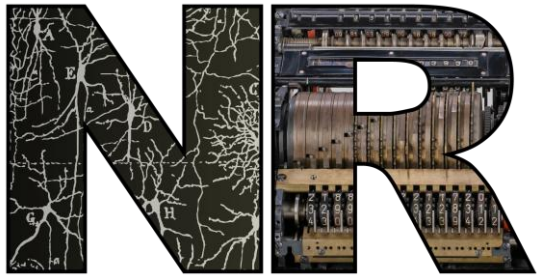
Untitled NeuroAI talk…

Dan Goodman

Imperial College London

Department of Electrical and Electronic Engineering

IMPERIAL



neural-reckoning.org

Starting with an apology

Dan… why doesn't your talk have a title? 🤔

Specifically, why doesn't it have the title listed? 🤔 🤔

Well…

I was supposed to give an introduction and history of NeuroAI, but…

I'm not sure I know what NeuroAI actually means.

Let's try: NeuroAI =

Computational neuroscience with bigger models and trained with backpropagation?

This might be (A) a useful demystifying perspective, or (B) embittered snark

Assuming it's roughly correct then there are:

Opportunities, **challenges** and **potential pitfalls**

Challenges

Challenges are not unique to NeuroAI, but some of them are particularly acute

Key challenge:

How do we know that a model has improved our understanding?

Key risk (I claim):

NeuroAI is particularly susceptible to the garden of forking paths

The garden of forking paths: a digression

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- Paper: Gelman and Loken (2013, unpublished)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?

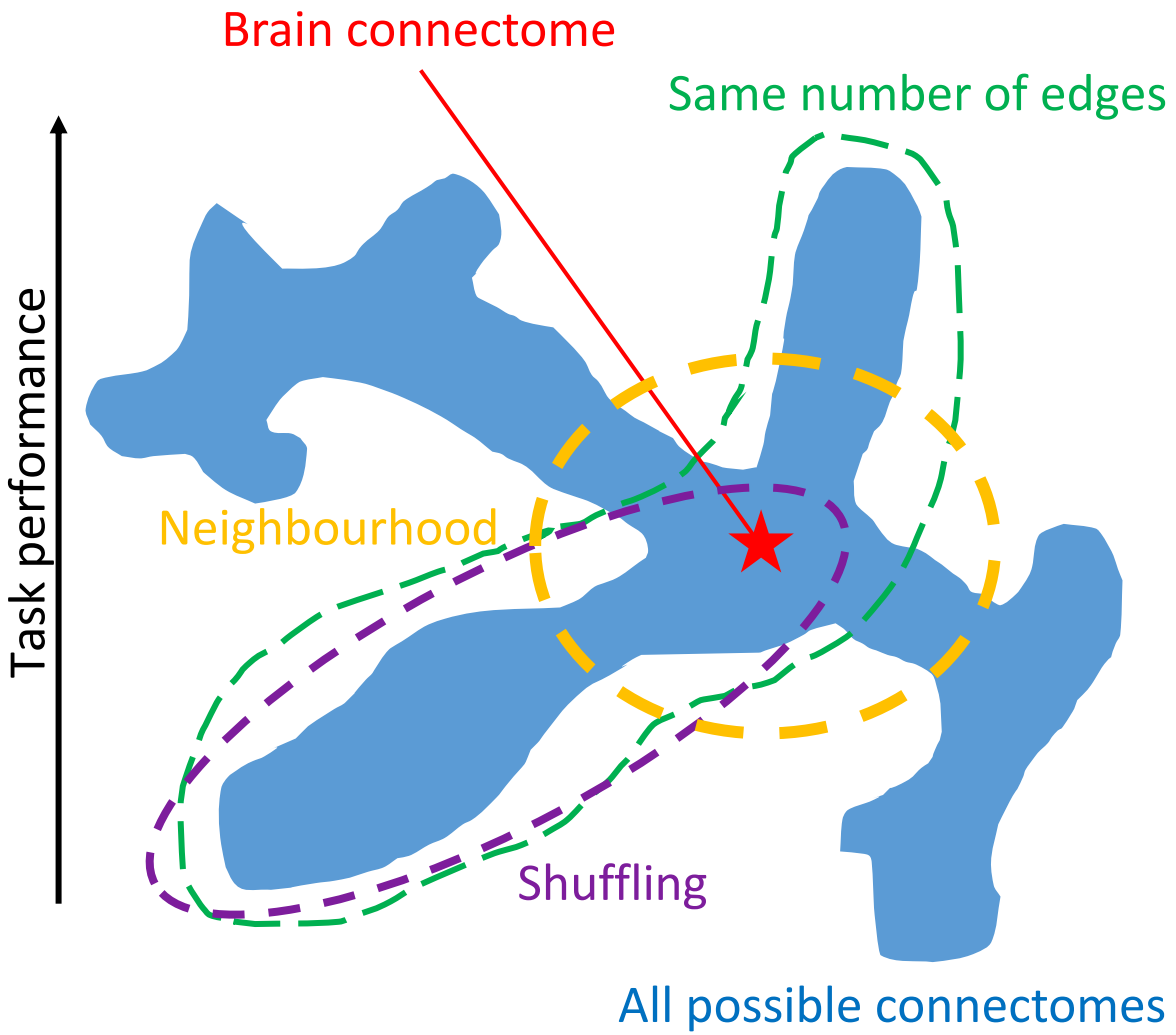
Task performance is middling, but maybe this isn’t the right comparison because not all those graphs are geometrically feasible?

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

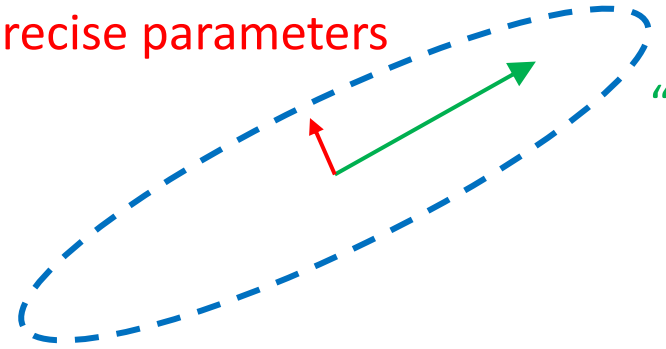
SAGE journals

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.

My take / advice / approach

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

(Yes this is an advert for Marcus' talk later on "Toy models")

Or… consider making different kinds of claims (remainder of talk = example of this)

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

Opportunities, optimism and advertising my work

So far I've focussed on the challenges and pitfalls, but **NeuroAI is a huge opportunity** too

If we want to understand the brain, we have to study how it does hard tasks

With small models or no backprop algorithms, this was very difficult

Now is the best possible time to be doing modelling in neuroscience

My approach:

- Deep exploration of models (follow the paths)
- Spend as much time trying to prove myself wrong as right (at least this is my ambition)

My aim:

Enrich our understanding of how various mechanisms can contribute to function

Examples in current work: neuromodulation, energy supply, heterogeneity

Today I'll just talk about **modularity**

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?

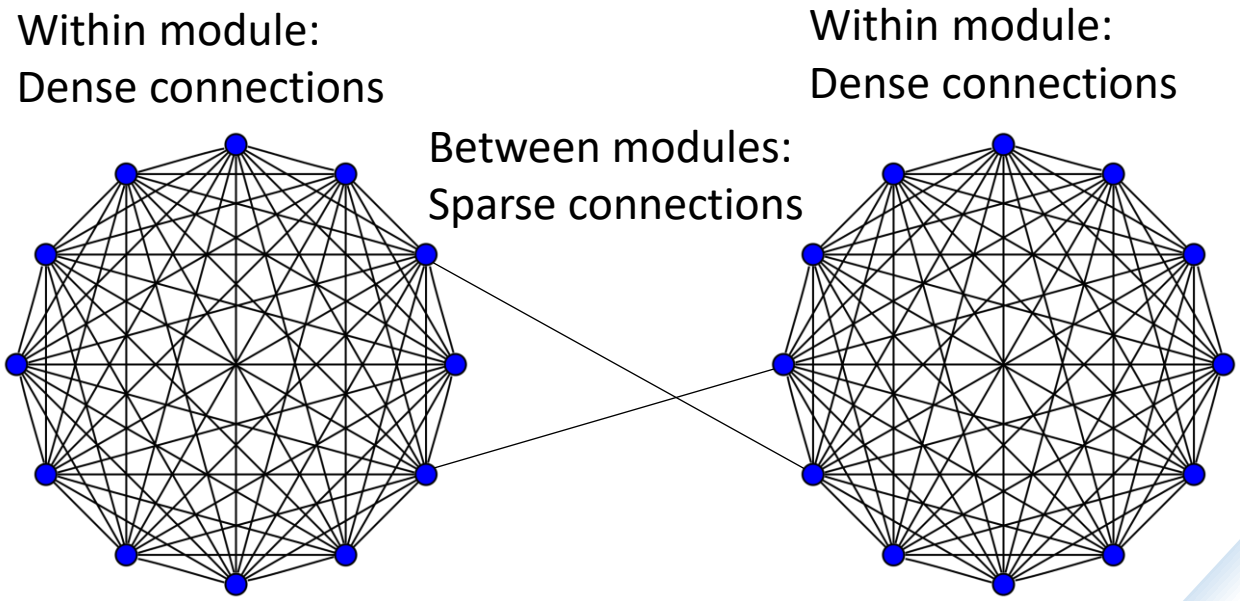
Our hypothesis: yes!

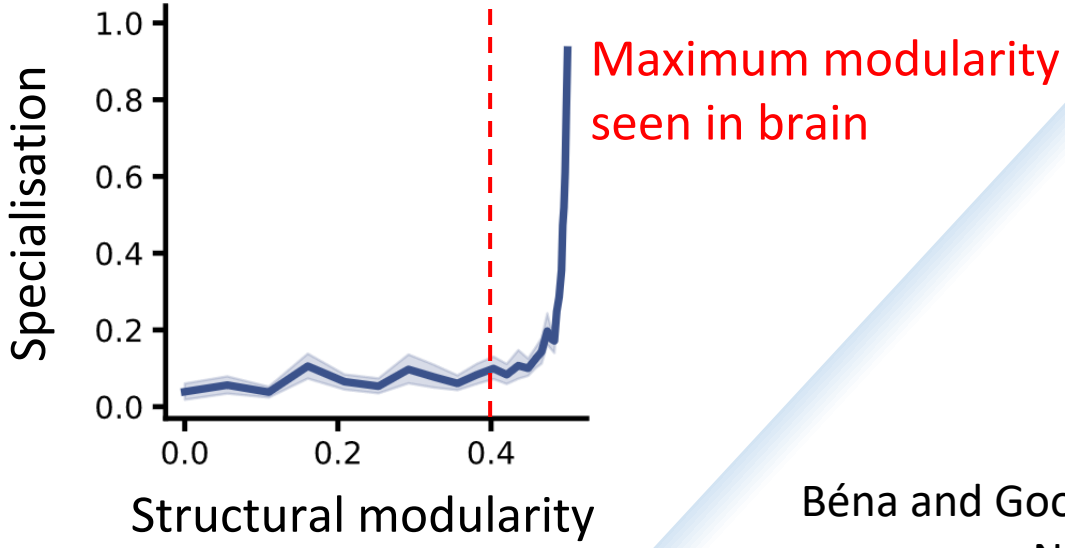
We tested this using ANNs

Answer is not necessarily.

At least, for levels of modularity in brain.

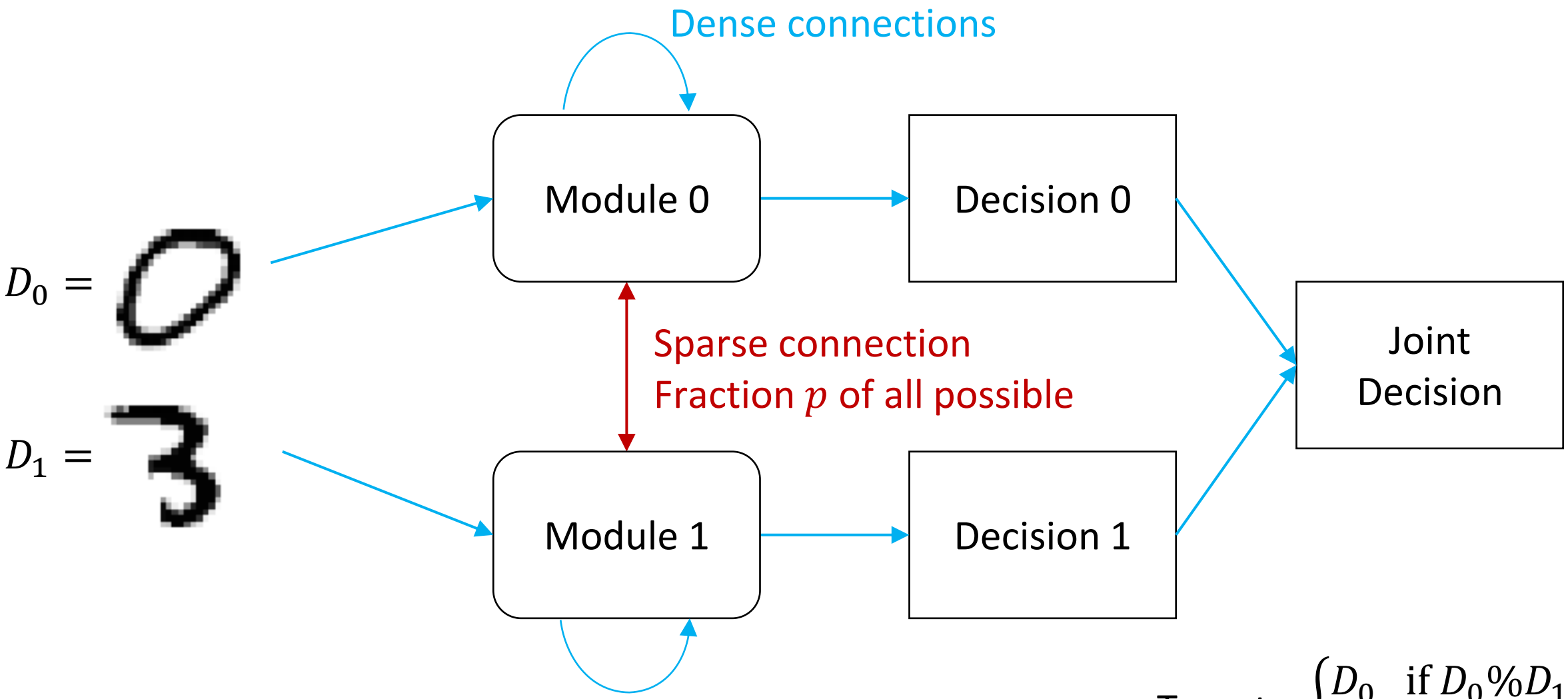
Need additional mechanisms





Béna and Goodman (2025)
Nature Comms

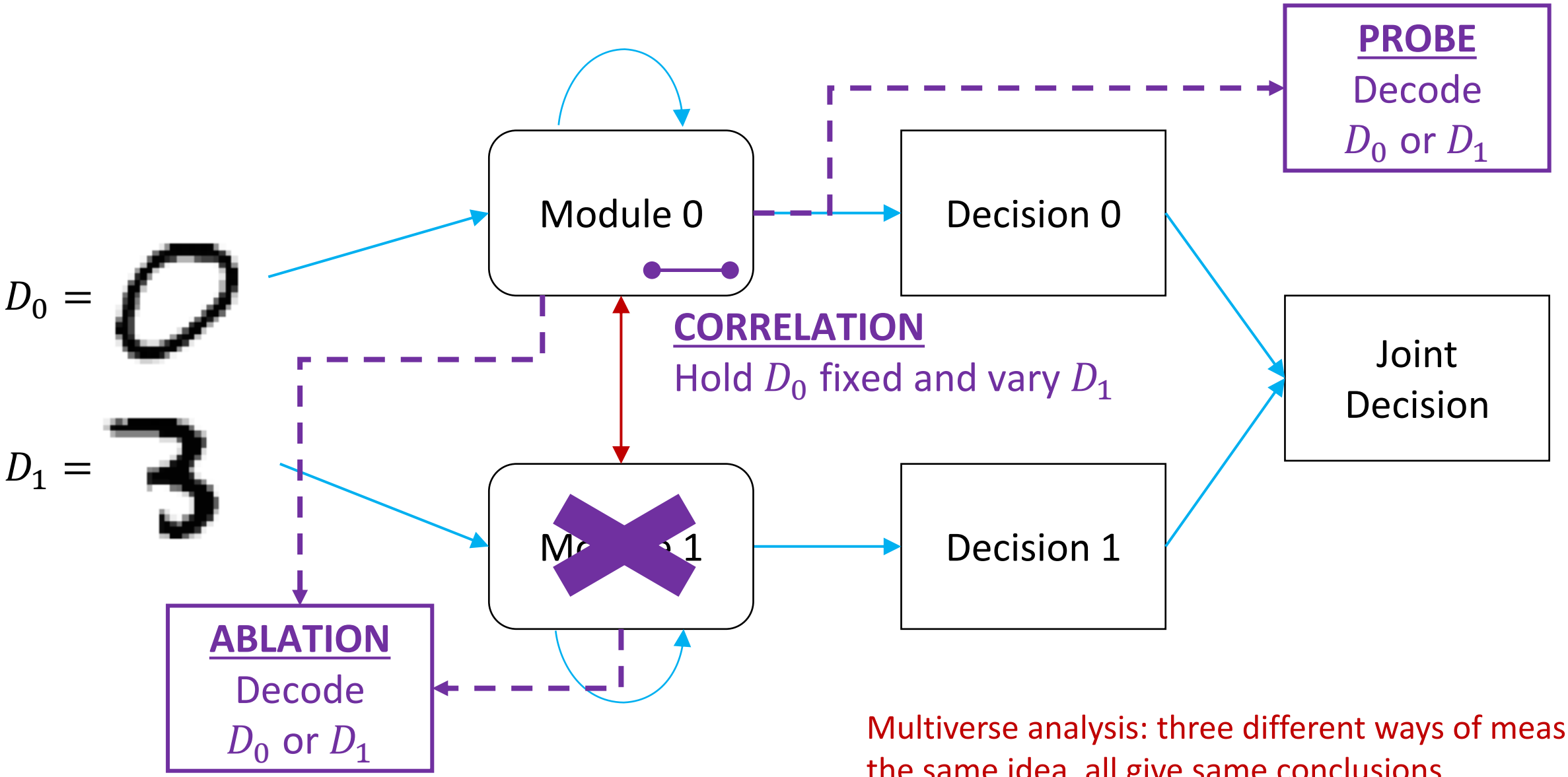
Task and model setup



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

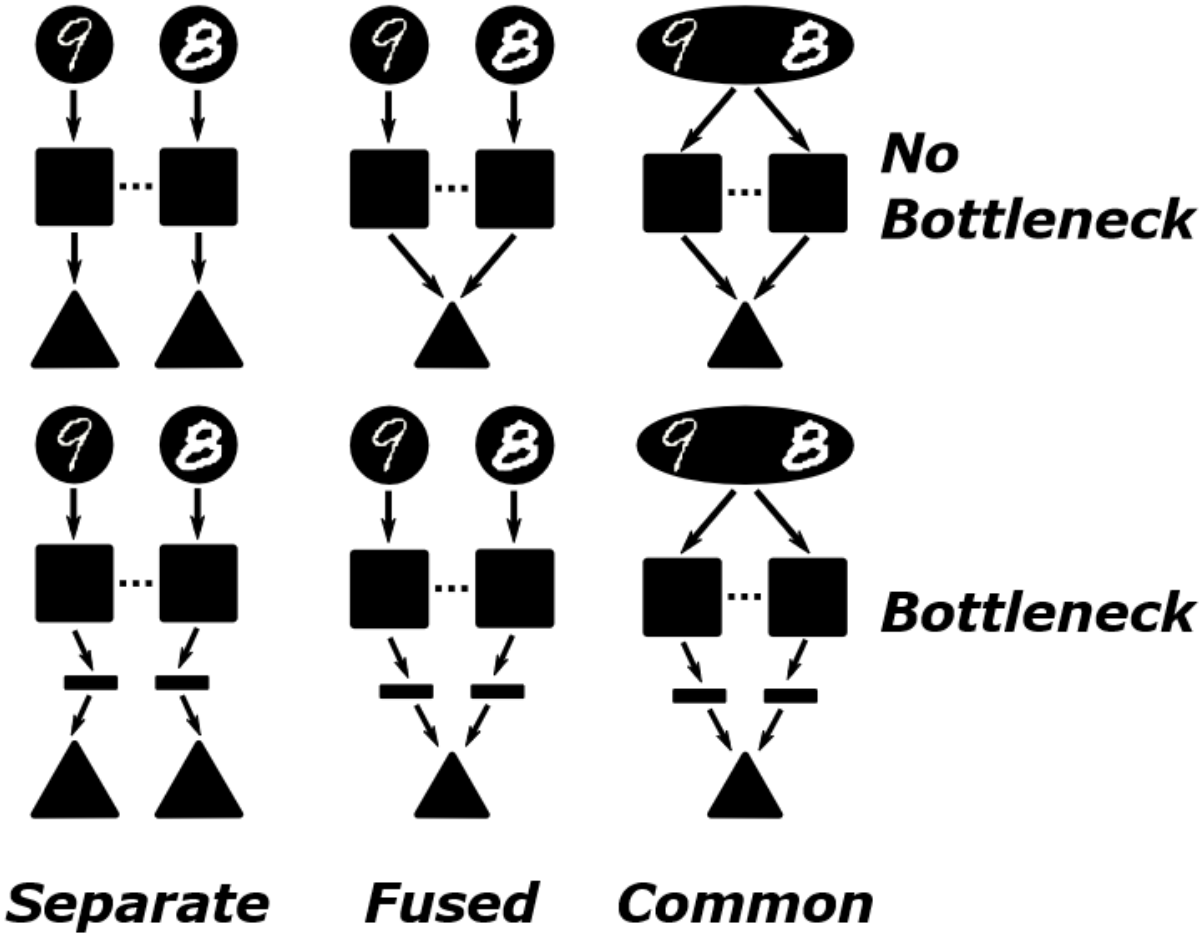
Notes from the garden: we took other paths too, will come back to this

Measuring functional specialisation



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

Model variants



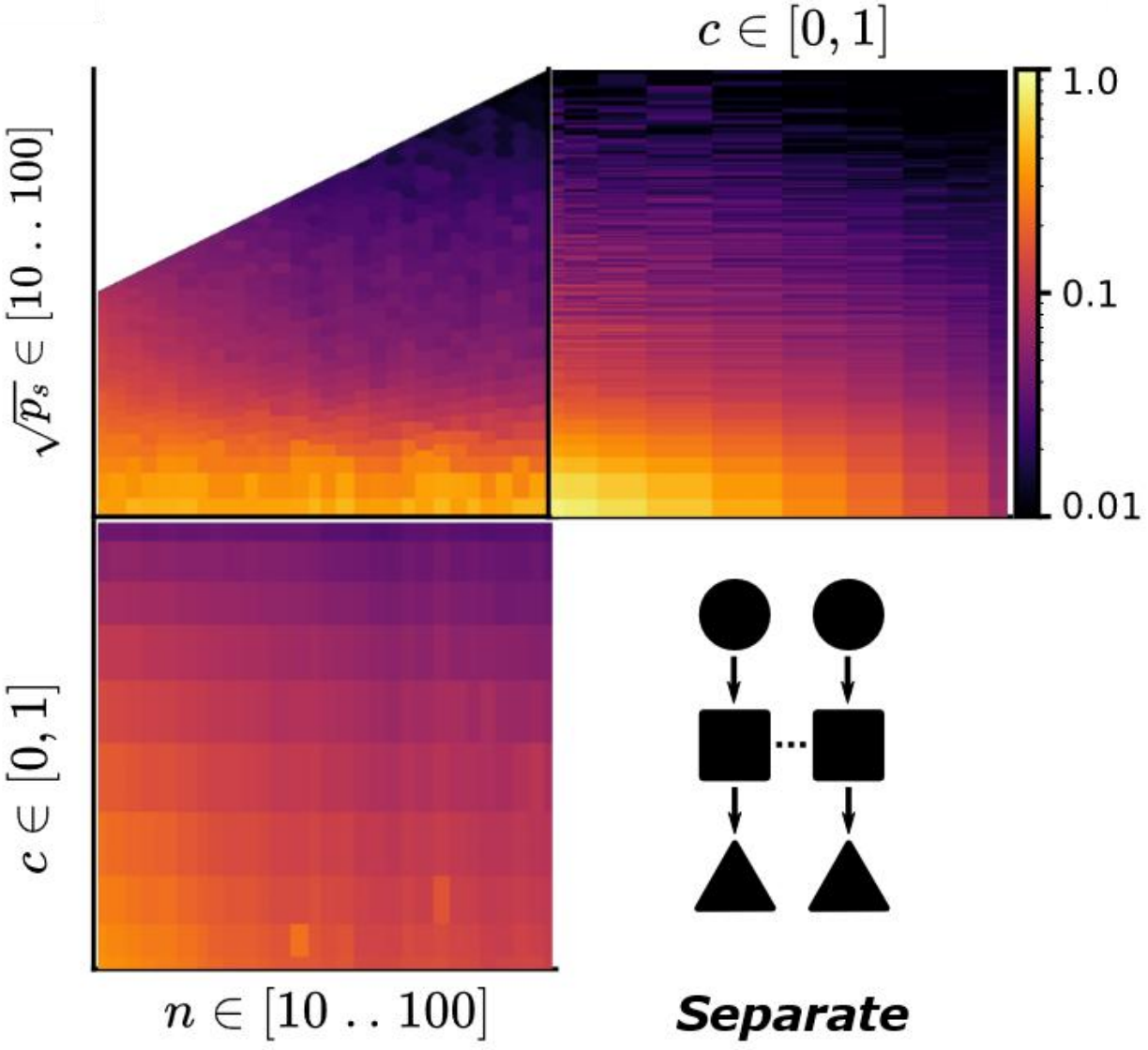
6 variants on this model

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 advice page: 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…

Marcus Ghosh

Toy models and structure-function

Postdoc fellow 2021-2025

Now started his own research group at Imperial

Gabriel Béna

Modularity

PhD student 2021-2026

Now postdoc in Zurich

Béna and Goodman (2025) Nature Comms

And thank you for listening

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