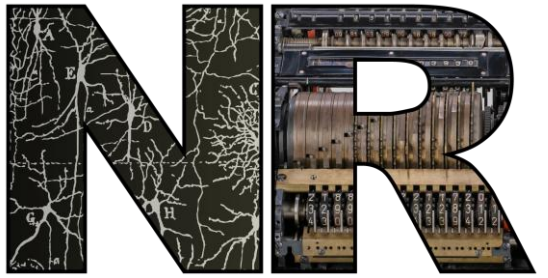
Modules in the garden of forking paths

Dan Goodman

Imperial College London

Department of Electrical and Electronic Engineering

IMPERIAL



neural-reckoning.org

Outline

Part 1

Philosophy

Sorry, but it’s for your own good 😉

Part 2

Research

Resource efficiency of the brain and modularity

There is a link between these… I think.

The garden of forking paths

P-Hacking: selecting data to get a statistically significant result

"Model hacking": same thing but using a model

But these assume that the researcher is unethical

Garden of forking paths: the same thing can happen unconsciously

- Origin: The Garden of Forking Paths (Borges, 1941)
- Paper: Gelman and Loken (2013, unpublished)

Happens whenever researchers make data dependent design choices

This happens all the time in modelling and is unavoidable: we refine our model as we go

Is there a solution?

Maybe, we'll come back to that.



The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time*

Andrew Gelman† and Eric Loken‡

14 Nov 2013

This photo is not mine, but it's from a hedge maze I used to go to as a child, on the Isle of Wight

“NeuroAI” in the garden of forking paths

Imagine we want to test the hypothesis “brain connectome is optimal”

We think about the space of all 2^{n^2} possible connectomes on n nodes: it’s too much!

Just consider graphs with the same number of edges?

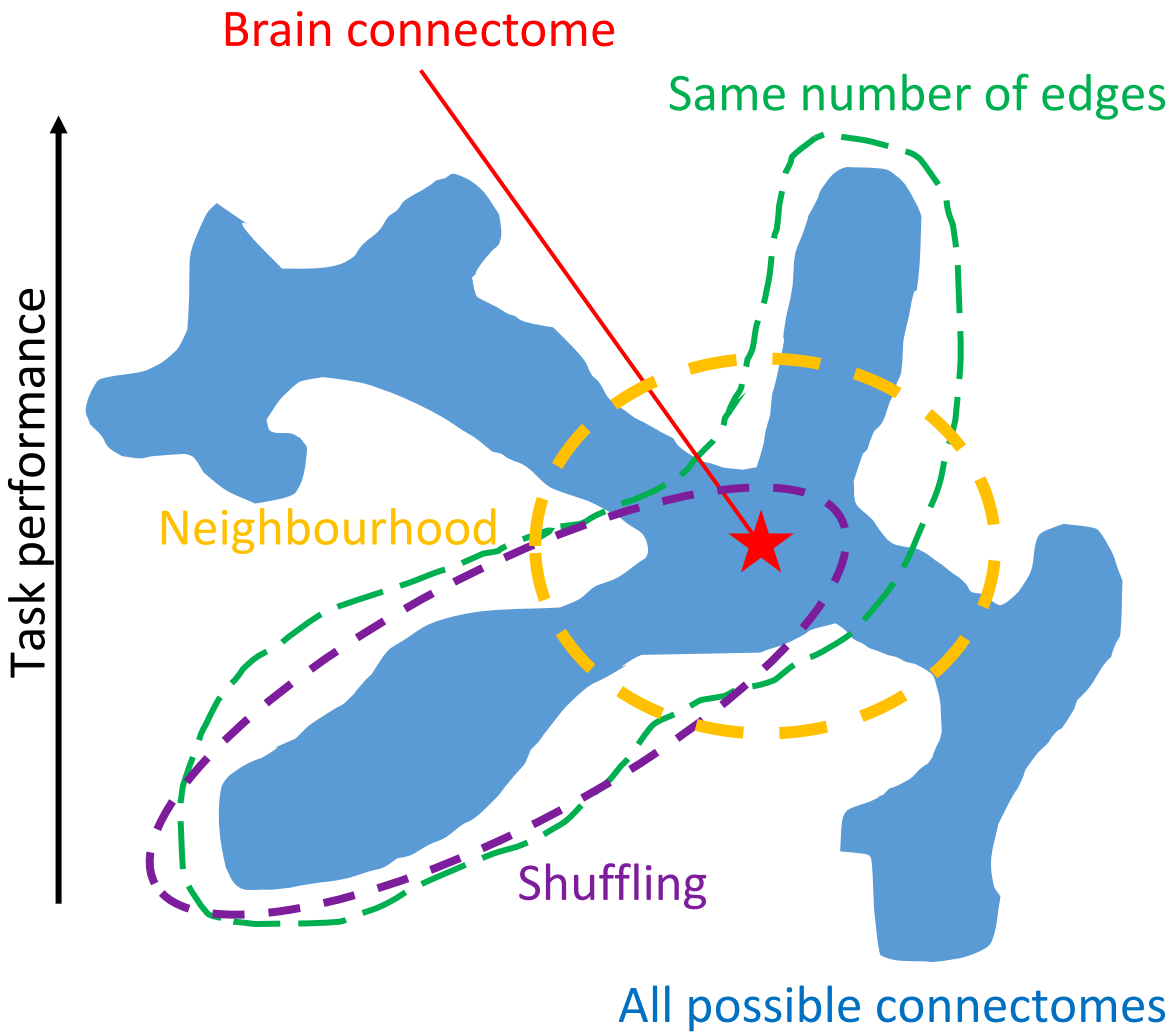
Task performance is middling, but maybe this isn’t the right comparison because not all those graphs are geometrically feasible?

How about graphs you get from shuffling?

This feels right and now we get the result that brain connectome is optimal. Publish it! 🎯

But what about the paths we didn’t take?

e.g. “all connectomes in a neighbourhood”



Possible solution: "Multiverse analysis" 🤔

Basic idea: test if result is true across all possible paths through the garden

O'Leary et al. (2015) suggest searching parameter spaces and looking for "sloppy" and "stiff" directions.

Perspectives on Psychological Science
Volume 11, Issue 5, September 2016, Pages 702-712
© The Author(s) 2016, Article Reuse Guidelines
https://doi.org/10.1177/1745691616658637

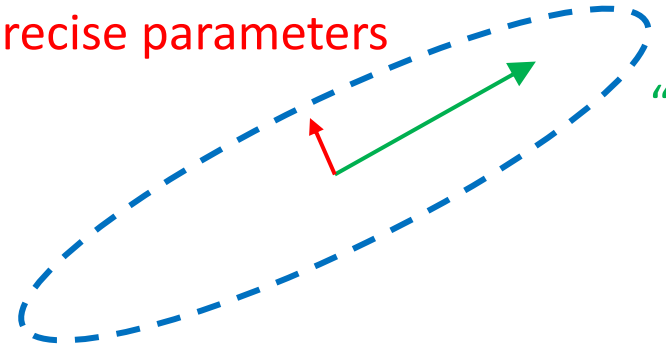
SAGE journals

Special Section on Improving Research Practices

Increasing Transparency Through a Multiverse Analysis

Sara Steegen[1], Francis Tuerlinckx[1], Andrew Gelman[2], and Wolf Vanpaemel[1]

"Stiff": result depends on precise parameters



"Sloppy": Result doesn't change much

Boundary of region in parameter space where result is true

Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Neurobiology
ELSEVIER

Computational models in the age of large datasets

Timothy O'Leary, Alexander C Sutton and Eve Marder

CrossMark

But large NeuroAI models limit how much of this is feasible.

Practicalities

There’s no perfect solution to this problem

🤠 **With your reader’s hat on:**

Prefer papers that suggest they’ve at least tried to address it

Red flags: only presenting one set of parameters and not talking about why

Green flags: several alternative choices, parameter sweeps, etc.

🤠 **With your researcher’s hat on:**

Think about alternative choices you could have made

Follow some of those paths, do parameter sweeps where feasible

Show the results for the unsuccessful paths in your paper

Consider using simpler models so you can go down all the paths

Or… consider making different kinds of claims

⚠️ **Warning** ⚠️
I don’t always do this even though I know I should.
It’s hard! 🥳

What kinds of claims can we make then?

So far we've talked about **normative models**: "brain/mechanism X is optimal for Y"

There are other ideas about what models are for:

Compressing and predicting data. Useful, but I'm not going to discuss these.

Mechanistic / bridge between levels. How an observation arises given components.

Two possible claims: this IS the connection, or this COULD BE the connection.

- "IS" : problem of multiple realisability (will come back to this).
- "COULD BE" : I would reframe this as…

Hypothesis generation. Create conceptual understanding that can lead to new work.

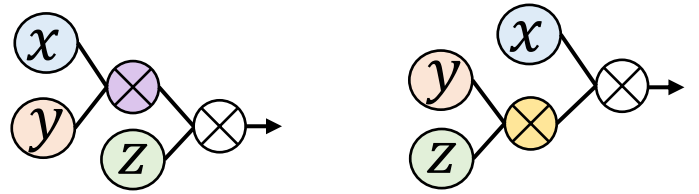
Can lead to both theory and experimental work.

Explicitly not an end point, but a necessary stopping off point on the way.

I claim: for neuroscience at the moment, this is core.

From multiple realisability to resource efficiency

Different mechanisms can give rise to the same outcome



Trivial example: $(x \times y) \times z = x \times (y \times z)$

More interesting: Eve Marder's work on CPGs 🦞 🦀

Claim: multiple realisability is a particularly hard problem for neuroscience

All network types that we study are **universal function approximators**

So knowing that you can implement function X with network Y doesn't tell you much

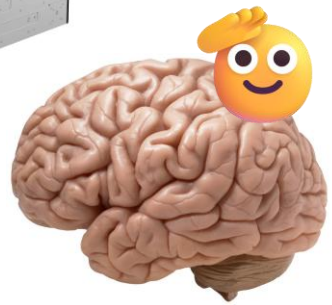
However: universal approximation often requires heavy resources (large N limit)

Adding a resource constraint makes the question interesting again



Important for neuroscience: the brain is heavily constrained

Compare: ~8W for the brain versus NVIDIA B300 GPU at 1000W+



So new question: how can we model brain given its resource constraints?

Sidenote: resource efficiency may *define* intelligence

"*Intelligence is the ability for an information processing system to adapt to its environment* ***with insufficient knowledge and resources***" (Pei Wang 1995)

Modern update from training LLMs: knowledge *is* a resource

From the point of view of an organism

Unlimited resources: no need to interact with environment, no need for intelligence.

Unlimited knowledge and computational power: easy to guarantee optimal choices.

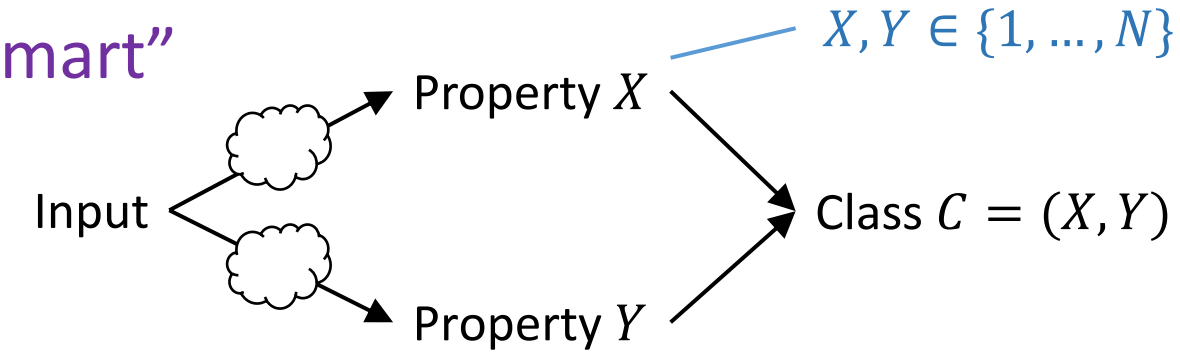
$$\text{selected choice} = \underset{\text{possible choices } c}{\text{argmax}} \ \text{reward given } c$$

From resource efficiency to modularity

"Dumb"



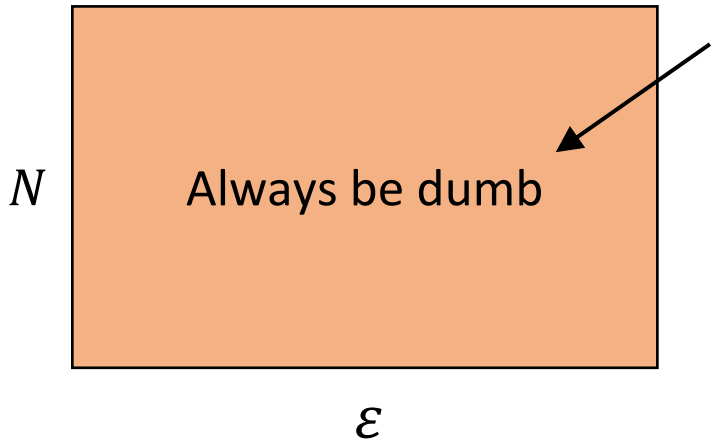
"Smart"



Resource cost cN^2 per trial
Needs to see every example (X, Y)
Error cost 0

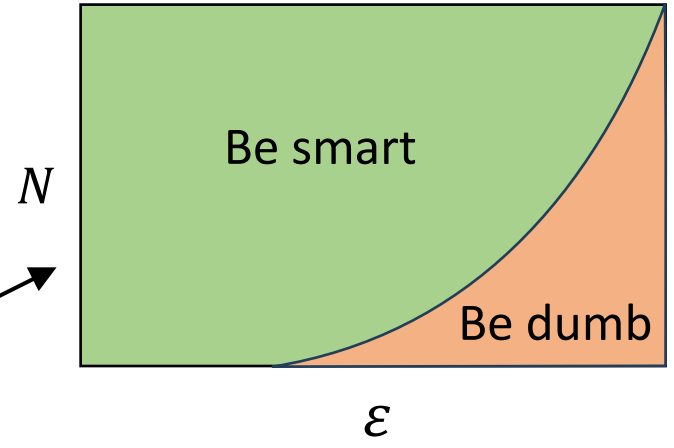
Resource cost cN per trial
Needs to see every X and every Y (generalisation)
Error cost ε per trial

If resource cost $c = 0$



Brains:
Sometimes smart,
sometimes dumb.

If resource cost $c > 0$



TL;DR version of the modularity project

Evidence:

Brain is “structurally modular”

Knocking out brain regions seems to lead to functional deficits

So…?

Functional specialisations a necessary consequence of structural modules?

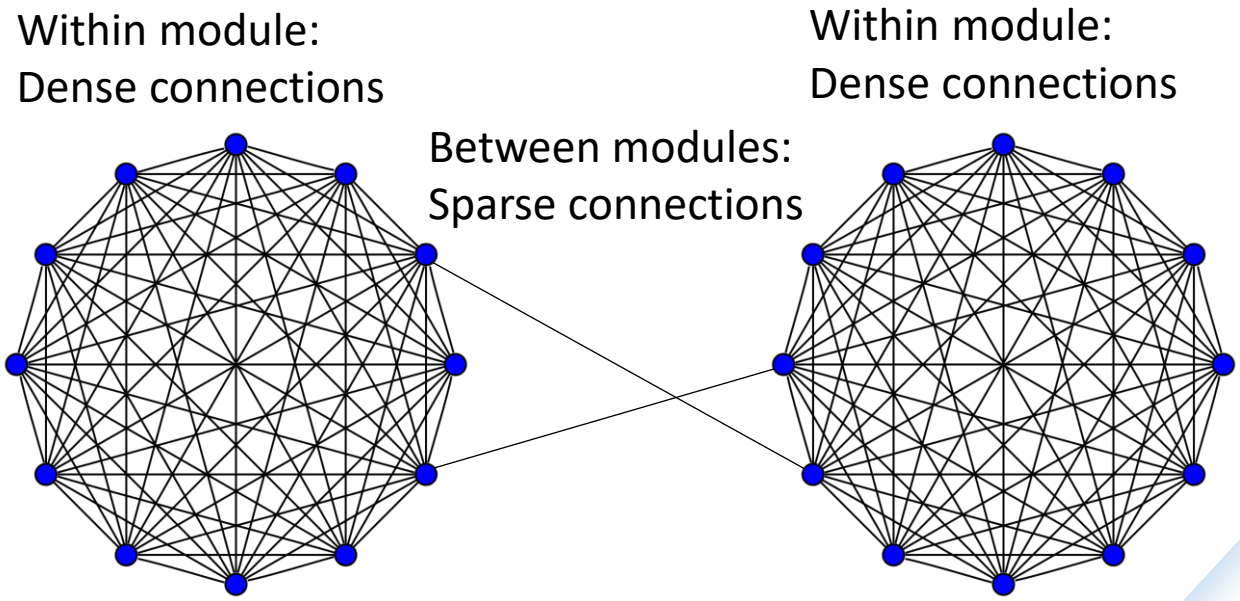
Our hypothesis: yes!

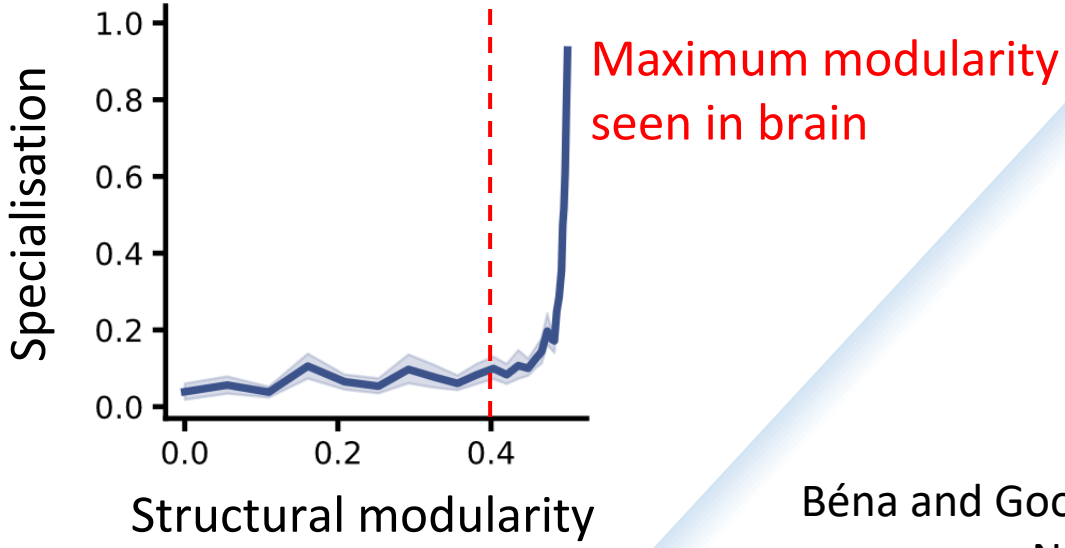
We tested this using ANNs

Answer is not necessarily.

At least, for levels of modularity in brain.

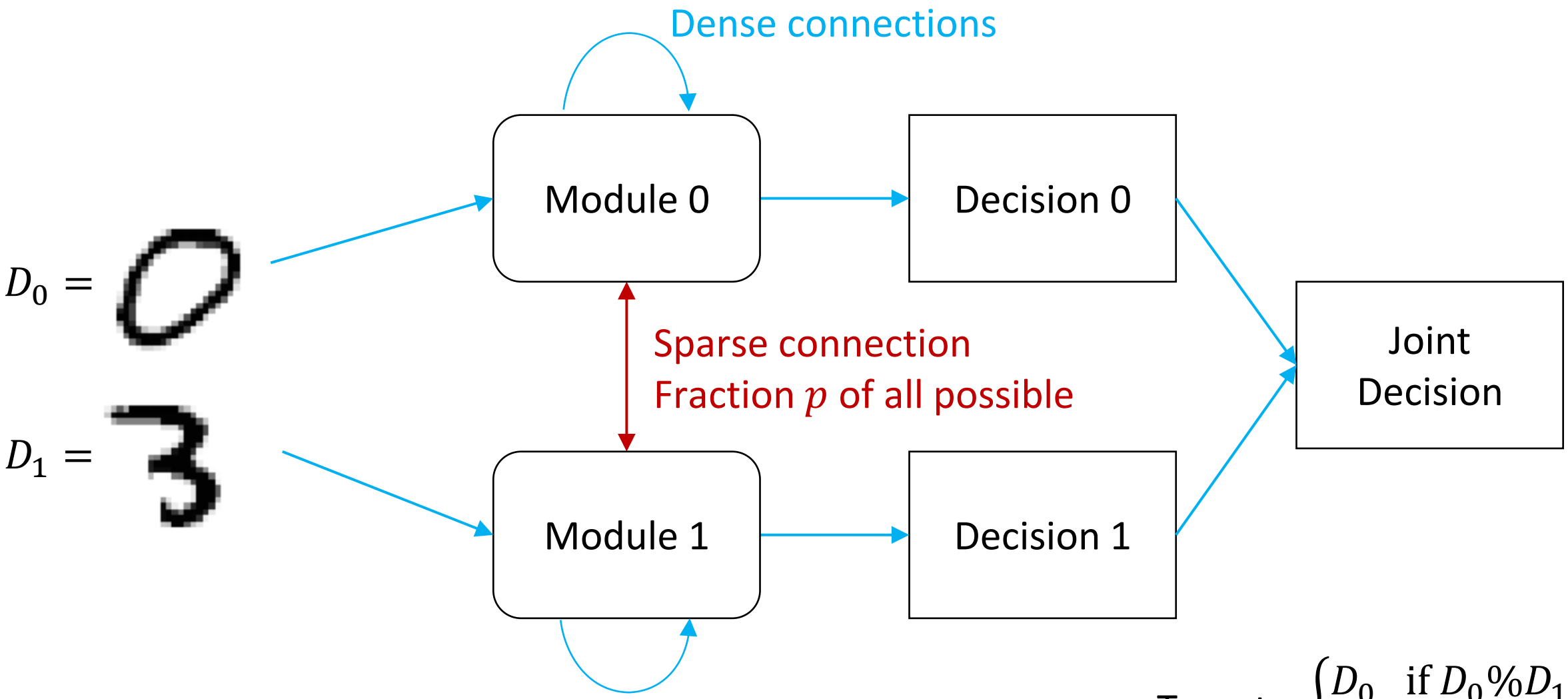
Need additional mechanisms





Béna and Goodman (2025)
Nature Comms

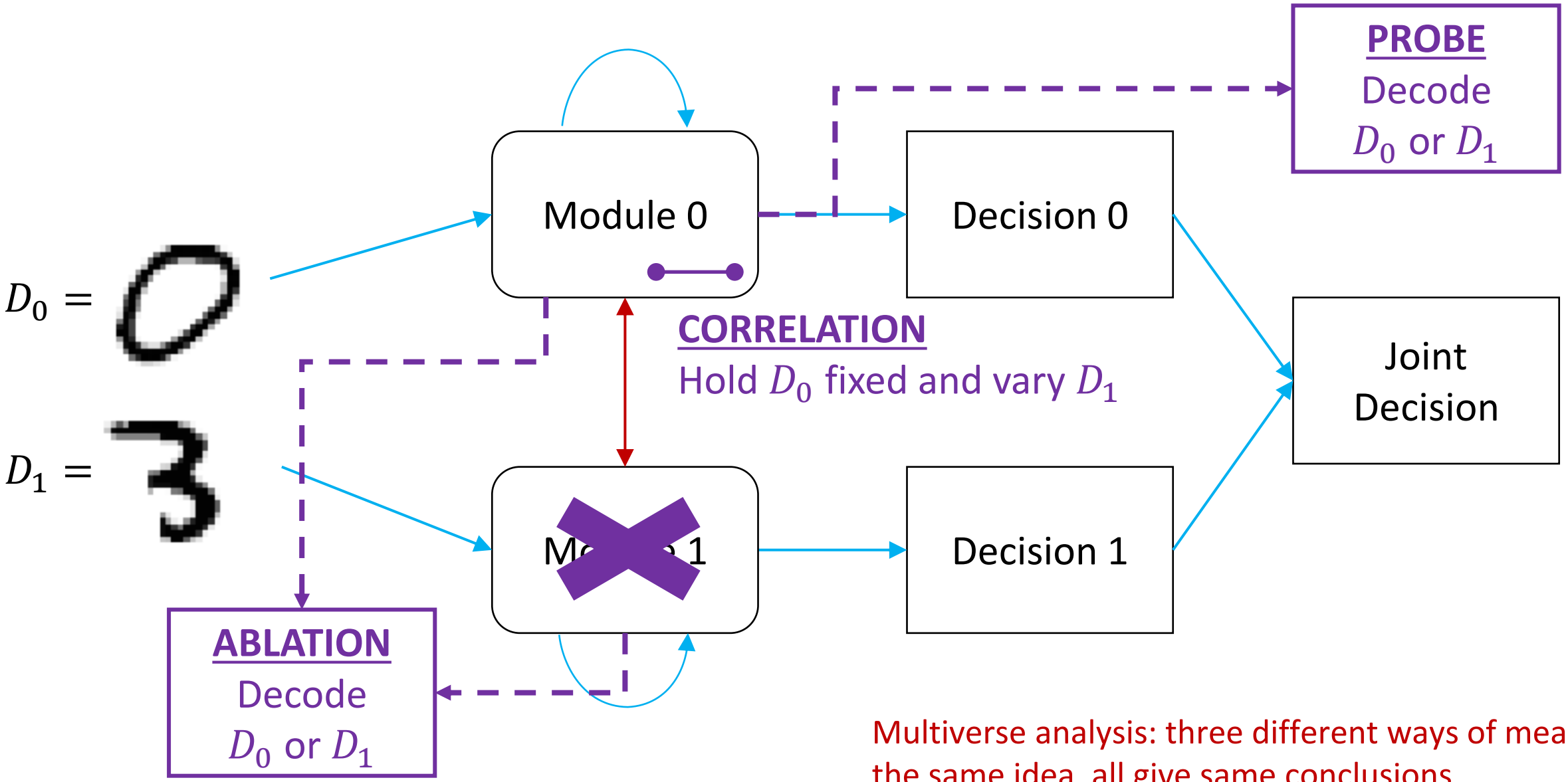
Task and model setup



$$\text{Target} = \begin{cases} D_0 & \text{if } D_0 \% D_1 = 0 \\ D_1 & \text{if } D_0 \% D_1 = 1 \end{cases}$$

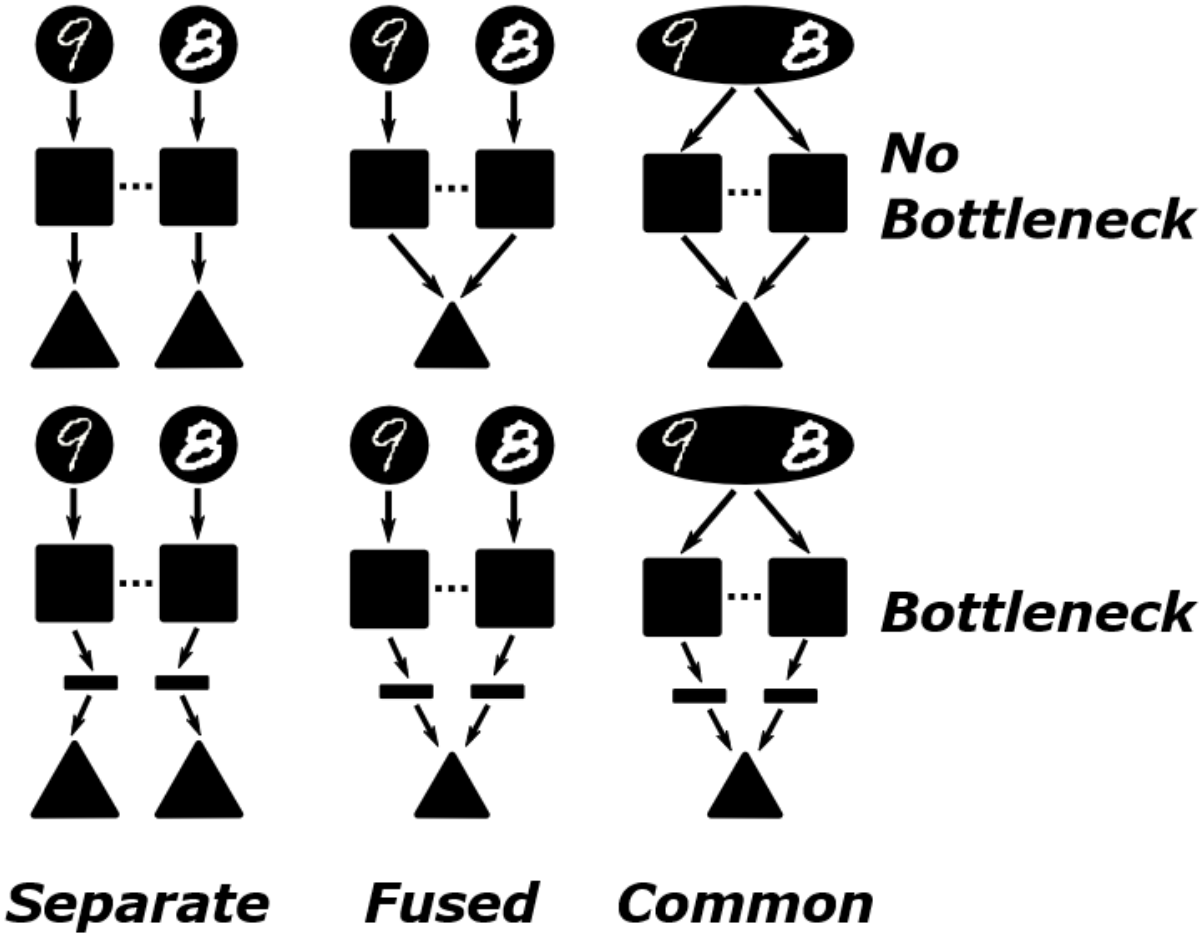
Notes from the garden: we took other paths too, will come back to this

Measuring functional specialisation



Multiverse analysis: three different ways of measuring the same idea, all give same conclusions

Model variants



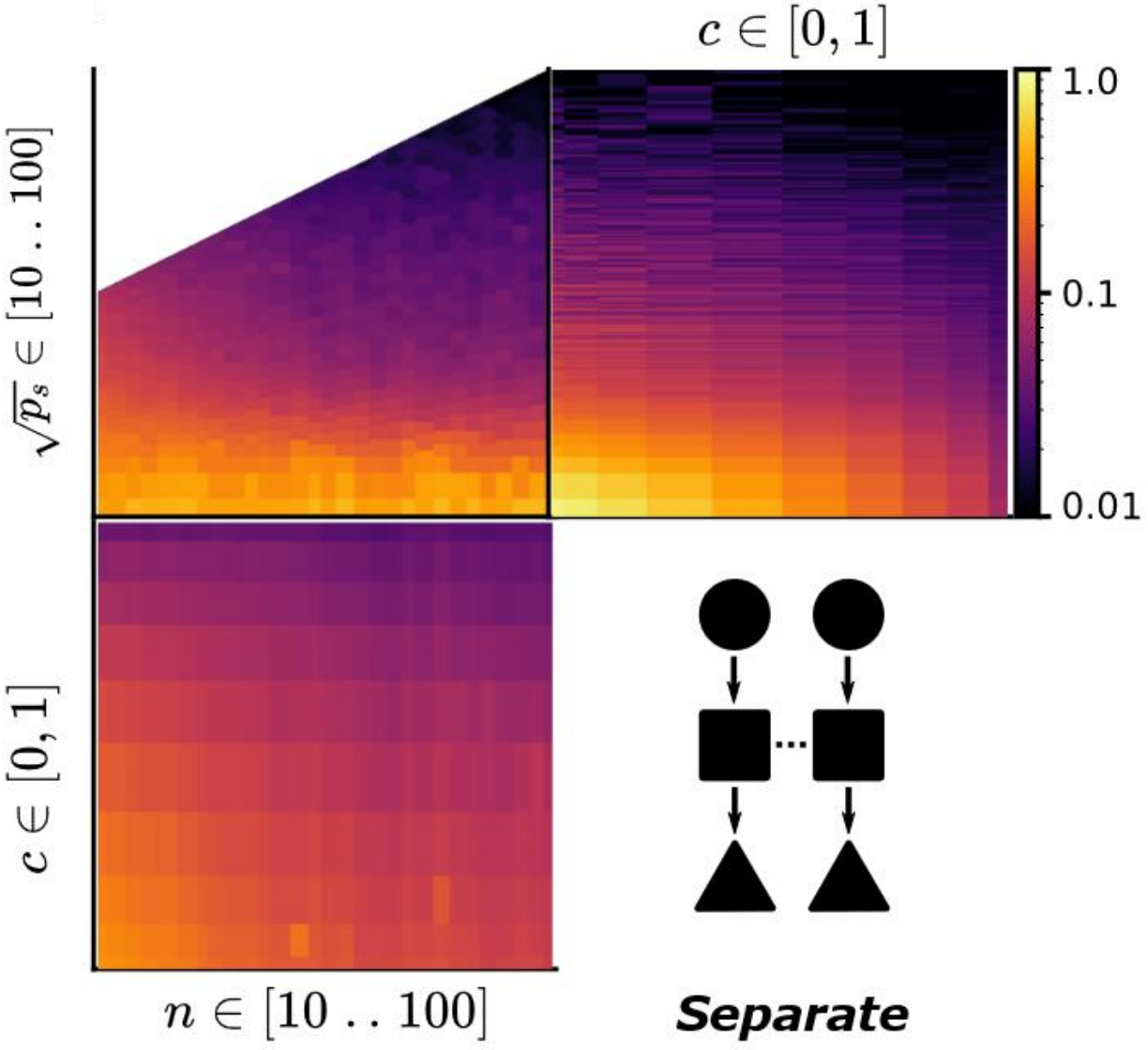
6 variants on this model

Separate pathways
Fused output pathway
Common input and output

In each case, with and without bottleneck

Gardening: These are some of the other forks we took.
Multiverse analysis: all give more or less same results.
Showing unsuccessful: some give less strong results

Detailed results



Varied parameters:

- c – correlation between input digits
- n – number of neurons in module
- p_s - number of inter-module synapses

Summary:

- Modularity can only emerge when task statistics are separable.
- When resources (n, p_s) are low, modularity more likely.
- Depends on architecture, but similar pattern for the others.

Parameter sweeps! There are many more in the paper

Conclusions

The observed structural modularity in the brain does not guarantee functional modularity.

(Note from the philosophy section: we're not making a strong claim about how the brain is)

However, that doesn't mean they must be unrelated! We know they are.

But, we need more than just the structure to get that.

Resource constraints is one possibility.

But there may be others:

Learning rules.

Environments (e.g. multi-task learning).

Training regimes.

Architectures.

WIP: A spatially distributed blood / energy supply may encourage functional modularity

All of our measures of specialisation can be experimentally measured.

Thanks to…



Béna and Goodman (2025)
Nature Comms

And thank you for listening

👉

Gabriel Béna
Modularity
PhD student 2021-2026
Now postdoc in Zurich

Nicolas Dundov
Blood and modularity



My current research group for working on all these ideas with me ❤️