PSYC 365 Class 13:
Sustaining attention

Mid-semester check in: Mental health

- You don't have to fail the course because your mental health is suffering!
- Email us! We'll work with you to find a strategy to get you through
- Check out wellness resources
- links to Psych Dept page with links to support resources on Canvas home page

The next 2 weeks

- Today: Sustaining Attention
  - Reading: Rosenberg et al.
- Tuesday, Mar. 4th: Emotion, Motivation and Attention
  - Reading: Inman et al., 2023 (Review paper)
- DEADLINE to submit MCQ for MT2 at 11:59pm on Tuesday
- Thursday Mar 6th: Questions and review
- **Tuesday Mar 11th: Midterm 2**
- **Tuesday, Mar. 18th: Fill out presentation group preferences on Canvas**

Reminder: How to navigate this course

- **Passingham text and some lectures**: Give background/context to the research we will look at more closely
- **Experimental papers**: zoom in on primary experimental material presenting important findings central to the course topics
  - **Goal**: Learn how to extract meaningful information from and critique scientific papers presenting key findings on the topics
  - **Tips and lectures:** Guide you to important information to focus on – separate out important info
  - **Lectures on the papers: My goal is to de-mystify them!** Cover the main information pointed to in the tips. Translate the key information in the papers out of jargon into more accessible language. This is modeling how to do it so you can learn to do it to!

Top-down processes: What influences what we see of the world?

- **Class 12:**
- **Prediction**: Egner et al. paper on predictive coding
- **Attention**: Background information on the cognitive neuroscience of attention covered in Passingham chapter 3
  - Covert/Overt attention
  - Selective attention
  - Biased competition
  - DAN & VAN
  - Sustained attention
- Experience of living with ADHD

Learning Objective

- Confidently describe recent research defining a neural marker of *sustained attention* and its relationship to ADHD

Reading question

According to Rosenberg et al., which specific attentional functions make up the process or *construct* of sustained attention?

a. Information selection

b. Divided attention

c. Inhibition of unselected information

d. a & c

Show of hands

a. Arrgggh. That nearly killed me!
b. Based on your tips I focused on the intro and discussion and key variables and it was challenging but manageable
c. Super clear. Sailed through. Whooo!
d. I haven't read it.
e. yep

Big Picture Problem: No *summary index* of attention

- Attention is key for perception and all kinds of cognition
- But different types of attentional process are measured in too many different ways
- So none of them will work as a measure of individual attentional ability!

In this case the big picture question involved an overall problem in the field. There is no summary index of attention. What is a summary index, you ask?. A single measure or number that *sums* up one person's general ability. In the Introduction, the authors talk about how we have a measure of overall intelligence, and a way to measure individual differences in working memory, but even though attention is *pervasive* – it underlies nearly every task in daily life -- there is no good measure of individual differences in attention. Yet [1] …But the problem is that [2]… and [3] . In this paper, as a summary index, they decided to focus on sustained attention because it plays a role in ADHD. And just to remind you sustained attention is keeping your mind on a task – often boring – for long periods of time. It involves selecting what to attend, and suppressing distracting information, and then refreshing that process repeatedly as your mind wants to drift.

Question

Why might they want to look at functional connectivity across the WHOLE brain and not in, say, the dorsal attention network?

a. Every cognitive function involves the whole brain

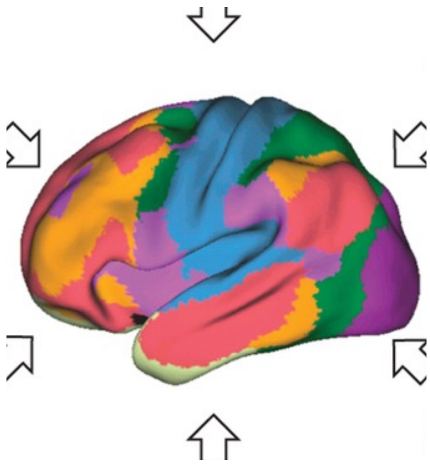
b. Sustained attention involves a variety of functions

c. Intrinsic connectivity data measures are very noisy so you need info from the whole brain

d. Attention research is very fragmented so we know nothing about underlying brain systems

(talk to your neighbor) b. And that variety of functions involves a wide range of regions, cortical, subcortical, and including the cerebellum

Goal: Identify a neuromarker of sustained attention based on whole brain *intrinsic connectivity*



Here they wanted to identify a neuromarker --- a pattern of brain activity that reflects each persons sustained attention ability -- by measuring functional connectivity within and also across canonical networks -- across the whole brain. Remember we talked about intrinsic networks when we discussed Poldrack? As a reminder, it's correlated activity observed with fMRI – slow BOLD fluctuations between regions that go up and down in unison when you're lying in the scanner and you're not focusing on completing a specific task. But you can also look at these patterns of intrinsic functional connectivity while people are doing a task. Here in search of their neural marker of attention they want to look at correlated activity not just in individual canonical networks across ALL of them. So on the left is an image of different canonical intrinsic networks on the outside of the circle and in the centre is an image of all of them together, added to one brain image like puzzle pieces [1]. Each network os in a different colour. **Q: Why do Rosenberg et al. say that a measure of attention based on resting state data would be a really useful thing?** Resting-state data is relatively straight-forward to collect and share across different locations and across language and cultural barriers.

Rosenberg: 3 assumptions

1. Individual differences in sustained attention will be reflected in complex patterns of correlated BOLD activity across brain regions
2. These patterns will be observed both when you're doing a task and at rest
3. If we can distill a signature of these patterns we should be able to use them to predict attention ability in a completely different set of people

The authors describe that their approach is based on 3 assumptions.[1] And will not even be confined to a single network. What is the value of looking across many networks? Passingham says: To put it bluntly, no behavioral task depends on the activity of one brain region in isolation, and one of the strengths of brain imaging is you can measure activity across the whole brain. It's about the degree to which a whole bunch of regions are talking or not talking to a whole bunch of other regions But this study takes it a step further, as often a cognitive process depends on more than just one network. One pretty obvious hypothesis you might have is that better sustained attention would be reflected in stronger connectivity within the DAN. But maybe by focusing on the DAN you'd miss important information. Here they wanted to look at patterns of activity beyond what was restricted to this network. So beyond looking at the whole brain they said they assumed [2]. This is true of canonical networks and they thought it would be true of their whole-brain neuromarker as well. [3]... The challenge was to get a handle on those complex patterns. Which I think they did quite successfully in this study. This paper has been cited 917 times since 2016. The method has been successfully adopted for other studies, has gotten attention for giving us a useful approach to using brain connectivity measures as "neuromarkers" of clinical conditions.

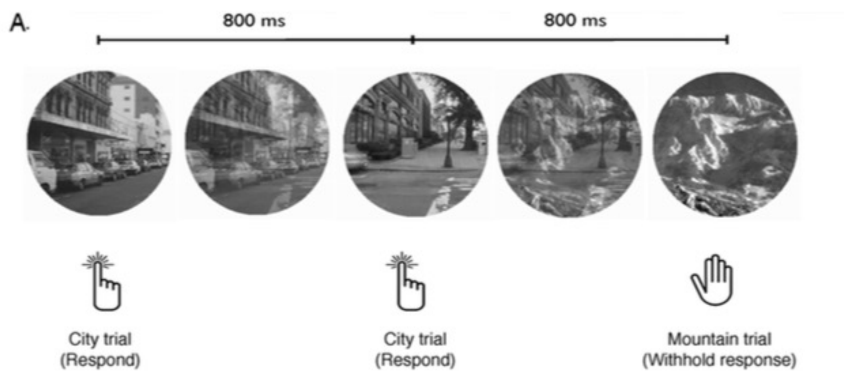
Rosenberg: Questions to be answered

- Big picture question
  - Can we find a *neuromarker,* or brain-based measure of general attentional ability?
- More specifically…
  - ISO reliable marker of overall attentional ability observable in resting brains and generalizable to new populations
- Specific question
  - Can we get take a ***data-driven*** approach to pulling out patterns of network activity to give us a marker of *sustained attention that* will generalize across populations?

27/2/25 | PSYC 365: Class 14 | 13

How can we find a consistent measure of how good you are at it? Big picture, like an IQ test for attention.[1] **Q: Why in resting brains?** Because you can measure resting state activity in populations who have different levels of ability to perform tasks in the scanner– small kids or clinical populations who can't lie still and focus for a long time – but you can still get them to lie still for 5 minutes to measure resting state data. **Q: What do they mean by generalizable to new populations?** It doesn't just tell you something about the brains of the exact people you're measuring -- it is useful for predicting what other brains will do under the same circumstances [3]... **Q: So what do they mean by a data-driven approach?.** Data driven is an exploratory method. Instead of going in with very clear hypotheses (e.g., I think THESE regions will predict attention ability and ADHD symptoms) you measure brain activity when people are paying attention and see what the data tells YOU!. And of course they chose to measure sustained attention because it's linked to ADHD.

Methods: Graduated Continuous Performance Task (gradCPT)



The experimental task they used was the graduated continuous performance task or gradCPT. What this involves is you lie in the scanner and watch images of cities and mountains as one image slowly dissolves into another. Your task is to hit a button when you identify a scene as a city scene and do NOT respond if it's a mountain scene. And 90% of the scenes are cities so it is not very often that you need to withhold the habitual action of pressing the button when the scene changes. Easy right? But you have to keep on paying attention when not much is going on – kind of like the curtain changing colour in the gorilla video. Your attention easily misses things if you don't really work to pay attention and so you make more mistakes, either pressing when you shouldn't or not pressing when you should. The number of mistakes you make is the measure of your attentional ability. Think about the following questions **Q: Why do we need a brain marker when this seems to give us a particularly good behavioural marker? Q: Why would they make 90% of trials city trials?**

Methods

- Yale participants: 25 students
  - gradCPT task
    - Performance measure: d' or sensitivity = hits–false alarms
      - How accurate is each person taking into account tendency to just hit the button
    - fMRI collected while they were performing the task
  - Resting state fMRI data
- Beijing participants: 113 kids & teenagers with ADHD and typical controls (mean age 11 years)
  - Resting state fMRI data
  - ADHD-RS scores (inattention & hyperactivity/impulsivity)

In this study there were two sets of participants, who did slightly different things. The first set was 25 students from Yale [1] who performed the grad CPT task as a measure of sustained attention. The dependent variable they used as a measure of performance was d' or sensitivity, which is basically a measure of your accuracy that takes into account your tendency to hit a button whenever you're in doubt. It was the number of hits minus the number of false alarms. Hits are when you correctly hit the button -- in this case for a city. False alarms are when you incorrectly hit the button -- in this case for a mountain. So it measures how accurate people are at hitting a button when they should taking into account their tendency to just hit the button all the time. Higher sensitivity, or d', means you are better at the task. In addition to collecting fMRI data while participants did the gradCPT task they also measured resting state fMRI data. [2]. The second set of participants was quite a different population from the first one. It was 113 kids and teenagers, mean age around 11, some of whom had ADHD diagnoses and some of whom did not. In this group they JUST collected resting state data in the scanner. And instead of gradCPT performance they measured scores on the ADHD-RS questionnaire. The ADHD-RS is a measure of the degree of ADHD using bunch of questions assessing attention. More specifically, the ADHD-RS, which is filled out by parents, has 18 questions, nine of which assess inattention, or how children attend to tasks or play activities, such as the degree to which a child “fails to give close attention to details” or is “easily distracted by extraneous stimuli.” The remaining nine assess hyperactivity and impulsivity levels, such as the degree to which a child “fidgets with hands or feet or squirms in seat” or “interrupts or intrudes on others.” Overall ADHD score is calculated as the sum of all responses. Higher scores mean more severe and/or more frequent symptoms.

Variables

Independent/Predictor

- gradCPT conditions: Go vs. Nogo (city vs mountain)
- Choice of 268 network nodes
- Participant group (neurotypical adult, neurotypical youth, ADHD youth)
- *Summary* measures of network connectivity based on BOLD time courses
  - Positive network strength
  - Negative network strength

Dependent/Outcome

- BOLD response
- Performance on gradCPT (d')
- ADHD-RS Score
- *Summary* measures of network connectivity based on BOLD time courses
  - Positive network strength
  - Negative network strength

Variables operationalize cognitive processes

- d' (performance on gradCPT)
- ADHD-RS Score
- SAN models: Built from correlations between BOLD time courses

- Capacity for sustained attention
- ADHD symptoms (associated with problems with sustained attention)
- Neuromarkers of high and low levels of sustained attention ability

So one thing to bear in mind is that the dependent variables are behaviours or brain responses we can actually measure that we are pretty sure measure the cognitive process of interest. So let's link each of the DVs in this study. d' or performance on the gradCPT task measures [1] capacity for sustained attention in the Yale group. The ADHD-RS score measures [2] ADHD symptoms that in part measure problems with sustained attention in the Beijing group. So now we have two different behavioural measures of attention ability: One is a laboratory measure measuring your actual performance in that moment. The other is a questionnaire measuring YOUR PARENT's opinion of your attentional abilities and ADHD symptoms in general. So they are quite different ways of getting at that information. Finally, the SAN models, which are created out of correlations between BOLD time courses that predict attentional ability are the [3] neuromarkers of sustained attention ability. And again, I expect most if not all of you have no idea what the SAN models are at this point, but we're going to cover that in more depth.

3-step analysis approach

- Step 1. Derive a network of regions whose connectivity strength predicts gradCPT task performance (d') in Yale group.
- Step 2. See if statistical model based on network strength while performing gradCPT can be applied to resting-state data to predict gradCPT task performance.
- Step 3. See if SAN model can be used with resting state data to predict ADHD scores in kids and adolescents in Beijing.

27/2/25 PSYC 365: Class 14 18

There are several levels to the ways they analyzed their data, as they took a 3 step approach to their data analysis. [1] Step 1. Model relationship between BOLD connectivity strength and gradCPT task performance in the Yale group. Again gradCPT gives you a *behavioural* measure of how well you sustain attention. So you want to see if that is related to patterns of functional brain connectivity while people are actually using sustained attention and not just at rest. [c2] Step 2. See if the model created in Step 1 can be used to see if you can use resting state data to predict grad CPT task performance. So even when participants are not using sustained attention to do the task, when they're just lying there in the scanner. Because networks that fire together at rest tend to fire together at work. Step 3. [3] Finally, see if the model created in Step 1 can be used with resting state data to predict ADHD scores, a different dependent variable, in a whole other population -- A different age group in a different country and culture and with and without ADHD

What is the Sustained Attention Network (SAN) model and how was it derived?

- The SAN model is a brain network based statistical model
  - You take correlated activity across a bunch of regions across the brain and reduce it to two numbers (summary statistics)
  - Those numbers reflect degree of connectivity in a brain network associated with the capacity for sustained attention
  - SAN model uses brain network scores to predict an individual's attention performance

Before we get into the hard core details I want to give you a more general idea of what the Sustained attention network or SAN model is and very generally how they got it. [1] first... [2] ....[3].... [4]... Importantly, the model based on data from one group predicts another group's performance. It GENERALIZES.. Next we're going to take a closer look at how you get those two numbers.

Questions?

- Fuzziest point?
  - Where did they get and what did they do with the nodes?
  - What's an edge?
  - What information does that huge matrix illustrate??
    - Take a moment to think/write it down
    - Turn around and talk to your neighbor about what you're most confused about so far

Any outstanding questions?

Step I

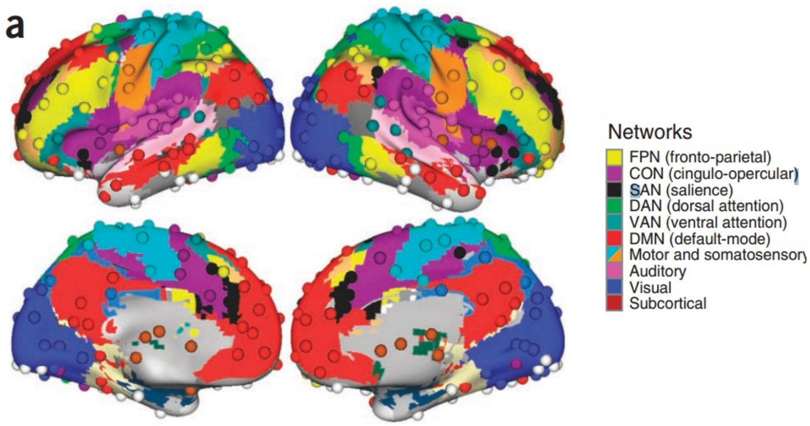
Assess what patterns of brain connectivity are relevant to attention

a. Measure correlations between pairs of nodes in each participant

b. Test which patterns of connection between regions is linked to good and bad gradCPT performance between participants

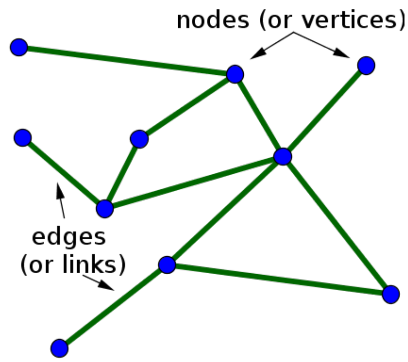
I'm now going to use a whole bunch of slides to illustrate information described in a couple of sentences in the paper, so you'll know what the sentences actually mean. They need to define the networks that predict sustatined attention. Step 1, to assess the patterns of brain connectivity were related to attention performance, was itself made up of two steps. The first was to measure correlations between pairs of nodes of brain networks, or pre-specified brain regions. But what are nodes? This requires some explanation.

Choosing Nodes (Regions) Across Canonical networks



What is a node? Each of those little dots or pimples on the brains is a node. In this study, 268 nodes were defined based on previous research. As you can see, each colour on the brain represents a different canonical network, and nodes were chosen from regions involved in all of those networks -- so the nodes are spread out over many canonical networks – they didn't restrict themselves to one or a subset but wanted to look at patterns of functional connectivity that could cut across networks. Across the entire brain. Once the they had these already-defined nodes , for each participant in the Yale group, they averaged BOLD activity across the voxels in each node and looked how it went up and down over the course of the gradCPT task and, for resting state data, also over time when participants were just resting.

Define Connections Between Nodes

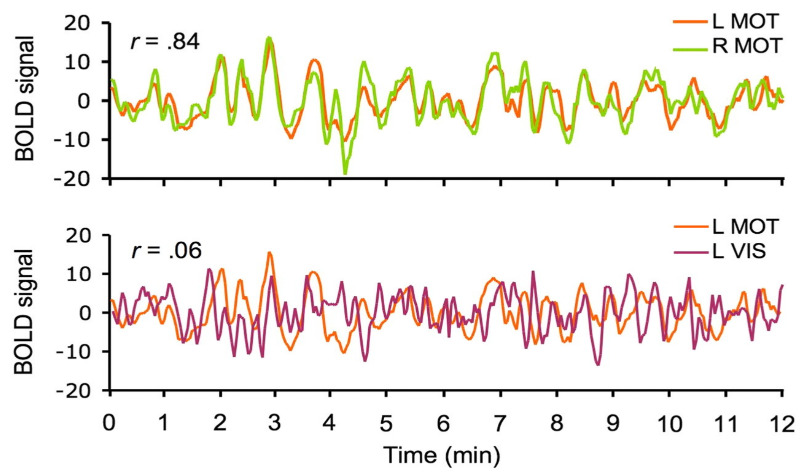


- Nodes – Brain regions where BOLD activity is measured
- Edges – connections between 2 nodes
- Correlations

27/2/25 PSYC 365: Class 14 23

How do they measure the strength of connections between regions or nodes that are what define the measure of network strenth? Nodes and edges are very basic concepts that come from graph theory, which is a technique used to measure patterns of connectivity. We've talked about what a node is. And as we just saw, the the first thing the researchers did is they extracted BOLD time series from a lot of nodes distributed across all the intrinsic networks. Once they had a measure of BOLD activation for each node as it fluctuated over time, they then calculated edges. To do that they looked at the strength of connectivity, which was the correlation between the activity over time between each region, between all of the pairs of nodes they looked at. On the left is a graphical depiction of nodes, or blue dots, and edges, or the correlated functional activity between two pairs. This correlated activity could be strong or weak.

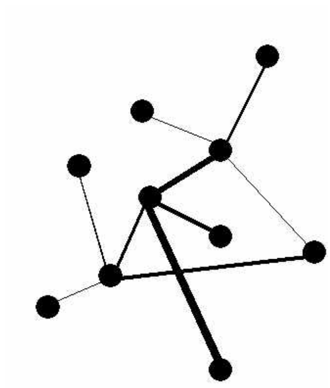
Edges: Correlated BOLD activity



27/2/25 PSYC 365: Class 14 24

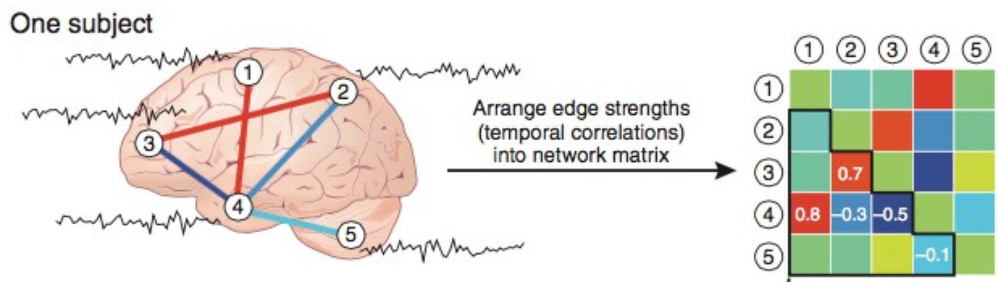
This is not from the paper but is to show you what we're actually measuring when we talk about the strength of an edge. Here is an example of what that correlated BOLD activity over time (called a time series) cam look like. At the top is an example of a high correlation in the time series for left and right motor cortex - which you'd expect to show similar patterns of activation. Time in minutes is on the X axis (remember the BOLD response is slow and we get a data point every 2 minutes). And on the Y axis is the level of BOLD activation. You can see that when the red line that is left motor cortex activity goes up the green line that is right motor cortex activity goes up and when red goes down green goes down. At the bottom is an example of a LOW correlations between left motor cortex and left visual cortex. The orange and dark red lines go up and down in a way that seems almost completely independent of each other -- the activity is unrelated. These regions belong to different intrinsic networks. Again, the correlation in brain activity over time between two nodes an edge. A higher correlation between 2 nodes is a stronger connection. The top two regions are working together, dancing in sync. The bottom regions are each doing their own thing.

Some correlations between 2 nodes are stronger than others



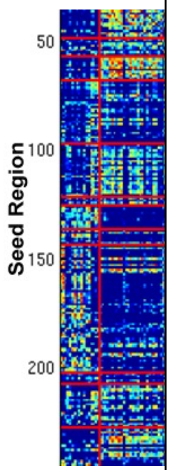
- Each "node" is one of these dots or regions
- Each "edge" is the correlation between 2 nodes. Thicker lines = stronger edges = higher correlation between nodes

Make a Matrix of Correlations

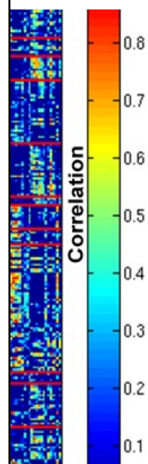


This is also not from the study but just an example of a simple correlation matrix. Ignore the negative correlations because in this study they treated them as positive. Now you get a number for the strength of a correlation for each pair of nodes -- for each edge== and you arrange those numbers in a matrix showing the strength of the relationship between every combination of pairs. Here in this example, which is not from the study, brighter colours indicate a higher correlation (positive or negative). So when you make a matrix with 268 x 268 nodes you get something like this.

Result: Really bi



It's OK!
It's just a whole lot of correlations!
We will use even more correlations to distill them into two simple "summary statistics" for "good" and "bad" attention



Questions?

Fuzziest point?

- What is a positive/negative tail?
- What is the summary statistic the calculated for each?
- What do the results scatter plots show?
- How did the SAN model predict resting state data?
  - Take a moment to think/write it down
  - Turn around and talk to your neighbor about what you're most confused about so far

Step I (part 2)

Assess what patterns of brain connectivity are relevant to attention

a. Measure correlations between pairs of nodes

b. Test which patterns of connection between regions is linked to good and bad gradCPT performance (d') *between* participants

Now the second part of step 1. Test which patterns of connection between regions is linked to good or bad performance on the task -- measured by d'. "First, robust regression between each edge in the connectivity matrices and d' performance was performed across subjects." They looked at the strength of the connection between every pair of nodes – each edge -- and tested whether the strength varied across subjects with attention (was higher or lower in the same way that the attention score varied (was higher or lower). A lot of tests! 268 x 268!

Correlate edge strength with good and bad gradCPT performance

- Positive tail: Edges most correlated with good performance (757 edges) – it's a network.
- Negative tail: Edges most correlated with bad performance (630 edges) – it's a network
- Network strength: *Summary statistic* calculated for each tail by adding up all of the correlations

27/2/25 PSYC 365: Class 14 30

Next Round of Correlations. Once they correlated each edge with performance on the sustained attention task they divided the edges into two categories. Edges whose strength was associated with better performance, and edges whose strength was associated with WORSE performance. Basically they took the correlations between pairs of nodes that were most correlated with performance on the task and sorted them into two piles: One pile predicted good performance and one pile predicted bad performance. They called each of those networks of nodes that fell out from their relationship with good and bad performance a tail. These tails reflected the strength of all of the correlated activity between the pairs of nodes whose correlated activity individually best predicted grad cpt performance. The two tails together made up < 8% of the total number of edges. [1] Magic! [2] you wind up with two simple numbers.

Summary Statistics

- Sum up the correlations of all the edges in each tail to get the summary statistics of network strength
- Positive network strength – predicts high attention!
- Negative network strength –predicts low attention!
- Both measures of network strength correlated with d'

Those numbers added up in each tail were the two summary statistics indicating **Positive network strength and negative network strength.** These statistics were a measure of how strong the connectivity was for each person in the positive and negative networks.
They then could use both positive and negative network strength as predictors with d' as the dependent measure of sustained attention performance between participants. So for the positive network if they had a higher score, which meant more strong correlations between the nodes in that tail, or higher network strength, they had better attention task performance. In the negative network, if they had a higher score, meaning more strong correlations between THOSE nodes, they performed worse. The researchers confirmed that the the **summary statistic for the whole brain was** the best predictor of gradCPT socres (better say than individual edges). Now we're ready to test our predictions, based on large-scale brain activity, against real data.

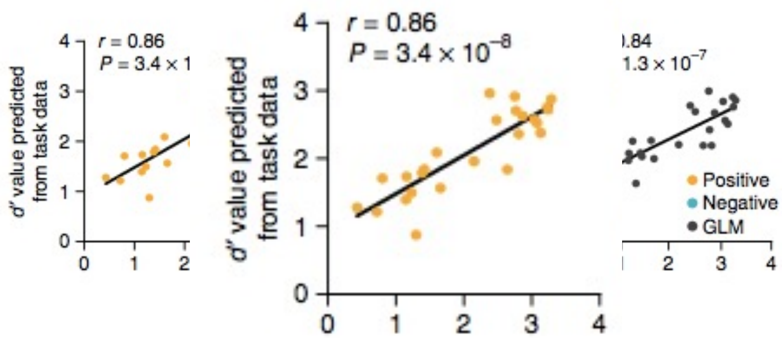
But what is the *SAN model*?

- SAN model
  - A statistical model that predicts an individual's sustained attention ability from *positive and negative network scores*
  - It is a model of the relationship between network strength and d' scores observed in the Yale group during the grad-CPT task

[1] [2] The model is trained on group data to predict an individual's data. Internal cross validation is a way of making sure the results are solid even when you are making a whole lot of statistical comparisons. That is to say, in the cross-validation procedure, they took one participant out and calculated the edges that best predicted high and low attention in the group without that person. They then used those groups of edges (or networks) to calculate a summary score for that person. Then they used a simple regression model to predict all but the left-out person's d' score from their summary scores in each tail. Those are the SAN models. And calculated the left-out person's predicted d' score from those.
So you can apply the model based on positive tail, negative tail, or both. That's what we'll see in the next slide.

Prediction of gradCPT performance from SAN model

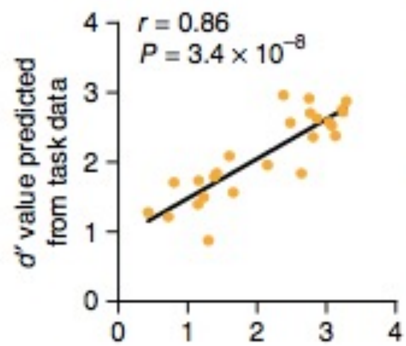
gradCPT fMRI data



THIS IS FROM ONE OF THE FIGURES IN THE PAPER. Now that they've trained the SAN model on the group data they can use it to predict any one individual's attention score based on THEIR pattern of brain activity in the positive and negative networks. So basically based on summary scores from the networks made up of edges that predict high and low attention in all but one of the group (that's 24 participants), you predict how a new person would do on the gradCPT task based on their individual brain network summary scores (a measure of the strength of connectivity in that network). And then you compare it to how the new person actually did on the gradCPT task. You do that for every participant -- taking their data out and predicting it with the other 24 participant's data to get predicted scores for everyone. You then correlate the predicted scores with the actual scores. That's what the scatter plots show. They show how well the scores for each person predicted by the model correlated with how well each person actually did [1] Let's just zoom in to see what we're looking at in all 3 plots. This shows correlations between the d' values that were predicted from the SAN measures of network strength (on the Y axis) and the actual gradCPT *d*' values (on the x axis). Here we see predictions by the positive network in yellow. [2]. Here is the same thing for the negative network in blue. The scatter plot with the the black dots on the far right represent predictions of general linear models (that's what GLM stands for) which are regressions -- statistical relationships that om this case took into account BOTH positive and negative network predictions when predicting attention task performance (instead of either one or the other. If you're thinking the negative network should show a negative correlation (negative network scores were negatively correlate with performance). bear in mind that what is on the Y axis is not the positive and negative summary scores themselves but the d' performance PREDICTED by those scores. [3] Take home: functional connectivity scores predict sustained attention performance**."** Looking good. That suggests the model works – at least with brain data from the gradCPT task.

Question

- Wait, isn't this is a small sample correlation!
- How did they control for false positives?



Cross validation. Leave one participant out, re-calculate the positive and negative tails based on all the remaining participants connectivity data, and predict the left-out participant's performance from the other participants' scores. Repeat with every participant having in turn being left out.

This is the section of the Results titled: *Internal validation: prediction from task connectivity*

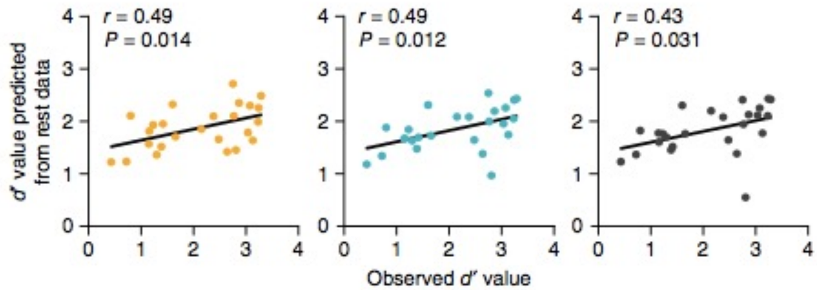
Step 2

- See if SAN model, based on brain networks identified during performance of the gradCPT task, can predict gradCPT task performance from tail scores calculated from **resting state** data in the same people

Now we have a way of calculating brain scores using peoples' brain activation while performing a sustained attention task and using those to predict peoples' performance on the task. But that's kind of circular. The authors say that this would tell us that this large set of brain connections is highly related to sustained attention ability as a basic property of the individuals' brains, and not just when they were using their attention muscles during the explicit attention task.

Prediction of gradCPT performance by SAN model using Yale resting state data

Resting fMRI data



SAN network strength from resting state data predicted grad CPT task performance
Predicted performance was reasonably correlated with actual performance

They then applied this approach to the participants' resting state data So they created summary scores from the SAME sets of edges (high and low attention) used in the SAN model from the brain activity of the Yale group at rest rather than while they were doing the gradCPT task to predict what each participants' performance should be on the task . Or as it says in the paper "in other words, the summary statistic of network strength was calculated based on an individual's resting-state connectivity matrix rather than his or her task-based matrix." This is the same kind of plot we saw before, with d' scores predicted from resting state data on the y axis and their actual d' scores on the X axis. Once again scores predicted from the positive network are in yellow, from the negative network are in blue, and in the regression that included both network scores in black. Here correlations are .49 and .43, so you can see the correlations were stronger when predicted d' was based on the brain data while they were actually doing task than when they were at rest. This is because the model was trained on task data not resting state data. So the patterns of network activation may be similar, but not identical during a task and at rest. But still, the predictions weren't bad. They also controlled for age and IQ and for hyperactivity to verify that the results are really about attention and not general cognitive ability or overall arousal/hyperness. So they conclude that [1]

Questions?

- Fuzziest point?
  - What exactly are the SAN networks made up of?
  - What exactly predicted ADHD scores in the Beijing group?
    - Think/write
    - Turn around and talk to your neighbor about what you're most confused about so far

Step 3

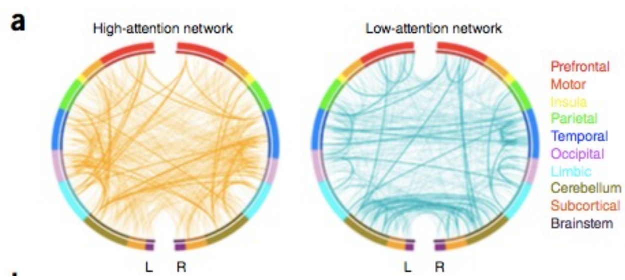
See if you can use the SAN model (created from Yale data during gradCPT task) to predict ADHD scores in kids and adolescents in Beijing based on network scores calculated from their resting state data.

So now we're using resting state data in a whole new population to calculate brain scores from these tails, or sets of edges, that were defined by the Yale group as the networks that best predicted good and bad task performance. And remember some of these kids had ADHD diagnoses and some of them did not.

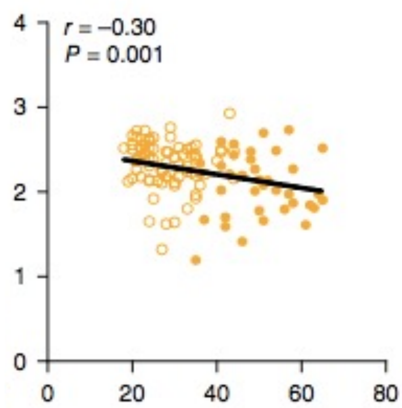
The SAN Networks (sets of nodes and edges)

- High Attention Network
- The 757 edges that most reliably appear in the positive network – stronger connectivity predicts higher d'

- Low Attention Network
- The 630 edges that most reliably appear in the negative network – stronger connectivity predicts lower d'



Beijing Group: SAN models predict ADHD scores



- *SAN network strengths predict ADHD scores in a whole new population!*
  - Higher ADHD-RS scores mean more severe deficit/ADHD symptoms
  - Higher predicted d' means lower ADHD scores
  - Relationship was specific to attention – it wasn't related to age or IQ!

So at this point we know that both task data and resting state data predict performance on the grad cpt in the group from Yale. But now the researchers are taking it a step further. They're activation in the SAN networks that were associated with gradCPT performance in the first group, are creating summary scores from the SAN networks for each participant based on resting state data from the Beijing group, and are now using this group's SAN scores to predict a new outcome variable -- the ADHD RS in the Beijing group. Let's look at the predictions of the positive network first. Now on the Y axis is the ADHD-RS score predicted by the summary score calculated from resting state data and on the x axis is the actual ADHD RS score. Solid dots are kids with ADHD diagnoses and hollow dots are kids without. The correlations are negative in this case because the models were trained on measures of GOOD attentional performance (d') which is negatively related to ADHD symptoms. So low ADHD scores are correlated with high positive network strength. [1] Here are the correlations with Positive or high attention and Negative or low attention network models and the GLM with both. [2]... The authors conclude there is a "Meaningful overlap between the neural mechanisms that are important for sustained attention and the neural patterns that lead to an ADHD diagnosis"!

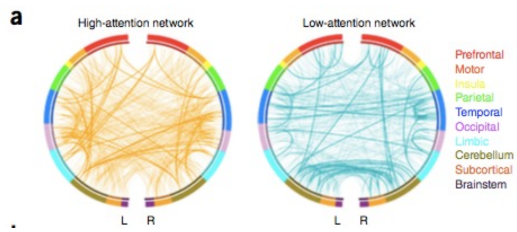
Summary

High Attention Network

- Individuals with stronger connectivity in this network showed *less severe* symptoms of ADHD (low ADHD-RS scores)

Low Attention Network

- Individuals with stronger connectivity in this network showed more severe symptoms of ADHD (high ADHD-RS scores).



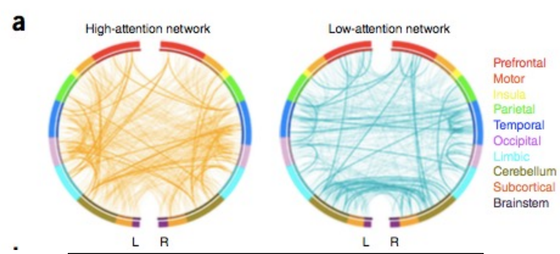
The SAN networks

High Attention Network

- connections between motor & occipital cortex & cerebellum
- connections between subcortical/cerebellar, motor and visual *networks*

Low Attention Network

- connections between temporal and parietal regions & between hemispheres in temporal and cerebellar regions
- connections within subcortical and between subcortical and frontoparietal *networks*



So you might be wondering – what regions were important to each SAN network? They divided the nodes into "macroscale regions" which are the bands of solid colour on the outer edges. [1] [2] [3] They note the importance of the cerebellum's role in cognition – in relation to the old view that the cerebellum is entirely about motor control. Interesting that poorer sustained attention was associated with MORE connections between the FP network, associated with cognitive control, and subcortical networks. They also "lesioned" canonical networks by taking them out of the data analysis one by one, and found that they were still able to predict performance so it didn't depend on any one canonical network. The importance of individual nodes to attention performance really depended on what other regions they were paired with not how many connections they had overall. Main difference between networks was not so much in the strength of connections but the nodes that were paired with each other.

Reading question

Based on their results, Rosenberg et al. believe that ADHD is best considered to be a:

a. All-or-nothing disorder
b. Disorder in general cognitive abilities
c. Continuum of neural and behavioural dysfunction
d. Disorder in sustained attention and cognition

Rosenberg Summary 1

- SAN model allows functional connectivity between many nodes from many canonical networks to predict cognitive ability across different populations
  - "Holistic neural index of sustained attention"
- Benefits of resting state data
  - Can be collected quickly and easily
  - Predictive not just descriptive
  - Not confounded by differences in task performance due to age or training
- Meaningful link between ADHD and sustained attention
- Does NOT suggest that attention is a unitary process!

"We demonstrate that the strength of functional brain networks predicts sustained attention in previously unseen individuals." [1] In the Discussion they really emphasize the importance of the finding that networks that were created based on brain activity during the gradCPT in a group of 25 Yale students predicted ADHD scores in a large group of kids and teenagers (with and without ADHD) in Beijing. That is, patterns of connections between the regions in each of the two tails that predicted better task performance in the Yale data set predicted less severe ADHD symptoms in the Beijing University data set, and connections that predicted worse performance predicted more severe ADHD symptoms. This suggests there is really a meaningful link between sustained attention and ADHD. [2] ...In general they point out that RS data is really useful in populations that have trouble doing tasks. **Q: What might be other populations like that?** [3] [4] It is likely that the overall sustained attention factor measured here includes many cognitive and attentional processes (such as inhibition) that are captured by the functional connectivity analyses.

Rosenberg Summary 2

- Sustained attention emerges from coordinated activity all across the cortex, subcortical structures and the cerebellum
- Underlying networks extend far beyond traditional regions and networks
- Serves one of the main goals of human neuroimaging:
  - Identify *neuromarkers* that predict educational or health outcomes
- Results suggest ADHD is a continuum of neural and behavioural function

What has the SAN model done for us lately?

Functional connectivity predicts changes in attention observed across minutes, days, and months

Monica D. Rosenberg, Dustin Scheinost, Abigail S. Greene, Emily W. Avery, Young Hye Kwon, Emily S. Finn, Ramachandran Ramani, Maolin Qiu, R. Todd Constable, and Marvin M. Chun

PNAS February 18, 2020 117 (7) 3797-3807; first published February 4, 2020 https://doi.org/10.1073/pnas.1912226117

Here's a more recent paper from the same lead author and lab. Here is what they say "Sustained attention varies across people and fluctuates over time. Patterns of functional brain connectivity predict a person's overall sustained attention ability, but do they predict changes in attentional state? Here, across five studies, we show that the same functional connections that predict overall sustained attention ability predict attention changes observed over minutes, days, weeks, and months. Furthermore, these functional connections are sensitive to cognitive and attentional state changes induced by anesthesia. Thus, fluctuations in the same functional connections that predict attention in part reflect fluctuations in attentional state."

If you want to measure sustained attention, look at the whole brain!

- In what ways could these methods or findings be applied to help people who struggle with ADHD?
- What about other real-world problems?

Discussion time!

Consider the TED talk by Anil Seth, "Your brain hallucinates your conscious reality." Explain his take that hallucination is uncontrolled perception while perception is controlled hallucination. How does this relate to what we have learned about the visual system, object recognition, and other modes of perception?

Discussion time!

You mentioned when talking about the Egner et al. paper that surprise has twice the impact then expectation on FFA activation, but do we know if this is always the case? Would surprise always be more impactful than expectation, or in something like a new situation where we don't have any expectations does this interaction change?

Also do we know if this ratio is held across brain regions where predictive coding is likely used, for example in other modalities like hearing?

Discussion time!

Since the predictive coding model relies on updated expectations of what is being perceived, how does the brain recognize objects of completely unfamiliar scenes or objects that do not fit into any existing categories? Does it create a new category or attempt to fit it into a pre-existing one? How is this process altered in individuals who have autism?

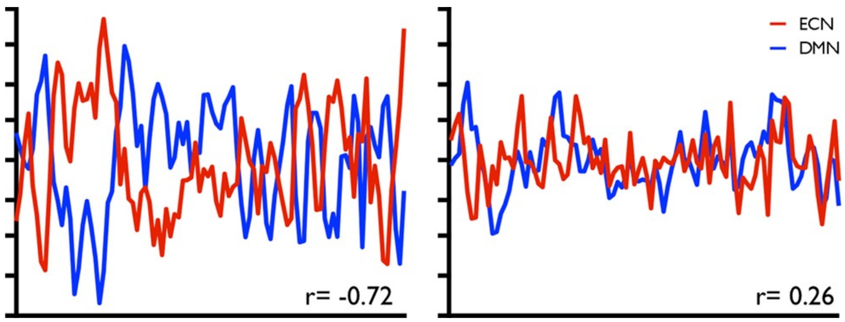
Discussion time!

I was once kayaking in a lake and saw a large figure in the not-too-far distance, but could not identify what it was. I continued staring at it and squinting and moving my head around to get a better look so that I could identify the object and I could feel the prediction programming in my brain working hard, but never was able to identify it. What is interesting about this experience is that I had so much discomfort in not knowing what it was. To this day I get anxiety thinking about my inability to identify this object despite perceiving it. **I am curious what you think might be going on chemically and functionally in one's brain when they sense this emotional and physiological unease from a failure to match up bottom-up information to top-down schemas of the world, and how this phenomenon might be explained psychologically?**

For next time

- Read Inman et al., 2023
- Submit multiple choice question for Midterm 2

(extra slide) fMRI Correlations



Negative Correlation | Small Correlation

27/2/25 | PSYC 365: Class 14 | 53

Classical fMRI correlational analyses don't handle phase very well at all. It simply takes a value for the BOLD amplitude at each time point and doesn't take lag into account. So a slight lag in time will look like a weaker relationship. And a 180 degree phase difference will look like a negative correlation. Newer techniques are now being developed to take phase information into account. They did a fisher's z transform of the correlations which renders the sign (neg or pos irrelevant). So \a negative correlation would be treated the same as a positive. And if a high negative just means its in a strong antiphase relationship that would make some sense. That would still be a strong relationship. But if you had a correlation that was weakly but consistently out of phase that would be a weaker correlation. But for this study the strength of any individual correlation doesn't really matter – it is simply that the pattern of ALL THE CORRELATIONS that are observed correlates strongly with behaviour. So in a sense the approach in this study is resilient against problems of phase that could seriously influence interpretation in other studies.