CBPD Functional Network Results

Ursula Tooley

8/3/2019

Contents

Data Descriptives

Motion

We start with n= 75. In these analyses, I used a stringent pipeline, censoring and interpolating over vols with $>$ FD of 0.5 mm or DVARS > 1.75. Those are excluded from the timeseries used to calculate the networks used here. I excluded anyone with $> 50\%$ frames censored (following Yeo, 2011) or mean motion > 1 mm, or max motion over 10 mm, leaving us with 70 participants, 26 of which have a second run.

The dataframe below shows all runs for subjects who meet this criteria. At the moment, I'm just examining the first run, not averaging in those who have a second run.

Data Frame Summary

motion

N: 92

```
<tr>
  <th align="center"><strong>No</strong></th>
  <th align="center"><strong>Variable</strong></th>
  <th align="center"><strong>Stats / Values</strong></th>
  <th align="center"><strong>Freqs (% of Valid)</strong></th>
  <th align="center"><strong>Graph</strong></th>
  <th align="center"><strong>Valid</strong></th>
  <th align="center"><strong>Missing</strong></th>
</tr>
```

```
<tr>
  <td align="center">1</td>
  <td align="left">age_scan
```

[numeric]

```
  <td align="left">mean (sd) : 6.98 (1.31)
```

min < med < max : 4.11 < 7 < 10.59 IQR (CV) : 1.95 (0.19)

```
  <td align="left">68 distinct values</td>
  <td align="center" border="0"><img src="data:image/png;base64,iVBORw0KGgoAAAANSUhEUgAAAJYAAABkCAYAAAAB
  <td align="center">92
```

(100%)

```
  <td align="center">0
```

(0%)

```
</tr>
<tr>
  <td align="center">2</td>
  <td align="left">run
```

[factor]

```
  <td align="left">1. run-01
```

2. run-02
3. run-03
   65 (70.7%) 26 (28.3%) 1 (1.1%)
   92 (100%)
   0 (0%)
   3
   pctSpikesFD [numeric]
   mean (sd) : 0.09 (0.1) min < med < max : 0 < 0.06 < 0.36 IQR (CV) : 0.11 (1.04)
   69 distinct values
   92 (100%)
   0 (0%)
   4
   relMeanRMSMotion [numeric]
   mean (sd) : 0.27 (0.16) min < med < max : 0.04 < 0.22 < 0.78 IQR (CV) : 0.18 (0.6)
   92 distinct values
   92 (100%)
   0 (0%)
   5
   nSpikesDV [integer]
   mean (sd) : 11.03 (9.16) min < med < max : 0 < 9 < 35 IQR (CV) : 15 (0.83)
   30 distinct values
   92 (100%)
   0 (0%)
   6
   relMaxRMSMotion [numeric]
   mean (sd) : 2.68 (2.3) min < med < max : 0.13 < 2 < 9.64 IQR (CV) : 2.99 (0.86)
   92 distinct values
   92 (100%)
   0 (0%)
   Generated by summarytools 0.8.7 (R version 3.4.3)2019-08-06

SES

```
## Adding missing grouping variables: `ID`
```

Data Frame Summary

ses

N: 70

```
<tr>
  <th align="center"><strong>No</strong></th>
  <th align="center"><strong>Variable</strong></th>
  <th align="center"><strong>Stats / Values</strong></th>
  <th align="center"><strong>Freqs (% of Valid)</strong></th>
  <th align="center"><strong>Graph</strong></th>
  <th align="center"><strong>Valid</strong></th>
  <th align="center"><strong>Missing</strong></th>
</tr>
```

```
<tr>
  <td align="center">1</td>
  <td align="left">ID
```

[factor]

```
  <td align="left">1. sub-CBPD0002
```

2. sub-CBPD0015
3. sub-CBPD0018
4. sub-CBPD0020
5. sub-CBPD0023
6. sub-CBPD0025
7. sub-CBPD0035
8. sub-CBPD0037
9. sub-CBPD0038
10. sub-CBPD0040 [66 others]
    1 (1.4%) 0 (0.0%) 0 (0.0%) 0 (0.0%) 1 (1.4%) 1 (1.4%) 1 (1.4%) 1 (1.4%) 1 (1.4%) 1 (1.4%) 63 (90.1%)
    70 (100%)
    0 (0%)
    2
    race2 [factor]
    1. White
11. Black
12. Other
    17 (25.4%) 41 (61.2%) 9 (13.4%)
    67 (95.71%)
    3 (4.29%)
    3
    ethnicity [factor]
    1. Not Hispanic or Latino
13. Hispanic or Latino
    60 (85.7%) 10 (14.3%)
    70 (100%)
    0 (0%)
    4
    has_diagnoses [integer]
    mean (sd) : 0.03 (0.17) min < med < max : 0 < 0 < 1 IQR (CV) : 0 (5.87)

0 : 68 (97.1%) 1 : 2 (2.9%)
70 (100%)
0 (0%)
5
parent1_edu [integer]
mean (sd) : 14.9 (2.71) min < med < max : 12 < 14 < 20 IQR (CV) : 6 (0.18)
12 : 26 (37.7%) 14 : 10 (14.5%) 16 : 13 (18.8%) 18 : 16 (23.2%) 20 : 4 (5.8%)
69 (98.57%)
1 (1.43%)
6
parent2_edu [integer]
mean (sd) : 13.46 (4.36) min < med < max : 0 < 12 < 20 IQR (CV) : 4 (0.32)
0 : 3 (5.1%) 10 : 4 (6.8%) 12 : 27 (45.8%) 14 : 4 (6.8%) 16 : 8 (13.6%) 18 : 7 (11.9%) 20 : 6 (10.2%)
59 (84.29%)
11 (15.71%)
7
income_median [integer]
mean (sd) : 62761.19 (55231.08) min < med < max : 2500 < 42500 < 2e+05 IQR (CV) : 67000 (0.88)
11 distinct values
67 (95.71%)
3 (4.29%)
8
monthslive_iflostincome [integer]
mean (sd) : 2.7 (1.47) min < med < max : 1 < 3 < 5 IQR (CV) : 3 (0.54)
1 : 20 (30.3%) 2 : 11 (16.7%) 3 : 16 (24.2%) 4 : 7 (10.6%) 5 : 12 (18.2%)
66 (94.29%)
4 (5.71%)
9
childaces_sum_ignorenan [integer]
mean (sd) : 1.11 (1.37) min < med < max : 0 < 1 < 6 IQR (CV) : 2 (1.23)
0 : 32 (45.7%) 1 : 16 (22.9%) 2 : 11 (15.7%) 3 : 7 (10.0%) 4 : 2 (2.9%) 5 : 1 (1.4%) 6 : 1 (1.4%)
70 (100%)
0 (0%)
Generated by summarytools 0.8.7 (R version 3.4.3)2019-08-06

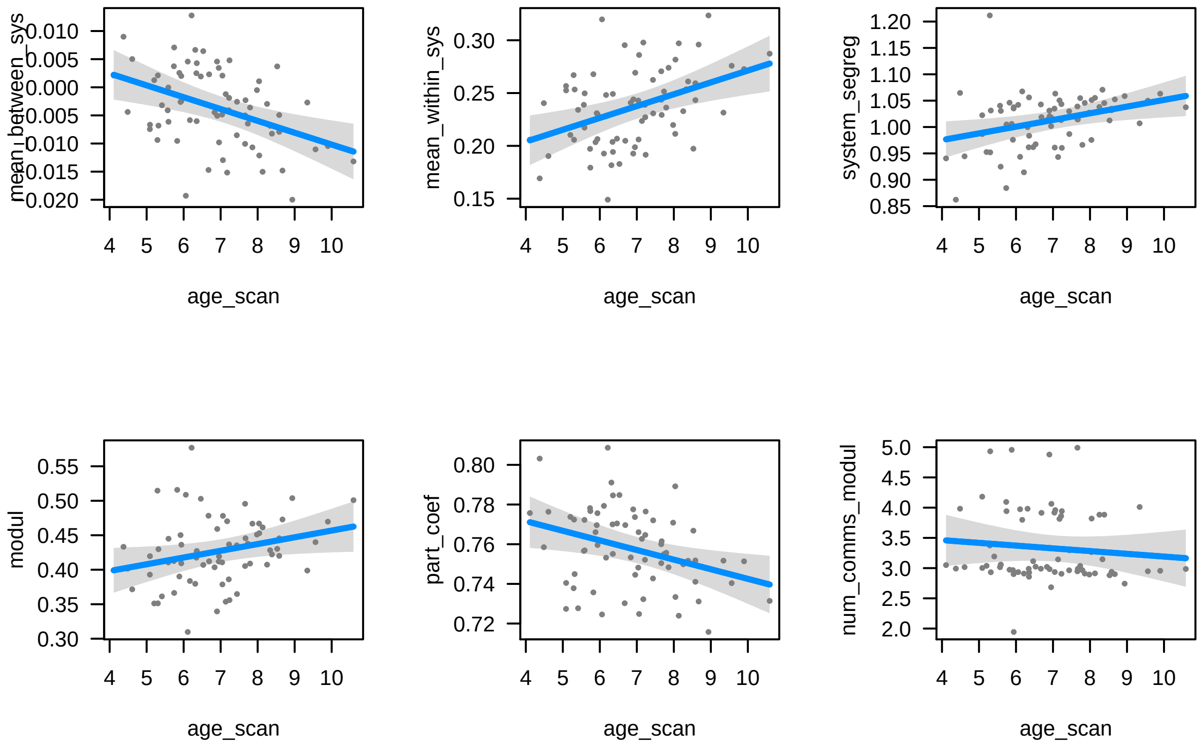
Cognition-NA for now

WPPSI and WISC scores are not totally finalized, so decided not to look at cognition for the moment, unless environmental effects are not interesting. For the future.

Age and measures of network segregation

We find that overall, all measures of network segregation significantly increase with age in our dataset, with remarkable consistency. We examined within-network connectivity, between network connectivity, system segregation (as calculated in Chan et al. 2018), the modularity quality index, and the participation coefficient (summed across negative and positive weights).

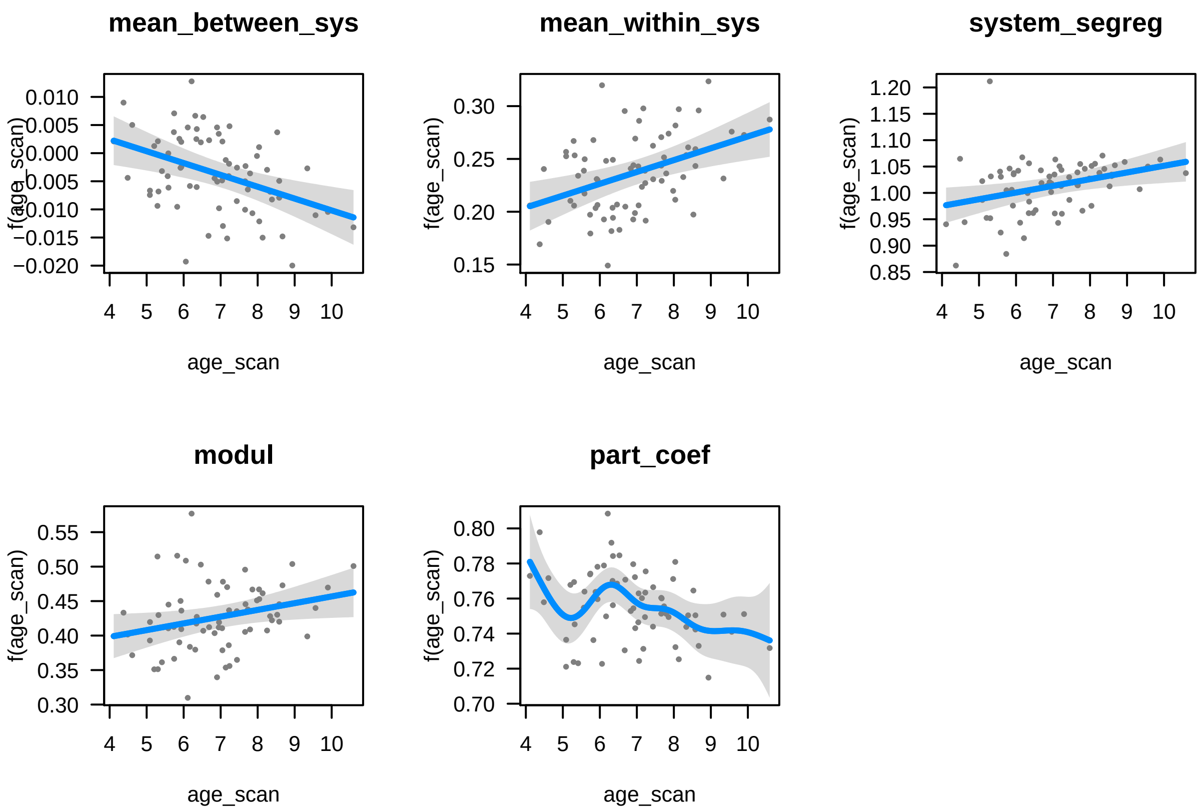
We controlled for sex, mean framewise displacement, percent of FD spikes in the data, the number of volumes a participant had, and the average weight of the functional network.



Interestingly, as a side note, the number of communities detected using modularity maximization on this data decreases slightly with age (range in k is 2, 5), but that looks to just be that we don't have many samples above 8 years old.

Non-linear effects of age?

We also used restricted least ratio tests to test for the presence of non-linearity in our data. As you can see in the plots below, which use GAMs, most measures are linear and look no different from the linear models above.

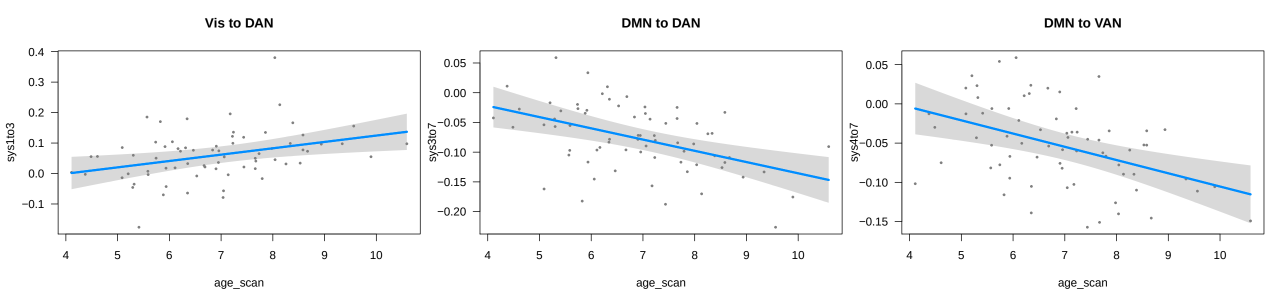


Tests of non-linearity confirmed that no non-linear effects are present. The wiggly participation coefficient line is just over-fitting.

Which systems are driving this effect?

We fit the same model, controlling for sex, mean framewise displacement, percent of FD spikes in the data, the number of volumes a participant had, and the average weight of the functional network, to between- and within-system connectivity for each of the Yeo 7 systems.

We find the strongest effects are in the default mode system, between default and attentional networks.



None of these showed significantly non-linear effects, either.

Effects on default mode connectivity survive FDR correction across the tests conducted, with DMN-VAN *p_fdr*=0.0078839 and DMN-DAN *p_fdr*= 0.0094245. Vis-DAN *p_fdr*= 0.0756036.

Environment and Measures of Network Segregation

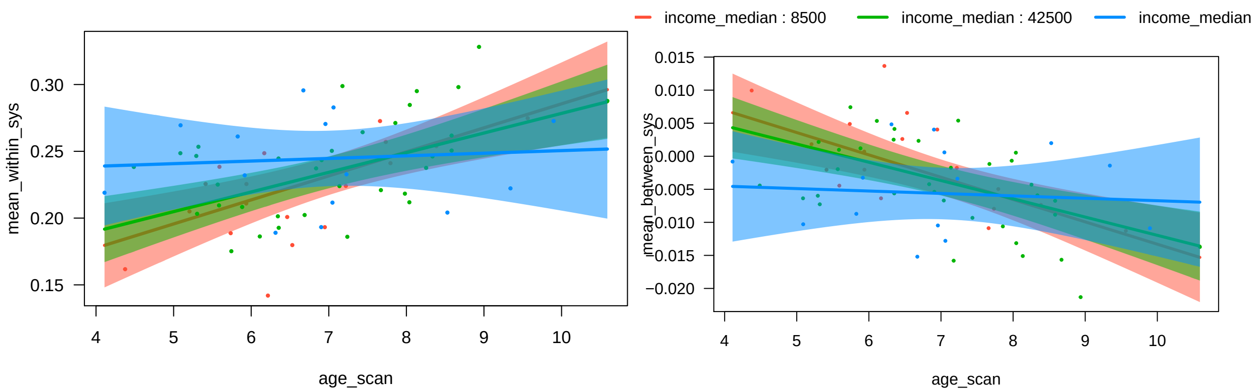
In summary, there are no main effects of SES, race, or ACES sum on any of the measures of network segregation.

Using the same models as above, `age_scan + male + fd_mean + avgweight + pctSpikesFD + size_t`, we see no significant associations between any of the measures of network segregation and either parent 1 education, family income (median of bin), or an SES composite of the two. Race is also not significantly associated with any of the measures of network segregation. I also looked at the sum of child ACES, which doesn't show any significant associations in the model above.

Didn't look at neighborhood questionnaire, not in the data that I have from Julia.

Interactions between environment and age

No significant interactions between any of the environmental variables above and age. However, age x income is marginally predictive of within- and between-network connectivity.



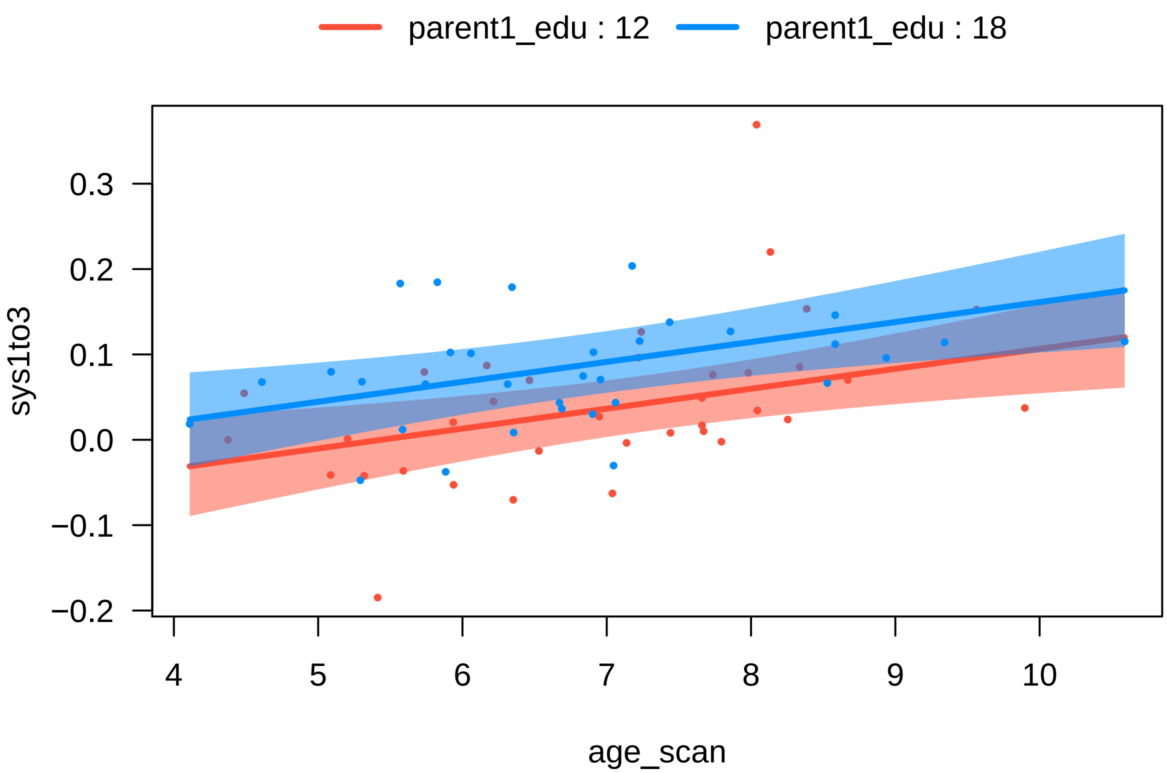
p=0.0884959 for within-network connectivity and p=0.0884959 for between-network connectivity. However, this does not seem consistent across different SES metrics (education vs. income, etc.)

System-specific connectivity and environment main effects

I decided to look specifically at the networks that show strong age effects in our age range, that is, visual to dorsal attention and DMN to attentional networks, with the idea that previously the systems that showed stronger age effects also showed environmental interactions.

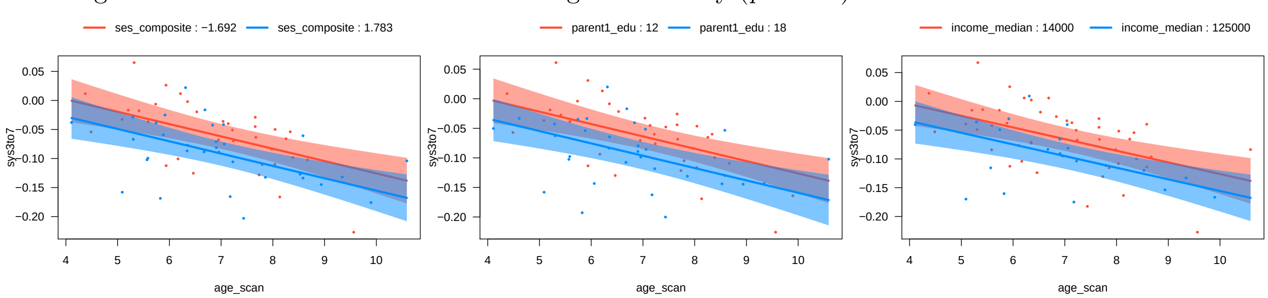
Visual to DAN

- Significant + effect of parent education on avg connectivity ($p = 0.01$)



DMN to DAN

- Significant - effect of ses composite on avg connectivity (p =0.02)
- Significant - effect of parent education on avg connectivity (p =0.02)
- Significant - effect of median income on avg connectivity (p =0.03)



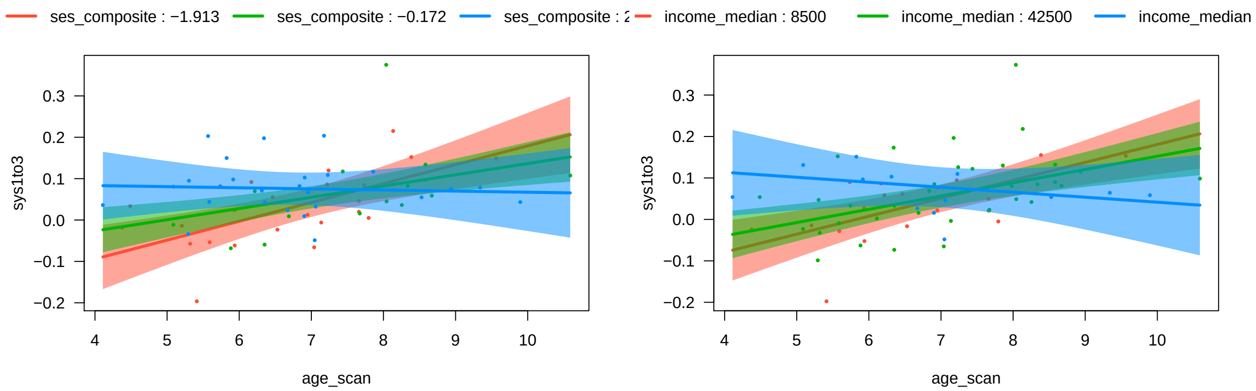
DMN to VAN

- Nothing

System-specific connectivity and environment interactions

Visual to DAN

- Significant interaction between age and ses composite on avg connectivity (p = 0.019)
- Significant interaction between age and income on avg connectivity (p = 0.011)



DMN to DAN

- Marginal interaction between age and ses composite on avg connectivity ($p = 0.09$)
- Marginal interaction between age and parent ed on avg connectivity ($p = 0.05$)

DMN to VAN

- Marginal interaction between age and ses composite on avg connectivity ($p = 0.08$)
- Marginal interaction between age and income on avg connectivity ($p = 0.09$)

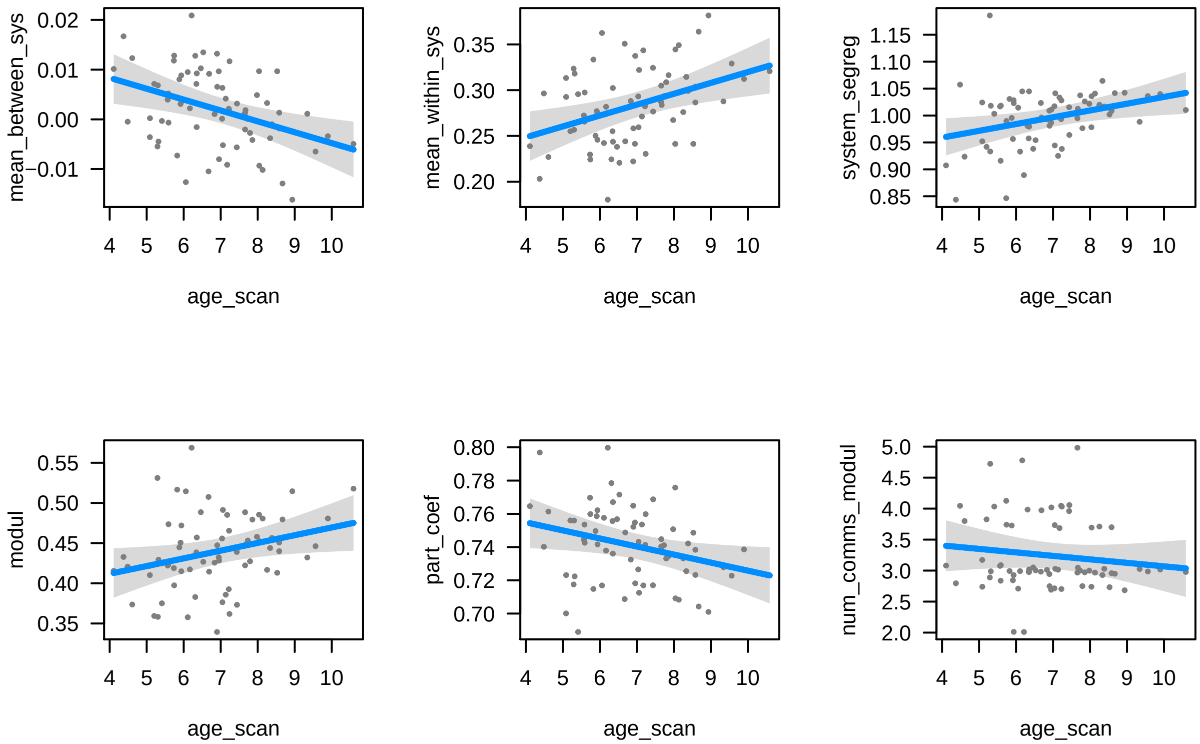
Questions, Notes, & Future Directions

- Other ways to make this more interesting/unique? Specific directions to investigate?
- Sig SES interactions are in visual!
- There are also sometimes outliers on some of the measures that flatten out age or environment trends, I have been looking into using robust regression to examine these trends instead, but did not go through all the trouble to bootstrap p-values for this time around. However, I think it's possible that robust regression might be a more defensible choice.
- Permutation tests?
- Spin tests for parcellations/cortical regions: https://github.com/frantisekvasa/rotate_parcellation
- Can you do this for networks somehow–is this relevant? Graham has done it before, maybe?

Replication parcellation: Schaefer200

Age and measures of network segregation

We used Schaefer200 as the replication parcellation, since it has nice correspondence to Schaefer400 and the Yeo7 systems.



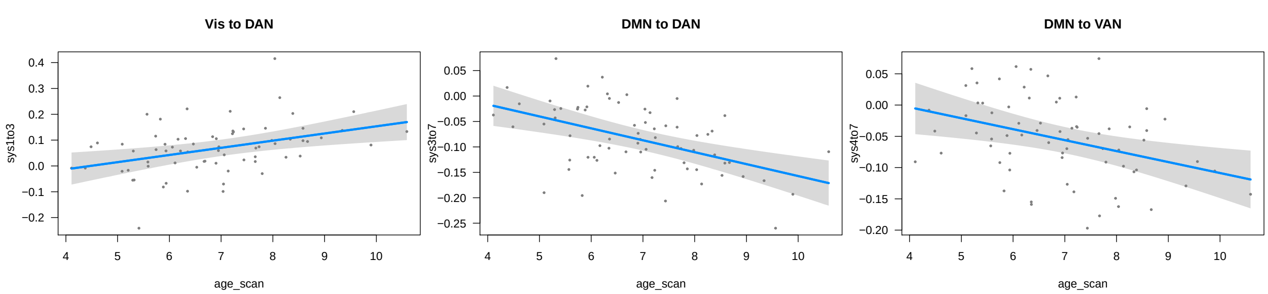
All previous findings significant increases in network segregation with age hold in our replication parcellation! The only finding that is inconsistent is the number of communities detected using modularity maximization decreasing with age, which was likely spurious anyways.

As a reminder, we controlled for sex, mean framewise displacement, percent of FD spikes in the data, the number of volumes a participant had, and the average weight of the functional network.

Also, no non-linear effects of age.

System-level effects

Again, we find the strongest effects in Vis-DAN, DMN-DAN, and DMN-VAN, which are the only effects that pass *fdr* correction.



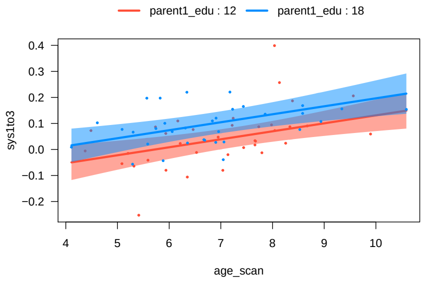
All three effects survive FDR correction across the tests conducted, with DMN-VAN p_fdr=0.0030375, DMN-DAN p_fdr= 0.0394162, Vis-DAN p_fdr= 0.038873.

System-specific connectivity and environment main effects

Again, the effects shown above replicate.

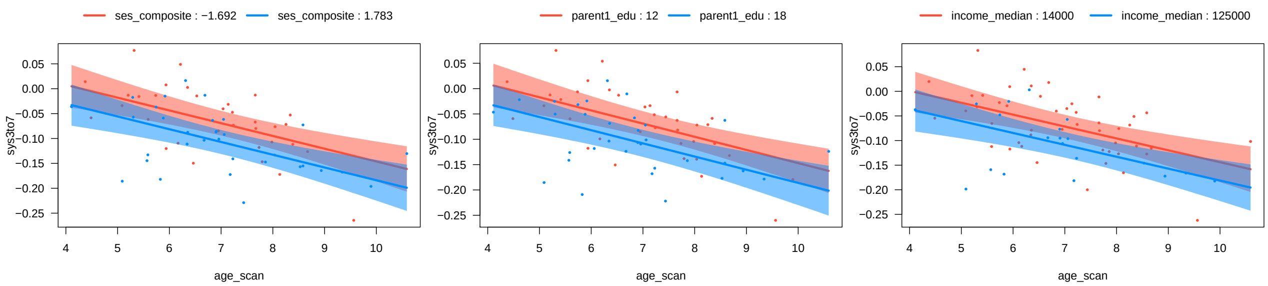
Visual to DAN

- Significant + effect of parent education on avg connectivity ($p = 0.01$)



DMN to DAN

- Significant - effect of ses composite on avg connectivity (p =0.01)
- Significant - effect of parent education on avg connectivity (p =0.02)
- Significant - effect of median income on avg connectivity (p =0.02)



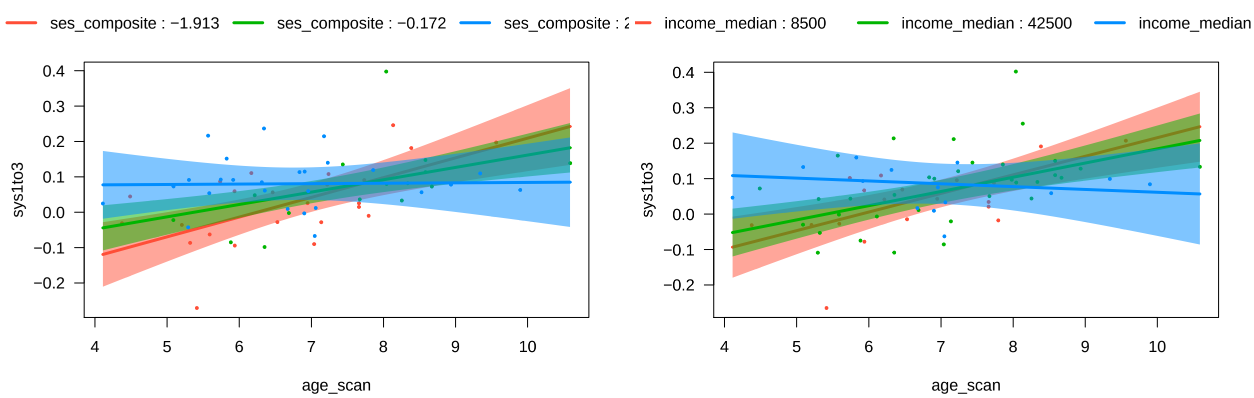
DMN to DAN

- Nothing

System-specific connectivity and environment interactions

Visual to DAN

- Significant interaction between age and ses composite on avg connectivity ($p = 0.02$)
- Significant interaction between age and income on avg connectivity ($p = 0.018$)



DMN to DAN

- Marginal interaction between age and parent ed on avg connectivity ($p = 0.08$)

DMN to VAN

- Nothing.