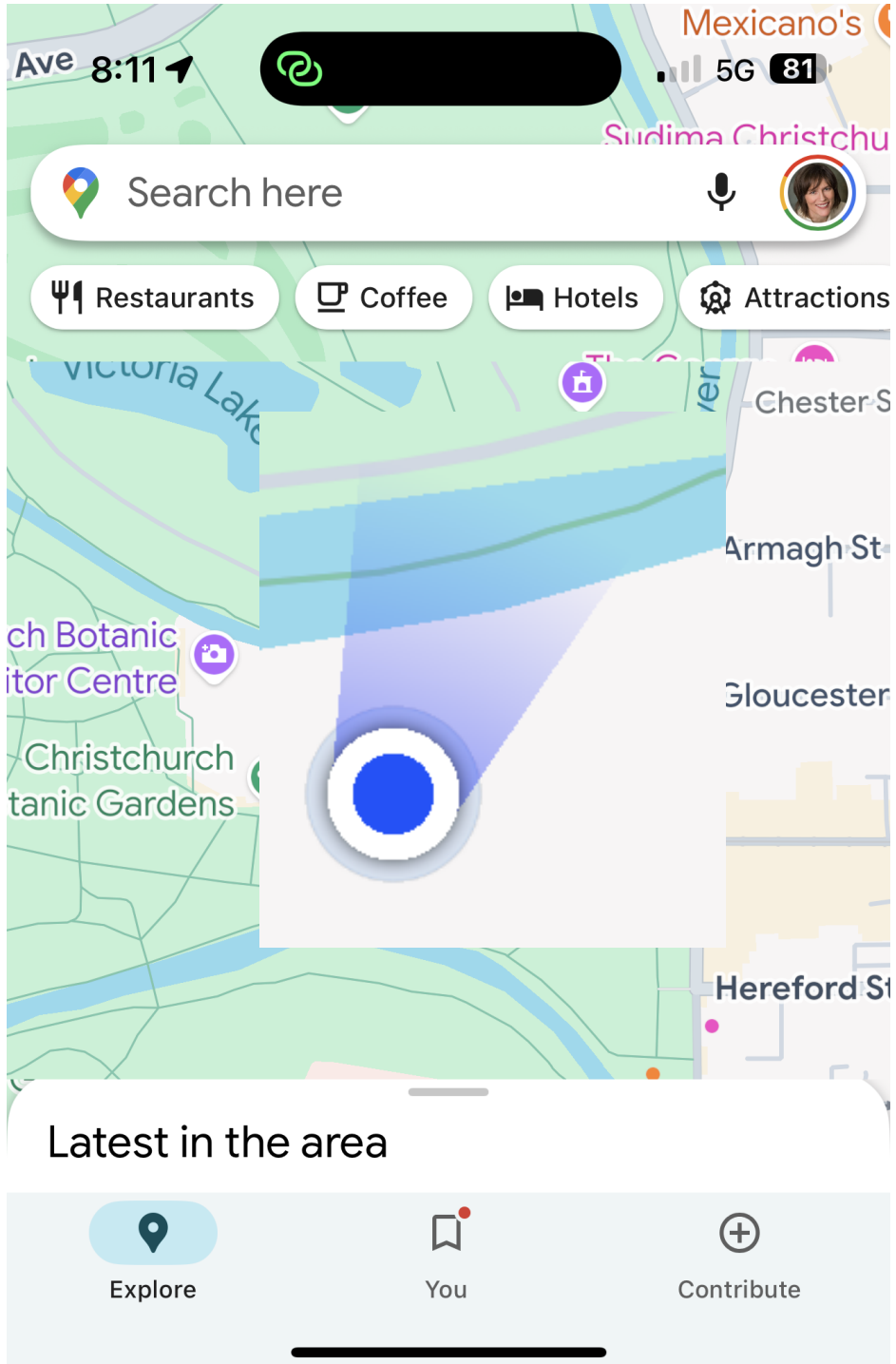


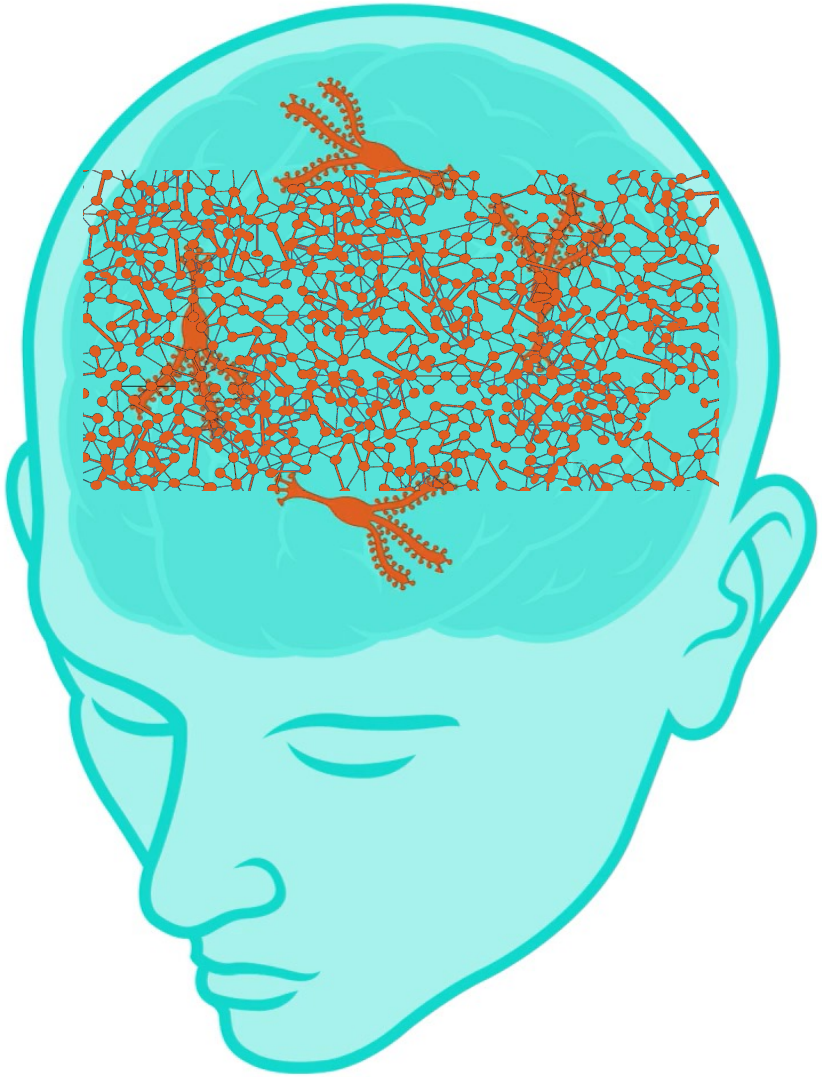
Problem-solving

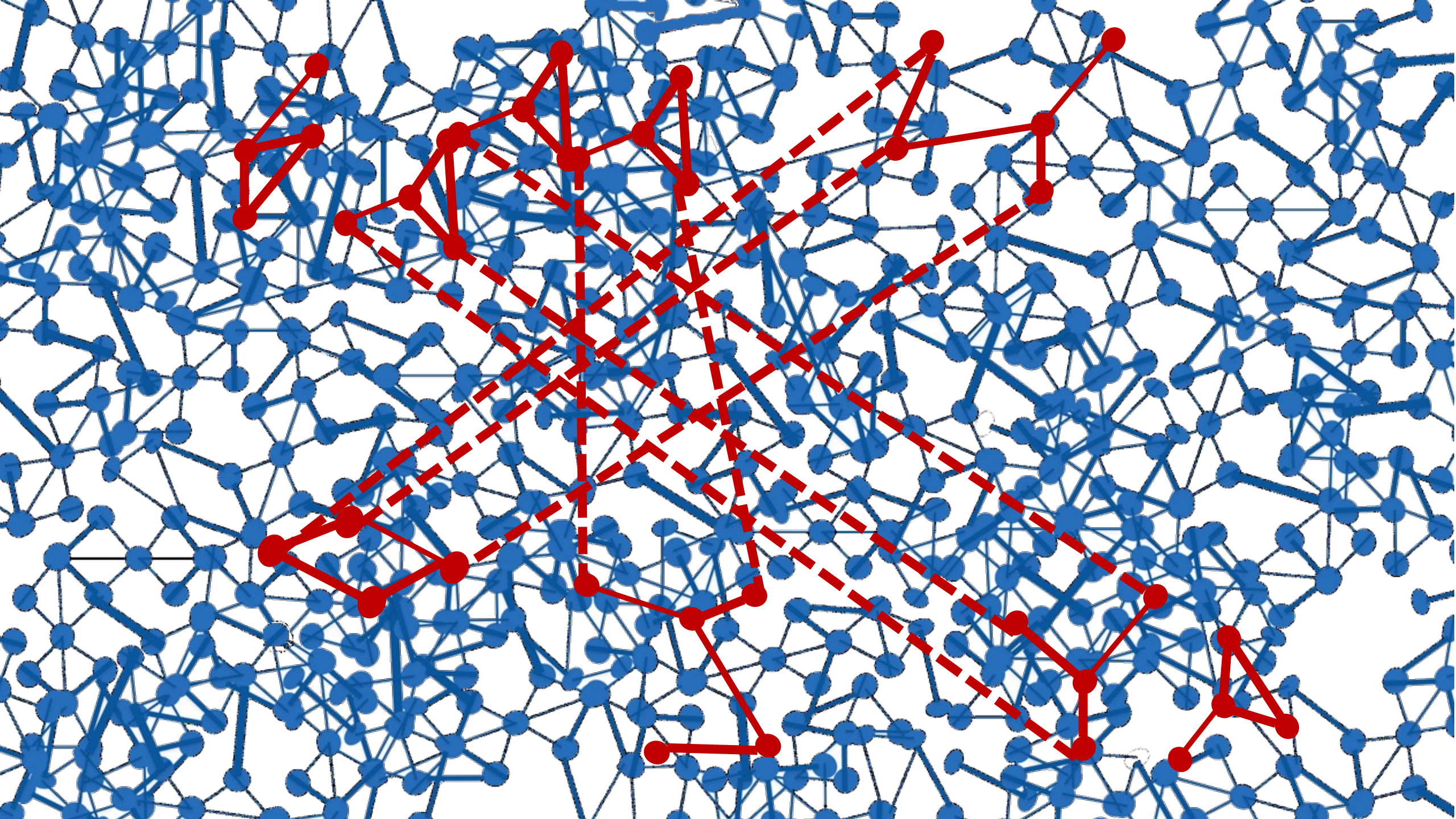
Ask ChatGPT!

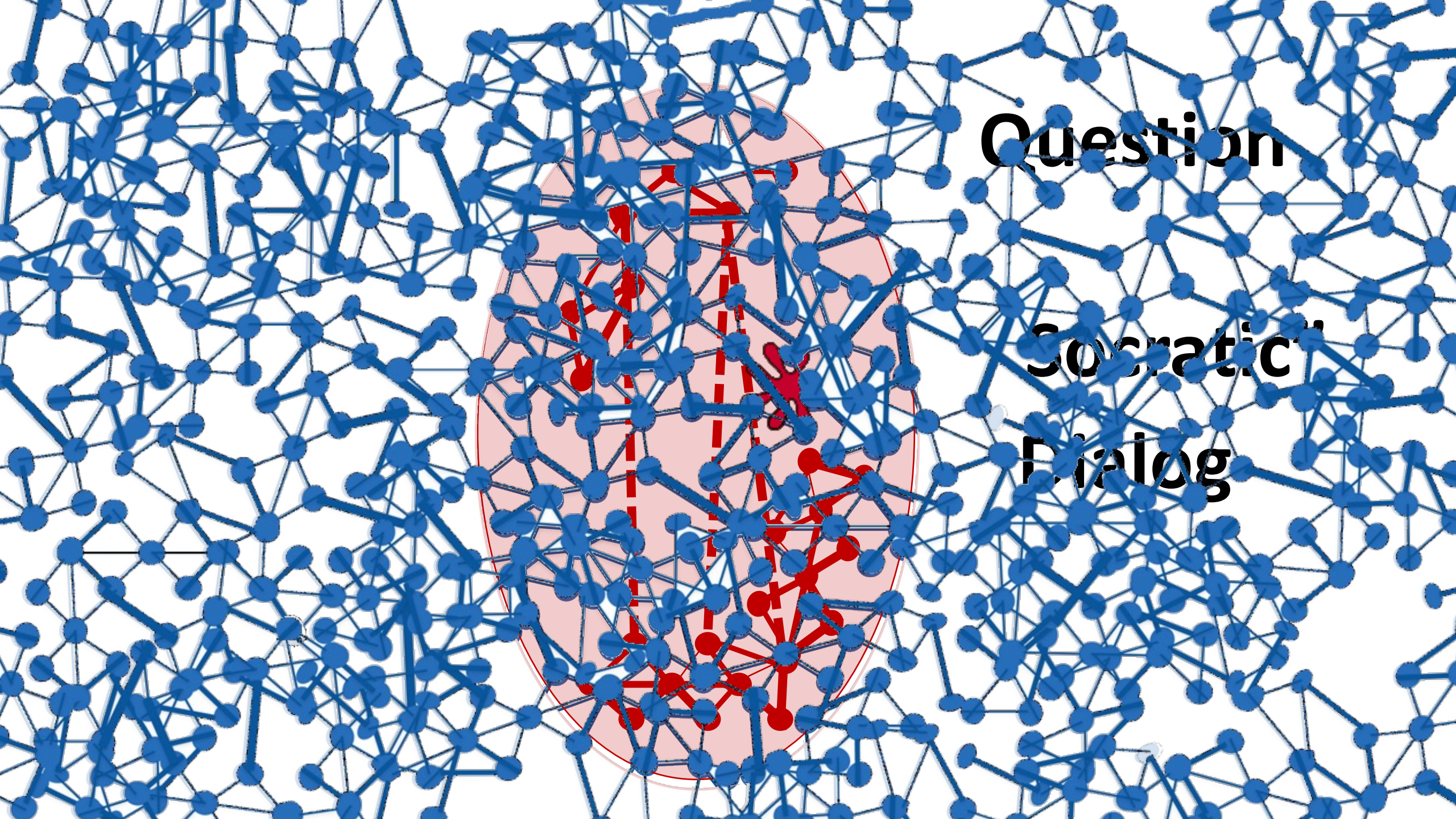
BO My Mongolian college sophomore students are super bored when I try to teach them about main guard versus scope in Python programming. How can I connect these ideas to the nomadic ways of life that my students grew up with?

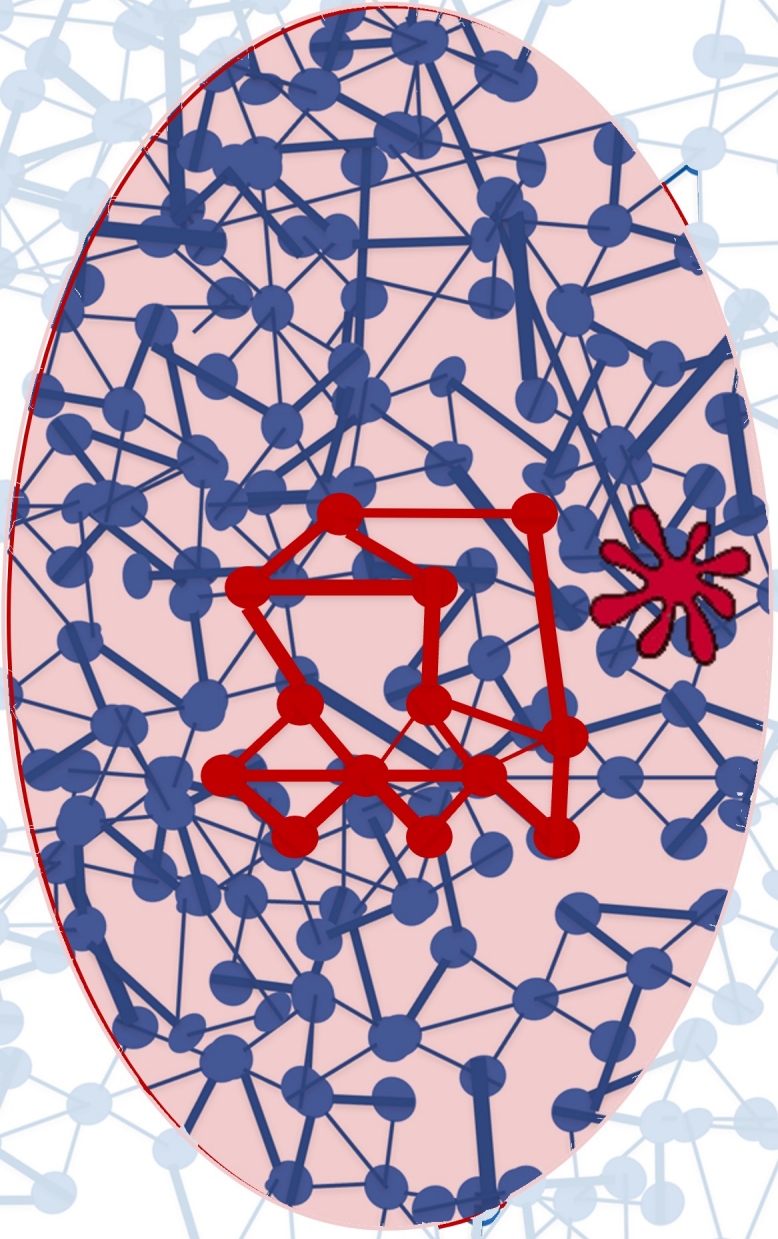



Mexicano's
Ave
8:11
5G
81
Search here
Restaurants
Coffee
Hotels
Attractions
Victoria Lake
Chester S
Armagh St
ch Botanic
itor Centre
Gloucester
Christchurch
tanic Gardens
Hereford St
Latest in the area
Explore
You
Contribute

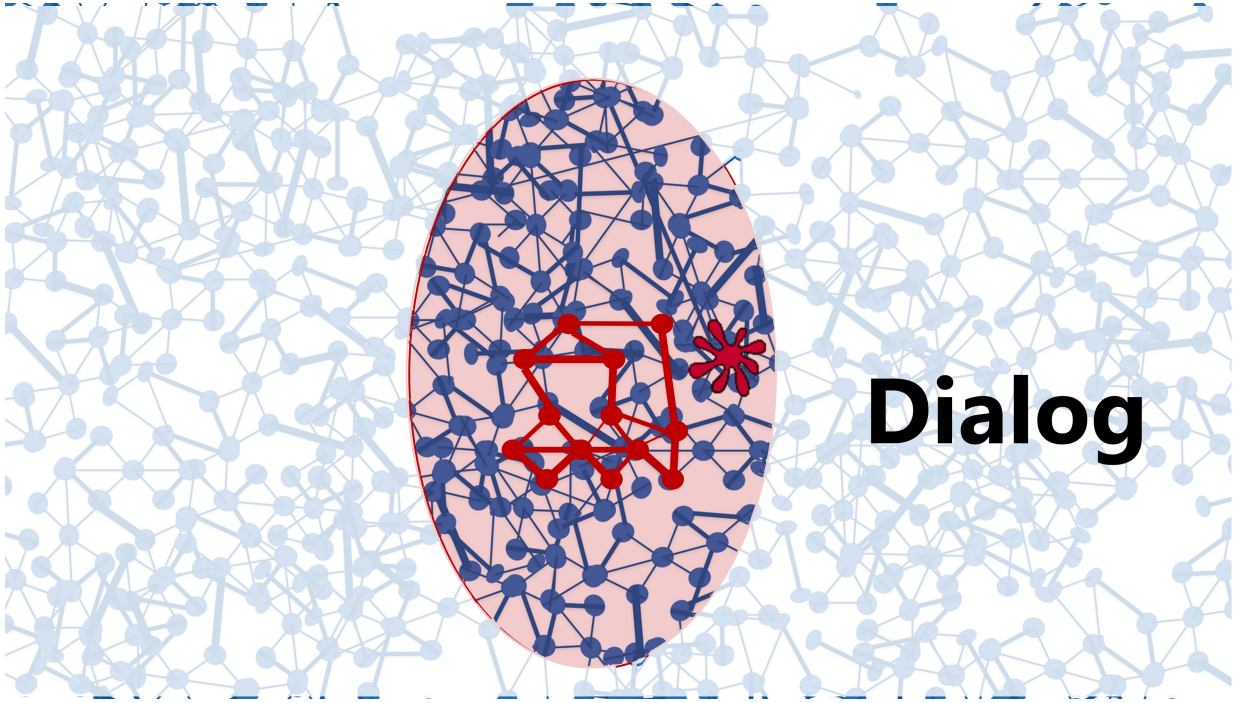


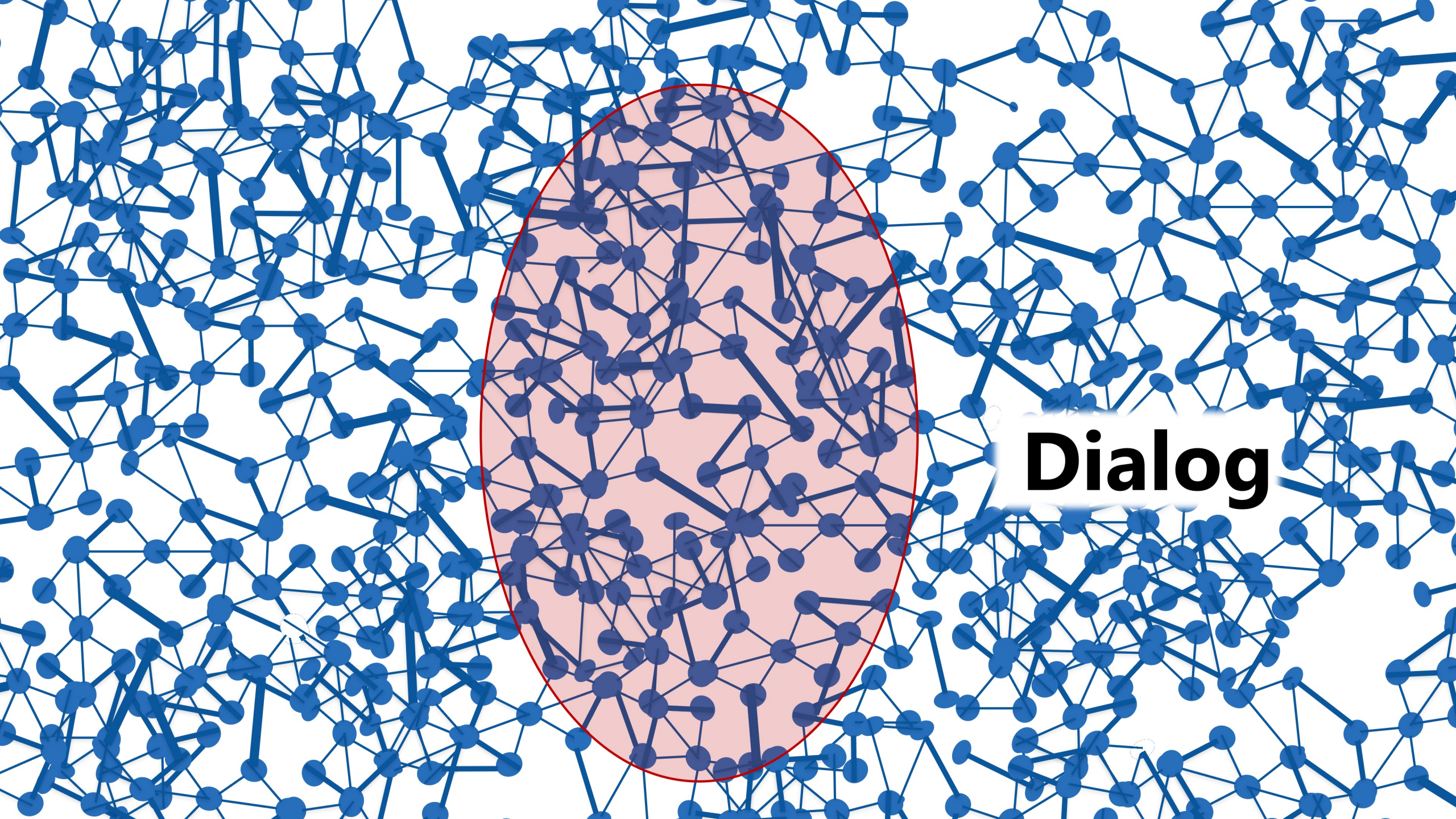


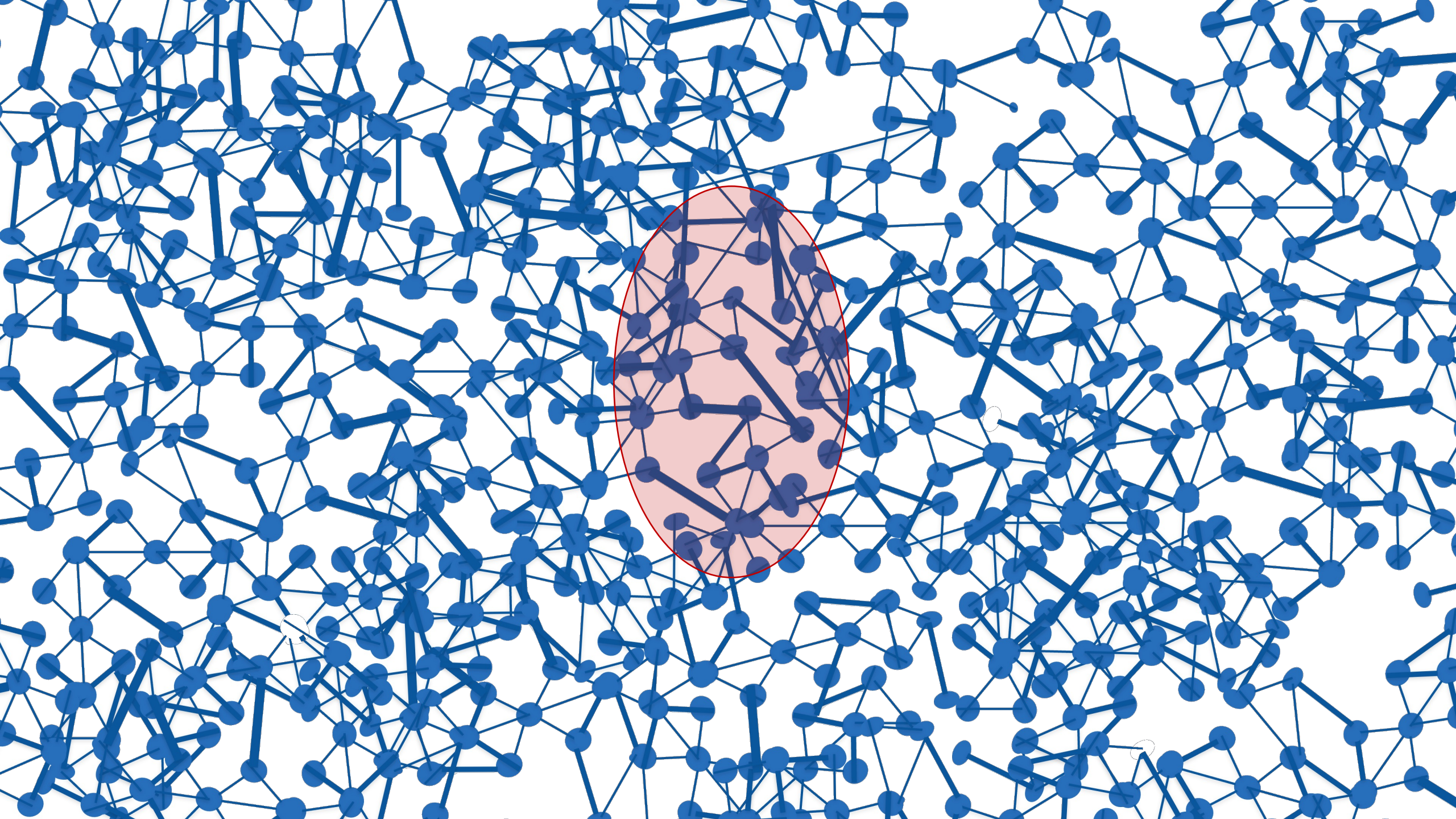

Question
Socratic
Dialog

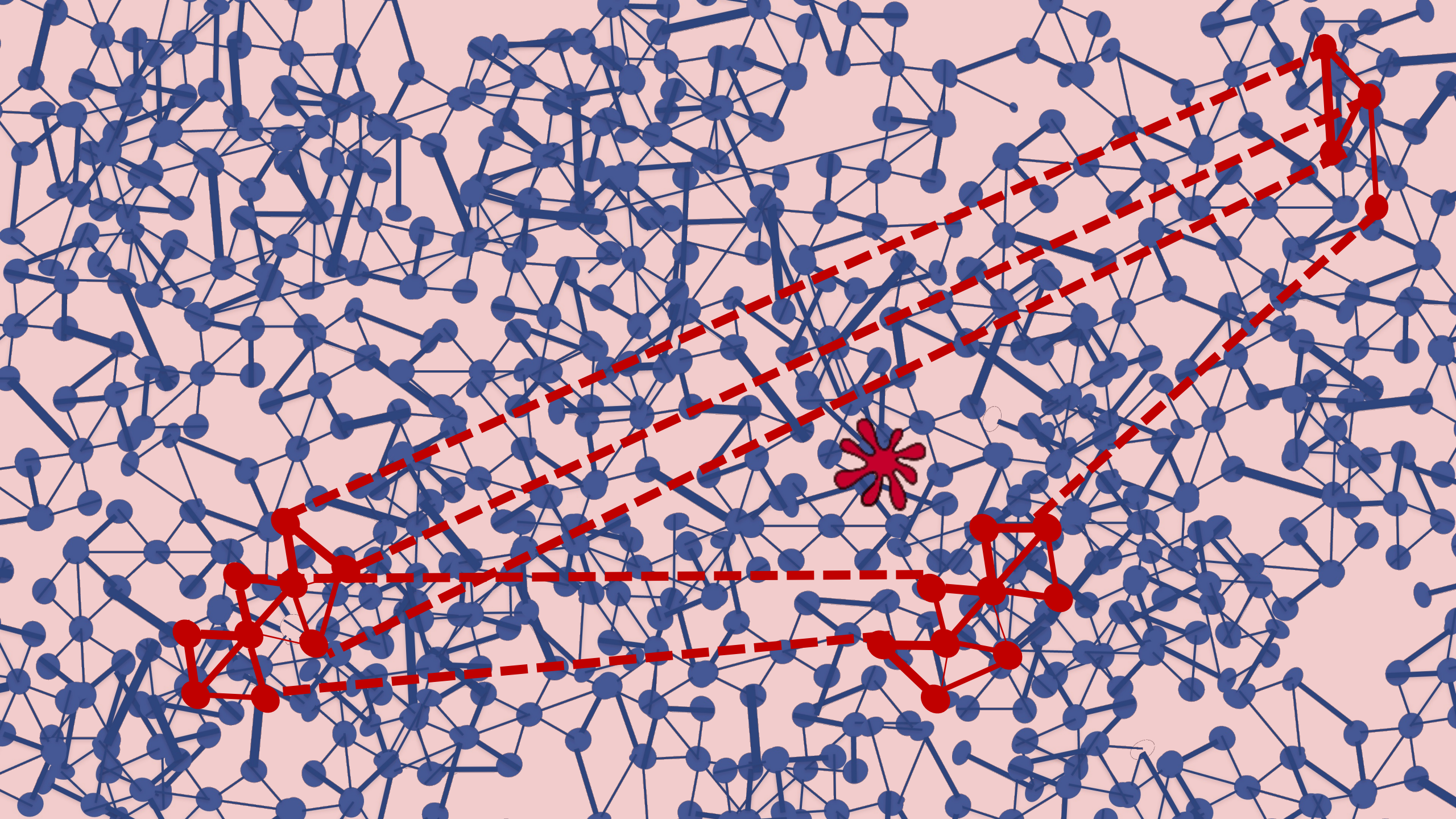


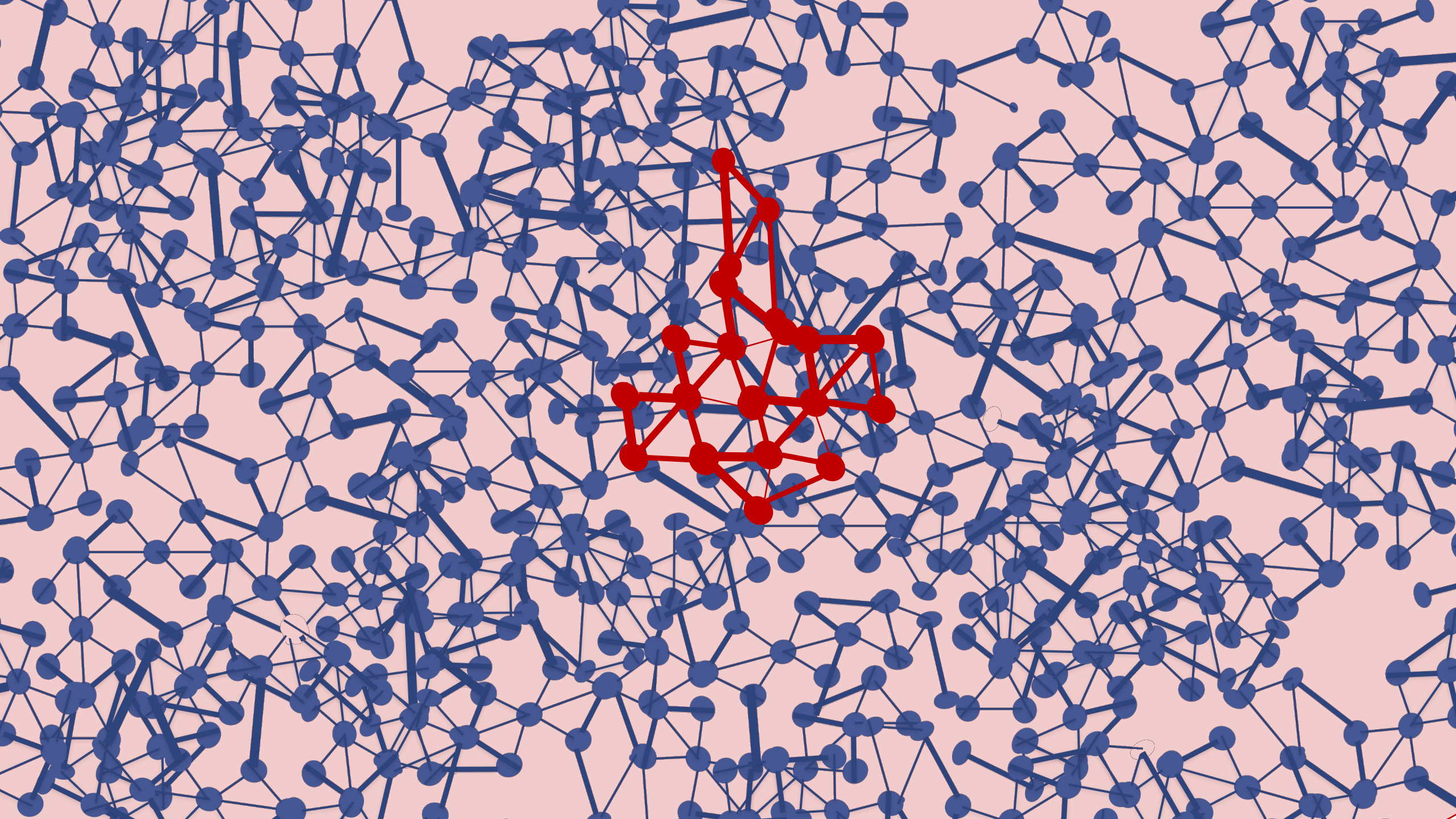
Dialog

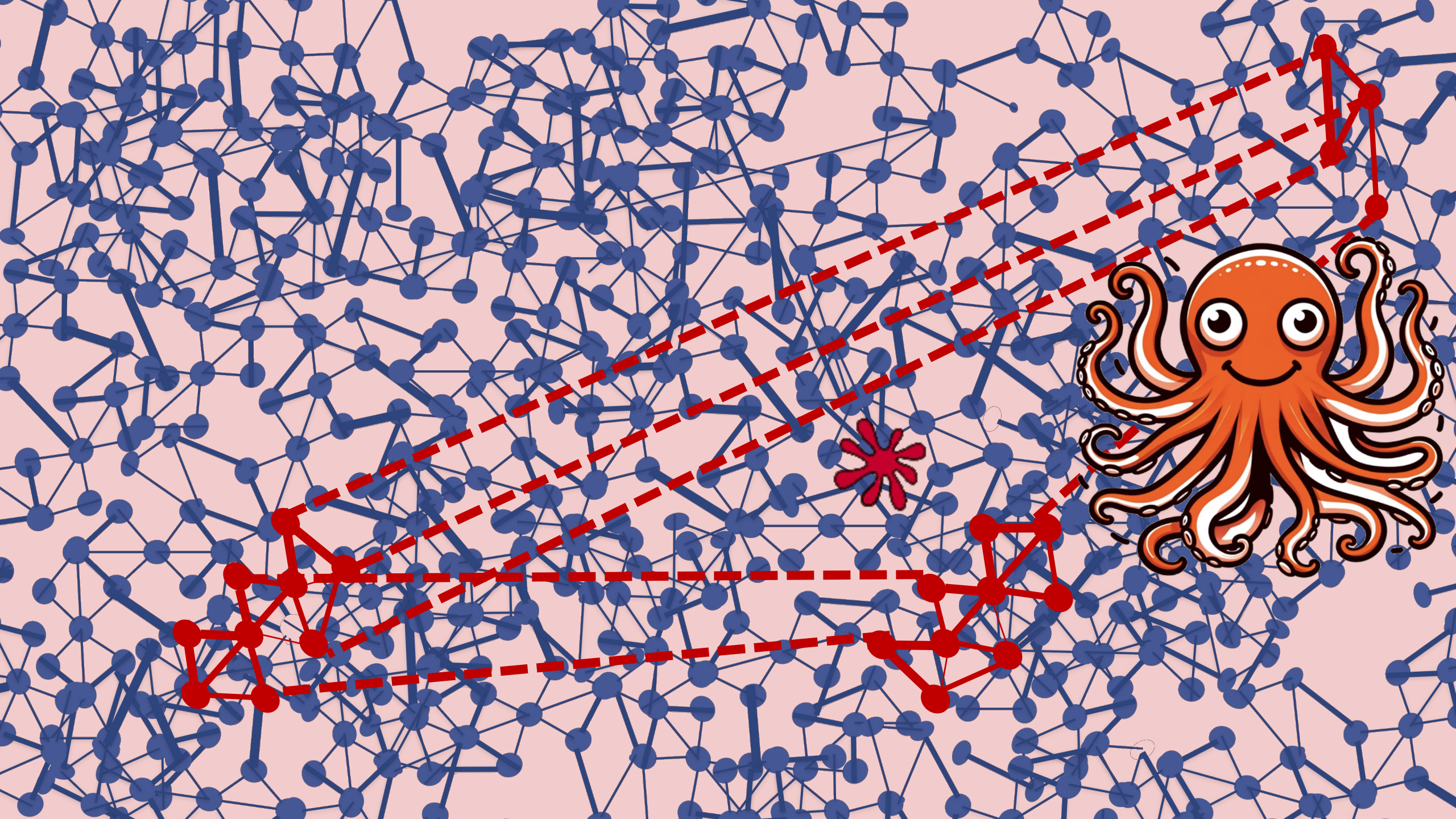

Dialog

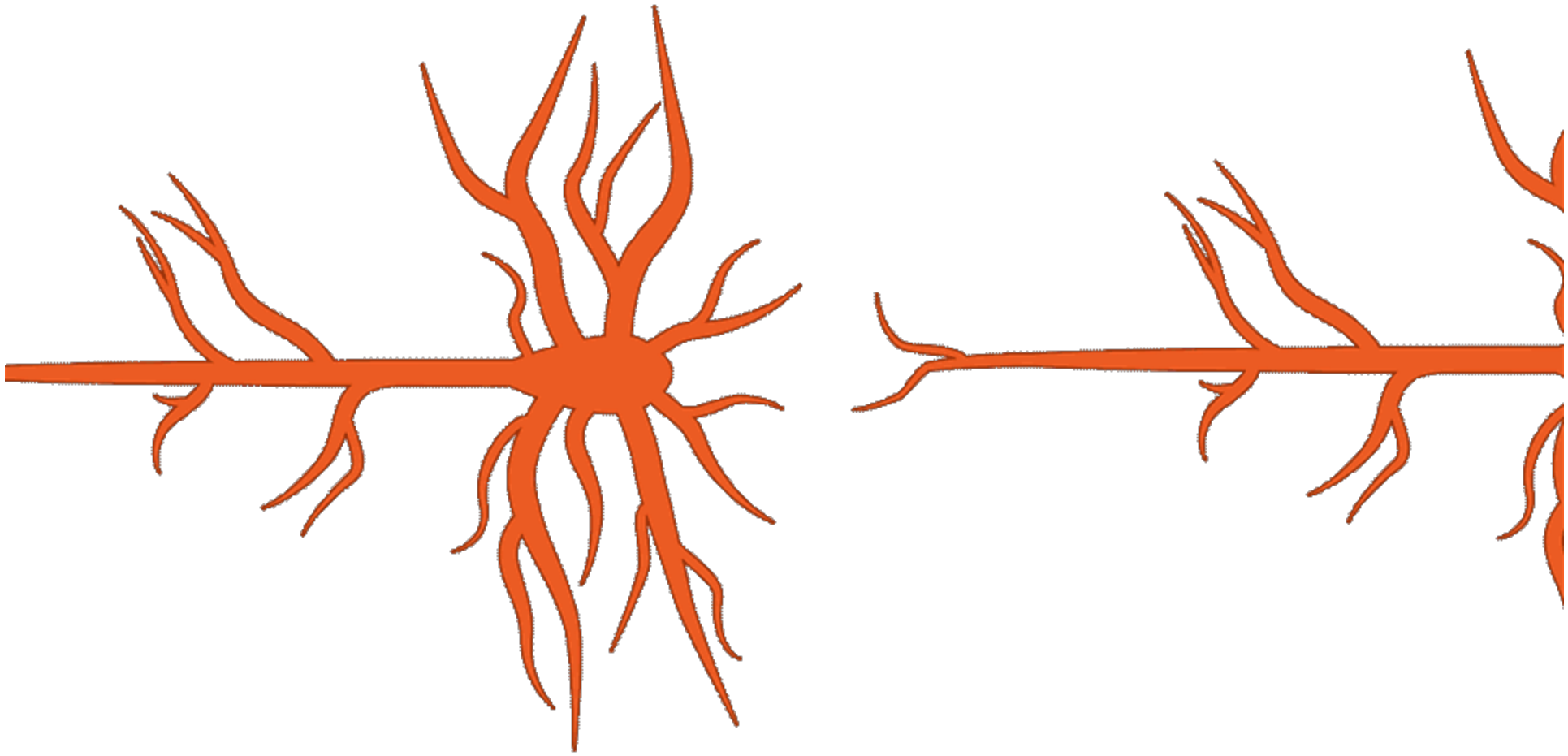

Dialog

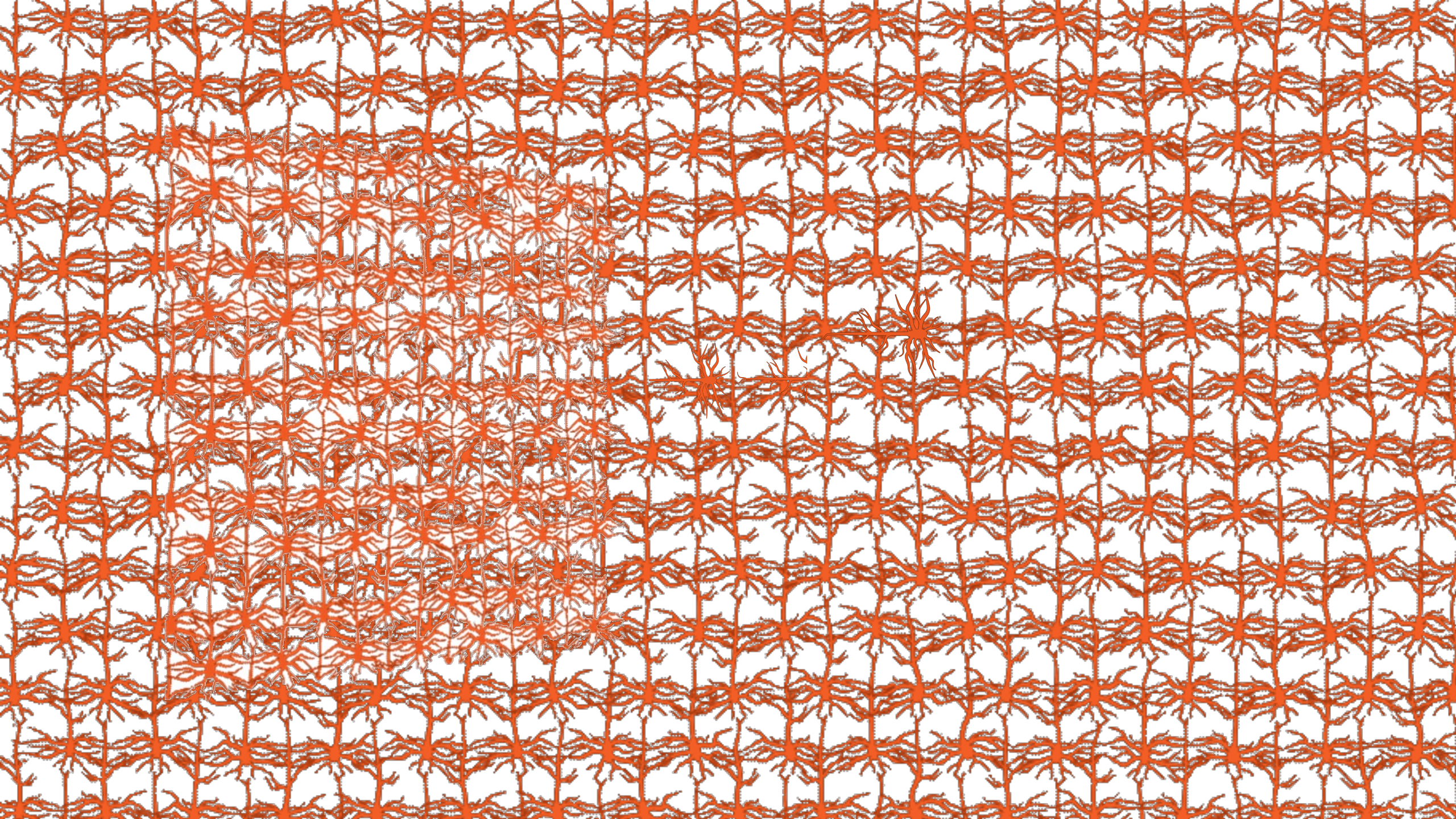


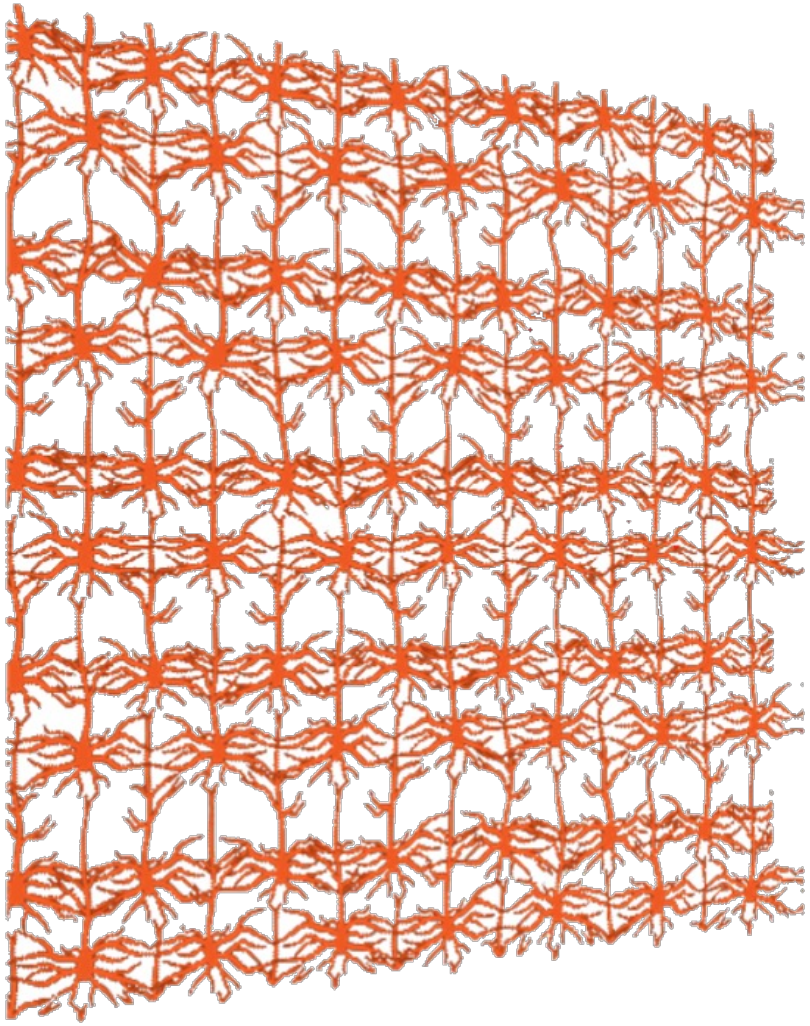


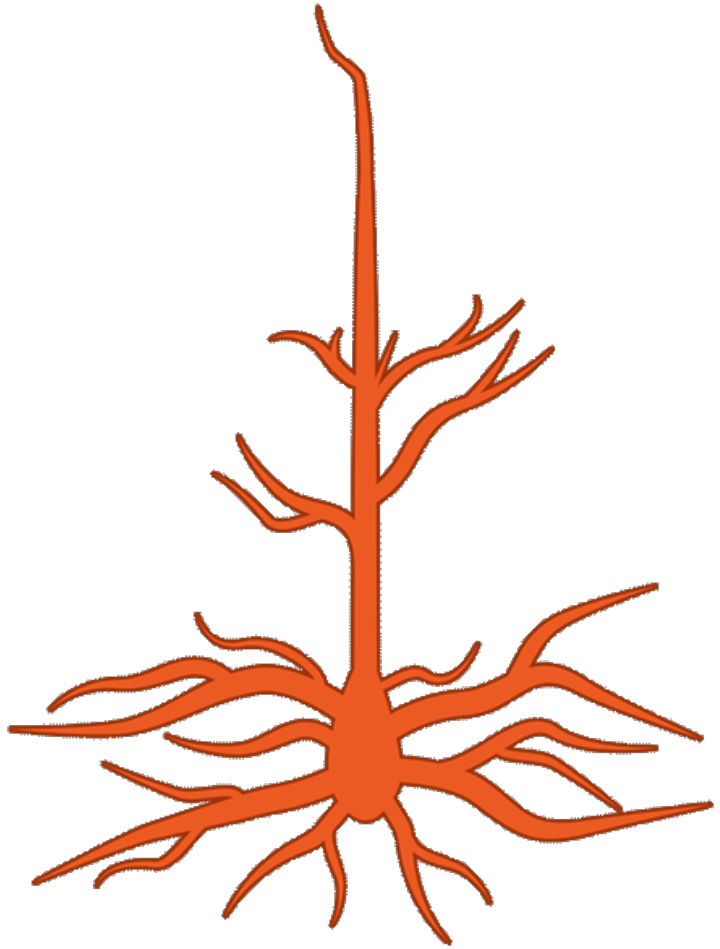




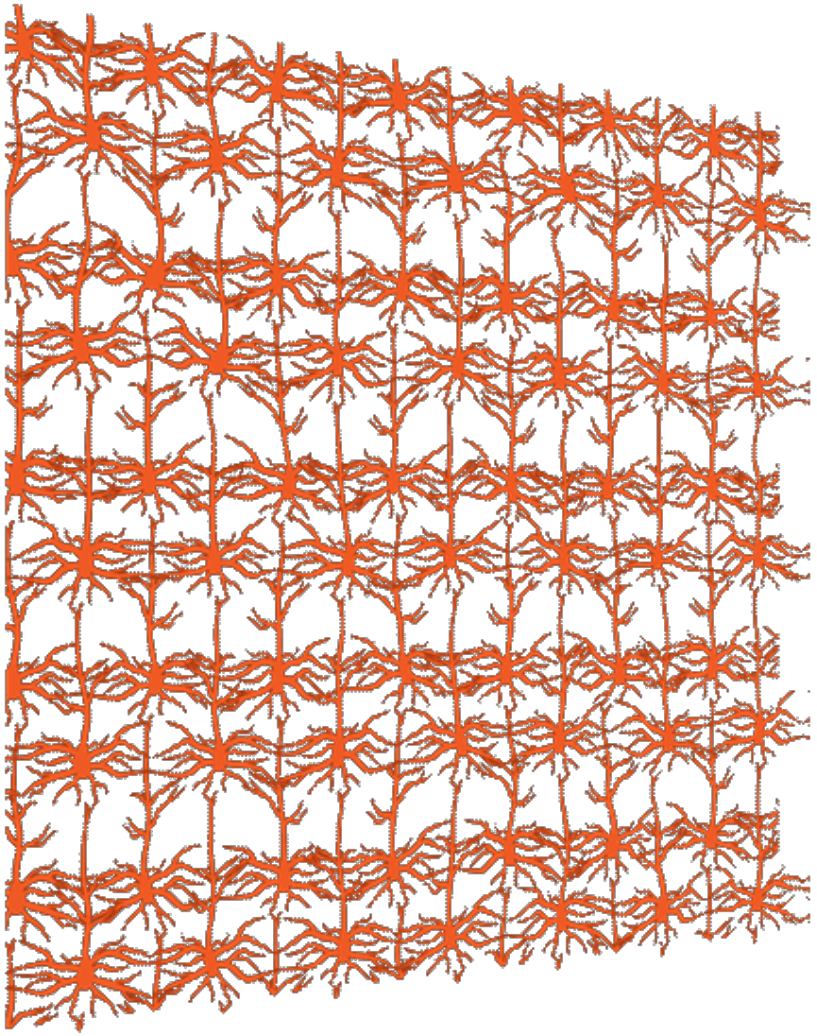








New learning boosts spirits



QUESTIONS?
ASK ME IN THE APP
App Store
Google Play
Download the presentation

QUESTIONS?

ASK ME IN THE APP

Download on the App Store

GET IT ON Google Play

Download the presentation

RATE MY TALK

GET REWARDED

Grokking the Brain:
What AI Reveals About How You Really Learn

Train an AI model long past the point where it has memorized its examples, and something strange happens — it stops memorizing and suddenly "groks" the pattern beneath them. That hard-won pattern can then be "distilled" into a smaller, faster model that runs it with ease. These are ideas from AI — but they may give us our clearest view yet of the human brain. We grok a hard skill slowly, with conscious effort, then distill it into a fast, automatic system that needs no effort at all. That is what expertise is.

The same “machinery” has a quieter side. A model is aligned by its training data — and in an analogous fashion, so are humans. Give me a child to the age of seven, the old saying goes, and I'll show you the adult. But adults are aligned too, usually without knowing it — and a brain aligned a little too well can harden into a mental fortress, deflecting new ideas.

Curiosity provides a way out. Dopamine—reward-based learning—fuels the brain's ability to rewire — it helps us build new links, keeps the fortress gate open, and keeps hard-won knowledge from setting into a wall. We'll look at what all this means for how you learn, how you use AI, and how you keep thinking for yourself in an age that makes it easy not to.