Characterizing the Network Architecture of Emotion Regulation Neurodevelopment

João F. Guassi Moreira[1], Katie A. McLaughlin[2], & Jennifer A. Silvers[1]*

[1]Department of Psychology, University of California, Los Angeles
[2]Department of Psychology, Harvard University

Summary

The ability to regulate emotions is key to goal attainment and wellbeing. Although much has been discovered about how the human brain develops to support the acquisition of emotion regulation, very little of this work has leveraged information encoded in whole-brain networks. Here we employed a network neuroscience framework in conjunction with machine learning to: (i) parse the neural underpinnings of chronological age and emotion regulation skill acquisition, and (ii) build a working taxonomy of brain network activity supporting emotion regulation in youth. We were able to predict age and emotion regulation ability using metrics from six whole-brain networks during an emotion regulation task. Further, by leveraging analytic techniques traditionally used in evolutionary biology (e.g., cophenetic correlations), we were able to demonstrate that brain networks evince reliable taxonomic organization to meet emotion regulation demands in youth. This work shows that meaningful information about emotion regulation development is encoded in whole-brain network activity, and opens up a realm of possibilities for network neuroscience applications in the study of emotion regulation development.

Introduction

Emotions provide humans with useful internal signals that promote adaptive behavior (e.g., Gershman, Monfils, Norman, & Niv, 2017; Martin et al., 2015). However, emotions can become detrimental if left unregulated. The acquisition of emotion regulation is a protracted process that emerges across child and adolescent development, and is highly consequential for wellbeing, including health risk behaviors and mental disorders (Calkins, 1994; Casey, Jones, & Hare, 2008; Thompson, Lewis, & Calkins, 2008). Over the past fifteen years, substantial efforts have been made to examine brain regions involved in the development of emotion regulation, with most work focusing on activation within, or dynamics between, the amygdala, ventromedial prefrontal cortex (vmPFC), and lateral prefrontal cortex (lPFC) (Ernst, Pine, & Hardin, 2006; Gee et al., 2014; Kim et al., 2013; McRae et al., 2012; Perino, Miernicki, & Telzer, 2016; Pitskel, Bolling, Kaiser, Crowley, & Pelphrey, 2011; Silvers et al., 2017). However, surprisingly little is known about the extent to which whole-brain networks—suites of brain regions that reliably activate in concert (e.g., default mode, frontoparietal)—support the acquisition of emotion regulation skills. This omission is noteworthy given work in adults showing that many emotional, as well as non-emotional, brain functions are supported by whole-brain networks (Gratton et al., 2018; Kleckner et al., 2017). Leveraging network neuroscience – an approach that prioritizes networks rather than individual brain regions as the unit of analysis – in conjunction with machine learning applications may help to fill two major gaps in knowledge regarding the neurodevelopment of emotion regulation. First, this combination of approaches may improve our understanding of how correlated but distinct developmental phenomena such as chronological age and emotion regulation ability are represented in the brain. Second, they may reveal the underlying, formal organization (i.e., taxonomy) of different whole-brain networks involved in the development of emotion regulation.

Using network neuroscience together with machine learning is a promising approach for parsing distinct but related developmental phenomena (Richardson, Lisandrelli, Riobueno-Naylor, & Saxe, 2018), such as chronological age and skills that tend to improve with age (in the present study, emotion regulation ability). The present study focused on two network properties that have been particularly important for understanding cognitive development in other domains and thus are likely to be important for delineating chronological age and emotion regulation ability. The first network property is network differentiation—the extent to which activity within an individual network is similar or distinct across different task states, perhaps reflecting the extent to which a network is "specialized" for a given task. Examining network differentiation in combination with behavior can help determine which networks become specialized to support skill development (i.e., improved emotion regulation ability), regardless of age. While some patterns of network differentiation may correspond with changes in specific skills, others that support more general aspects of development may show increased specialization with age rather than tracking with performance on a given task. For instance, prior research shows greater specialization among pain perception and mentalizing networks tracks with chronological age, but not skills related to representing others' states (Richardson et al., 2018). Extrapolating to emotion regulation, this implies that some networks might support domain-general development with chronological age (e.g., the default mode network), whereas others specifically support

emotion regulation abilities (e.g., the salience network). The second property is network community structure, or the way in which brain networks are divided into sub-structures known as “communities” or “modules” (Sporns, 2010; Yeo et al., 2015, 2011). In the context of emotion regulation, a network may possess multiple modules that support different emotion regulation sub-processes (e.g., holding emotion regulation goals in working memory or self-monitoring to evaluate emotion regulation success). We assessed network community structure here by examining the number of communities that emerged in a given network between task states as well as modularity, the strength of intra- versus inter-module connections. A network’s community structure (number of communities, modularity) and the suite of psychological processes it supports, might track differentially with age and ability, helping reveal the manner in which domain-general maturation (chronological age) differs from growing expertise in a skill (ability). Given that other lines of research have successfully leveraged network differentiation and community structure metrics to build increasingly accurate models of brain function, cognitive skills, and development (Baum et al., 2017; Richardson et al., 2018; Sizemore et al., 2018), it stands to reason that such an approach may also benefit the study of emotion regulation development.

Taxonomical organization—classifying phenomena based on relevant shared characteristics—has been a goal of science for centuries (e.g., Linnæus, 1758). While human neuroscience research has recently begun to adopt taxonomic approaches, such as categorizing task- and rest-based network activity (Bolt, Nomi, Yeo, & Uddin, 2017; Braunstein, Gross, & Ochsner, 2017; Eisenberg et al., 2018; Power et al., 2011; Yeo et al., 2011), no prior work has attempted to generate a taxonomy of whole-brain networks involved in the neurodevelopment of emotion regulation. Understanding how brain networks developmentally organize themselves may reveal how cognitive and affective subcomponents of emotion regulatory processes are related in ways that behavioral or univariate analyses alone cannot. For example, if a network involved in bottom-up emotion generation (e.g., salience) is found to be taxonomically similar to one involved in cognitive control (e.g., frontoparietal), that would imply that regulatory processes are executed by cognitively changing the bottom-up, ‘raw’ representation of a stimulus. Alternately, if the two networks in this hypothetical example are taxonomically distinct, that would imply that cognitions during regulation act on an emotional substrate via top-down or inhibitory processes. While neural taxonomies can be constructed at any developmental stage, they may be especially useful to establish in youth because taxonomies at this point are likely to capture information about ontogenetic relationships between networks. Creating a taxonomical organization of emotion regulation in a pediatric sample obviates the need for speculation about ontogeny in an adult sample and might yield a more useful baseline against which future, incremental work in adults can be compared.

In the present study, we sought to achieve two ends. First, we used the three aforementioned metrics of network activity (network differentiation, number of communities, modularity) in conjunction with classifier analyses to predict age and regulation ability in a sample of youth who completed an emotion regulation task during functional magnetic resonance imaging (fMRI) scanning. Second, we leveraged analytical techniques popular within

evolutionary biology (agglomerative clustering, cophenetic correlation coefficient) to build a working taxonomy of brain networks as they pertain to the development of emotion regulation.

Our analyses were centered on six whole-brain networks: the default mode network (DMN), the frontoparietal network (FPN), the salience network (SaN), the dorsal and ventral attention networks (DAN, VAN respectively), and a 'cognitive reappraisal network' (CRN; Buhle et al., 2014) commonly observed during cognitive reappraisal, a widely studied form of emotion regulation that was evaluated in the present study. The DMN, FPN, and SaN are relevant for emotion regulation processes insofar that they have been implicated in self-projection, executive function, and emotion generation (Buckner & Carroll, 2007; Feldman Barrett & Satpute, 2013; Marek & Dosenbach, 2018). The DAN and VAN were included because externally-evoked emotional responses necessarily rely on perceptual and attentional systems to instantiate both emotional and regulatory processes (Pollak, Messner, Kistler, & Cohn, 2009; Siegel, Wormwood, Quigley, & Feldman Barrett, 2018; Tamietto & De Gelder, 2010). The CRN was included because it is the most relevant network to the emotion regulation task incorporated in the present study. This is important because recent work in both adult and developmental samples suggests that task-specific information—and even activity from task-specific networks—is privileged over domain-general information and thus might be highly informative when differentiating between chronological age and task ability (Brody et al., 2019; Greene, Gao, Scheinost, & Constable, 2018). Because our questions were exploratory in nature, our study was guided by open questions about (i) how network properties relate to chronological age and emotion regulation ability, and (ii) the taxonomical organization of whole-brain networks involved in the neurodevelopment of emotion regulation rather than formal hypotheses.

Results

Analytic Overview. In the current study, we scanned 70 children and adolescents with fMRI while they completed four runs of a cognitive reappraisal task that involved alternately regulating (regulate condition) and reacting naturally (emotional baseline condition) to a series of negative and neutral images. Cognitive reappraisal is a widely-studied form of emotion regulation that involves reframing a stimulus as to alter its emotional import. Cognitive reappraisal abilities are known to improve with age during childhood and adolescence (Gross, 1998; McRae et al., 2012; Silvers et al., 2012).

After assessing data quality (QA) and preprocessing, we prepared the data for analysis by extracting beta-series from each node (i.e., ROI) across all six networks (DMN, FPN, SaN, DAN, VAN, and CRN). Correlations among each networks' beta-series were performed and used, in conjunction with representational similarity analysis (RSA) and graph theory, to yield our metrics of network activity. These metrics included network differentiation between task states (regulate v. emotional baseline) and two measures of network community structure: 1) differences in the *number of modules* within a network between task states (regulate v. emotional baseline), and 2) differences in *modularity* between task states (the degree of 'crosstalk' between modules for regulate v. emotional baseline). Notably, the number of modules and modularity are related but distinct—the former tells us how many 'divisions of labor' occur in a network and the latter reflects how rigidly the network appears to follow such divisions.

After examining bivariate associations between emotion regulation ability, age, and our network metrics of interest, we submitted multiple connectivity-based metrics to a parsimonious classification analyses (K Nearest Neighbor, KNN; rationale for this particular classifier provided in the Methods) to determine the extent to which networks encoded meaningful information about age and regulation ability.

Finally, we attempted to establish a working network taxonomy for emotion regulation in youth using methods typically used in evolutionary biology, namely agglomerative clustering (algorithmic grouping based on shared and unique characteristics) and cophenetic correlation analysis (statistical comparison of clustering solutions). The Methods section presents greater detail on each of these steps. Our results follow.

Behavioral Results. Behavioral results from our emotion regulation task have been reported elsewhere (see Guassi Moreira et al., 2018), but we briefly recapitulate them here for convenience. Participants' average negative affective ratings during the 'Far' and 'Look-Negative' trials were 1.95 (SD = .499) and 2.48 (SD = .562), respectively. This difference was statistically significant ($t(69) = -10.37$, $p < .001$). Participants' ability to engage in emotion regulation, quantified as the percent reduction in negative affect for the regulate condition relative to the emotional baseline condition, improved with age ($r = 0.344$, $p<.01$).

Network Identification and Validation. Detailed information on how networks were identified and validated is included in the STAR Methods and Supplementary Materials. Five of our networks (DMN, FPN, SaN, DAN, VAN) were defined via publicly available, meta-analytic maps pooled across over 14,000 studies (neurosynth.org). The CRN was defined using a prior meta-analysis (Buhle et al., 2014). We subsequently validated networks by removing ROI labels and submitting them to a clustering algorithm (walktrap) to determine whether it would recover the original networks. Results suggest that our defined networks were valid, as the algorithm tended to group ROIs from the same network together more frequently than with other networks (Table S10).

Using network approaches to parse chronological age and emotion regulation ability

Correlations Between Age, Emotion Regulation Ability, and Network Activity Metrics. Before training our age and emotion regulation ability classifiers, we examined correlations between our network activity metrics, age, and emotion regulation ability. Correlations (Spearman) were computed after visually examining data for outliers; outliers were winsorized by replacing them with the nearest value[1]. The full set of correlations, and their 90%[2] confidence intervals, are reported in Table 1. The following results were obtained using an edge-defining threshold of r=.5 and correlations using different edge-defining thresholds are reported in the Supplement (see "Varying Edge Defining Thresholds").

Age. Participant age was largely uncorrelated with network differentiation (dissimilarity of activation profiles between task conditions), number of communities between task states, and

[1] This was only needed for four datapoints across the entire set of 18 network level metrics.
[2] Given the exploratory nature of our research, we relaxed our CIs to 90%.

modularity (difference in the amount of intra- versus inter- module connections between task states). One exception to this was found in the DAN and FPN, where age was positively correlated with number of network modules ($\rho = 0.36, 0.16$, respectively). That is to say, older youth exhibited more modules within the DAN and FPN than younger youth when engaging in emotion regulation. By contrast, however, older youth showed *fewer* modules in the CRN while engaging in emotion regulation, compared to our emotional baseline (passive viewing; $\rho = -0.18$).

Emotion regulation ability. Emotion regulation ability was associated with several network activity metrics. First, youth who were better at emotion regulation (i.e., reported greater decreases in negative emotion when regulating compared to baseline) tended to have more differentiated connectivity profiles between the regulation and emotional baseline task conditions ($\rho = -0.19$) in the CRN. In other words, more specialization in the CRN was related to better emotion regulation ability. Better emotion regulation ability was characterized by having fewer communities ($\rho = -0.30$) that are highly modular ($\rho = 0.16$) in the CRN during emotion regulation. Fewer communities in the DAN and VAN during emotion regulation were also associated with better regulatory ability ($\rho = -0.15, -0.17$, respectively). On the flip side, youth with better emotion regulation abilities evinced more communities during emotion regulation in the SaN, DAN and DMN ($\rho = 0.16, 0.20, .16$, respectively).

K-Nearest Neighbor Classification.

We submitted the aforementioned connectivity-based network activity metrics—which all characterize an emotion regulatory state by contrasting it against an emotional baseline—to a KNN classifier to predict age and emotion regulation ability.

Age. We began by classifying age with varying *k* neighbor sizes (3, 5, 7). Three different age cut-offs were tested (mean split, median split, and 13 years) and while the results did not very tremendously between the cut-offs, the mean split (children defined as < 12.7 years, teens defined as > 12.7 years) yielded the best classifier success and are thus reported here (see STAR Methods for more details and Supplement for results with other cut-offs). To evaluate overall model strength, we created a composite F_1 ($F_{1.composite} = 1 / [(1 / F_{1.child}) + (1/ F_{1.teen})]$) based on the children and teen F_1 values. Just as with typical F_1 values, a greater score indicates better model fit. The composite F_1 for the $k = 5$ model (0.616) was the best among the possible neighbor sizes. Using $k = 5$, we were able to predict 73% (95% CI [.56 - .86]) of child participants correctly ($F_{1.child} = .675$), but only 51.5% (95% CI [.34 - .69]) of adolescents correctly ($F_{1.teen} = 0.567$).

We next used network activity features from each individual network to determine whether certain networks were better than others at predicting age. SaN had the best $F_{1.composite} = 0.614$ among all six networks. DAN was most predictive for children (70.3%, 95% CI [.53 - .84], $F_{1.Child} = 0.642$), whereas VAN was most predictive for adolescents (72.7%, 95% CI [.54 - .87], $F_{1.Teen} = 0.623$). Notably, CRN and SaN had teen classification accuracies that just narrowly missed significance (i.e., nearly above chance or 50% accuracy) (ERN: 60.6%, 95% CI [.42 - .77]; SaN: 66.7%, 95% CI [.48 - .82]). Precision, F_1 values, and confidence intervals for all

models and networks are reported in Tables 2 and 3; Classification accuracies are visualized in Figure 4. Supplemental analyses suggested that these results were consistent even when changing the edge-defining threshold in our graphs. This includes finding that the model was still better at predicting children (compared to adolescents) and SaN was still a highly predictive network (see "Varying Edge Defining Thresholds" in the Supplement)

Emotion Regulation Ability. Our best model for predicting emotion regulation ability. Low ability was defined as less than a 20% reduction in negative affect for regulate trials compared to emotional baseline, high ability defined as greater than 20% reduction, based on prior work in similarly aged samples (Silvers et al., 2012). We tested a slightly lower threshold (19%, corresponding to the median) and found comparable, albeit slightly less predictive, results (see STAR Methods and Supplement for details). Our best model used k = 3 neighbors ($F_{1.composite}$ = .517). Using k=3, we were able to predict which youth demonstrated low regulation ability with 69.4% accuracy (95% CI [.52 - .84], $F_{1.low}$ = 0.610) but were unable to reliably predict high ability youth (38.2% accuracy, 95% CI [.22, .56], $F_{1.high}$ = 0.448). In terms of individual networks, the CRN was the most accurate network in predicting youth with both high (52.9% accuracy, 95% CI [.35 - .70], F_1 = 0.545) and low (61.1% accuracy, 95% CI [.43 - .77], F_1 = 0.595) ability, and provided the best fitting composite model ($F_{1.composite}$ = 0.569). Precision, F_1 values, and confidence intervals for all models and networks are reported in Tables 4 and 5; Classification accuracies are visualized in Figure 5. Supplemental analyses revealed that these results were not driven by idiosyncrasies in processing network data (i.e., edge defining threshold). Across multiple edge defining thresholds (see "Varying Edge Defining Thresholds" in the Supplement for full details), the model was consistently better at predicting individuals with lower ability scores than individuals with higher ability scores and the CRN continued to be highly predictive of emotion regulation ability (indexed by $F_{1.composite}$).

Creating a taxonomy of emotion regulation neurodevelopment

The final aim was to provide a neurodevelopmental taxonomy of brain networks implicated in emotion regulation. The first section below describes the creation and validation of three working network taxonomies. The second section details comparisons between these taxonomies.

Taxonomy Building and Validation. We used agglomerative clustering to build three network taxonomies (see Figure 3 for associated dendrograms), one for each of our metrics of network activity: differentiation (dissimilarity of activation profiles between task conditions), number of network communities (difference in number of modules between task states), and modularity (difference in the amount of intra- versus inter- module connections between task states). As described in the methods, a clustering algorithm was used to group networks based on shared network properties.

To test whether the organization of the differentiation (similarity), modularity and community dendrograms were likely due to chance, we bootstrapped our sample, re-created the dendrograms at each iteration, and correlated them with our observed dendrograms (The cor_cophenetic()function of R's 'dendextend' package was used to compute cophenetic correlations, a metric that compares similarity of two dendrograms). The ensuing distribution of

correlations (*c*(observed, a bootstrapping iteration) constituted our empirical sampling distribution. If the 95% confidence interval did not include negative numbers or 0, then we may reject the notion that a given dendrogram's organization capitalized purely on noise and random chance. Results show that the structure of our dendrograms for similarity (c = 0.78, 95% bootstrapped CI [0.14, 0.99]) and modularity (c = 0.67, 95% bootstrapped CI [0.24, 0.99]) are unlikely to occur by chance. By this same logic, we cannot conclude that the dendrogram for network community is stable beyond being a product of noise and random chance (c = 0.78, 95% bootstrapped CI [-0.30, 0.99]).

An informal glance at the resultant dendrograms reveals consistencies across the three taxonomies based on different network features. For instance, the CRN and DMN are grouped in the same clade ("branch") for both the similarity and modularity dendrograms; SaN and DMN are grouped into the same clade for the similarity and community structure dendrograms; FPN and DAN extend from the same branch for all three taxonomies. A formal, statistical comparison of similarity between the dendrograms is detailed in the section below.

Taxonomical Comparison. In our next analytic step, we formally compared the structure of the three aforementioned working taxonomies to one another. Whereas our prior analyses revealed separate network taxonomies based on three unique features of network activity, this analysis compared the degree of similarity between the three different taxonomies. All network taxonomies were positively correlated, but only modularity and community were significantly so (c = .30, 95% bootstrapped CI [.08, .99]). The correlations between similarity and community (c = .42, 95% bootstrapped CI [-.34, .81]), and similarity and modularity (c = .25, 95% bootstrapped CI [-.23, .93]) displayed empirical sampling distributions whose confidence intervals included zero. Supplemental analyses revealed that these results were consistent across multiple edge defining thresholds.

Discussion

In this study, we set out to interrogate the neurodevelopment of emotion regulation through the lens of network neuroscience. By doing so, our goal was to parse the whole-brain network dynamics that contribute to correlated but distinct developmental processes (chronological age and acquisition of emotion regulation abilities), as well as establish a working neurodevelopmental taxonomy of brain network activity during emotion regulation. We found that whole brain network properties could accurately predict which participants were children, but not adolescents, as well as which participants demonstrated low, but not high, emotion regulation ability. Distinct individual networks predicted age and emotion regulation ability. While the SaN was the single network most predictive of age, the CRN was the most predictive for emotion regulation ability. We also found that the CRN, DMN, and, to some extent, the SaN evince taxonomical similarities while the FPN and DAN also appeared to be consistently organized with one another. These findings have implications for our basic understanding of (i) the neurodevelopment of emotion regulation, and (ii) how brain network dynamics may broadly change and interact across cognitive and emotional development.

How and What Brain Networks Encode Information about Development. Our classifier was able to predict which participants were children and which participants had low emotion regulation ability, but not those who were adolescents or had high emotion regulation ability. This speaks to the nature of both development and emotion regulation. First, adolescence is marked by tremendous flexibility across several domains (Blakemore & Mills, 2014; Fuligni, 2018; Hauser, Iannaccone, Walitza, Brandeis, & Brem, 2015; Pfeifer et al., 2009), including in connectome development (Kaufmann et al., 2017; Morgan, White, Bullmore, & Vértes, 2018). That is, brain networks exhibit temporarily less stable activity profiles, likely due to developmental demands from one's environment that require increased flexibility (Doremus-Fitzwater, Varlinskaya, & Spear, 2010; Spear, 2011). As such, one possible interpretation of the present results is that adolescents are more variable in their configuration of brain network activity, making it more difficult for a classifier to identify a consistent "pattern" of adolescent network activity. Second, that our classifier accurately identified individuals who were poor at emotion regulation but not those who were skilled at emotion regulation suggests that there are a number of different ways to successfully regulate one's emotions, but there are only so many ways to fail at doing so. A related inference is that successful emotion regulation might require a degree of flexibility (e.g., switching between different emotion regulation strategies), whereas unsuccessful emotion regulation may reflect a consistent tendency to mentally "check out" or fail to execute relevant cognitive processes (Aldao, Sheppes, & Gross, 2015). In both cases, the fact that the classifier failed to identify groups of individuals characterized by increased variability (e.g., adolescence is a developmental stage characterized by neural variability while high emotion regulation ability is marked by variability) might mean two things. First, it alludes to the existence of possible subtypes among participants such that development and skill acquisition progressively result in a variety of mature network activity signatures (i.e., multifinality). Second, it implies a hierarchical gradient of neural variability (Demirtas et al., 2019) such that variability at the cortical ROI level decreases with age and expertise (Guassi Moreira et al., 2019), while variability at the network level increases with age and expertise. Though the present data cannot directly speak to this possibility, this presents an enticing hypothesis to test in future research.

Intriguingly, we found that the SaN was the best individual network in predicting age, whereas the CRN was the best individual network in predicting emotion regulation ability. The former fits neatly with extant literature that child and adolescent development is marked by dramatic changes in affective responding (Casey et al., 2010; Ernst et al., 2006; Larson & Lampman-Petraitis, 1989). Because the SaN is theorized to be implicated in the bottom-up generation and maintenance of emotional responses (Feldman Barrett & Satpute, 2013), it is somewhat unsurprising that we observed it to be the best network for discriminating between the two age groups in our studies – especially given that data were collected during an affective task and emerging evidence suggests that information about development are uniquely encoded in network activity elicited by affective states (Rudolph et al., 2017). Future work might wish to examine whether similar age effects are observed in different networks during performance of social or cognitive tasks for which ability also changes across development (Blakemore & Mills, 2014; Crone & Dahl, 2012).

The notion that task-specific network connectivity is a salient marker of development is reinforced when considering our results with the CRN, which was the best predictor of emotion regulation ability, and outperformed other canonical networks that are also relevant for emotion regulation such as the FPN and the DMN. This result expands upon recent observations that task-related network activity tends to be highly predictive of related outcomes of interest (Greene et al., 2018). Here we show that activity from a task-relevant network is best at discriminating between individuals with high and low abilities in said task. Specifically, every CRN network metric examined was strongly associated with emotion regulation ability: better emotion regulation ability tracked with greater network differentiation, more local modules, and greater modularity in the CRN. These data point to a division of labor within the CRN, such that different modules become specialized to handle distinct features of reappraisal and this in turn promotes better regulation skills. For example, it might is possible that distinct modules contribute to monitoring regulation success and others to maintaining attention to regulatory goals. In combination with other emerging work in this space, the present results suggest that neural specialization – both locally (Guassi Moreira et al., 2019) and at the network level – support the acquisition of emotion regulation across development.

A (Working) Neurodevelopmental Taxonomy of Brain Network States During Emotion Regulation. The results from our taxonomy building analyses suggest that brain network activity during active regulation in youth evinces a coherent structure. That the DAN and FPN were consistently grouped together implies that cognitive control processes during emotion regulation are heavily reliant on perceptual representations of stimuli in a potentially in a feed-forward manner, and affirm that attention is an important supra-ordinate cognitive process for emotion regulation (Ochsner, Silvers, & Buhle, 2012). Likewise, that the CRN and DMN were repeatedly grouped together suggests that emotion regulation is supported by cognitive representations of the self given that the DMN is implicated in self-processing and reappraisal is thought to require self-monitoring and projection. Another intriguing finding is that networks were clustered hierarchically. This implies the possibility of 'supra-networks', such that certain networks are potentially sub-networks of a higher level network. Alternately, this also potentially suggests that the ontogenetic emergence of certain brain networks (i.e., those that are not part of core sensory/autonomic processes) begins as a one or two network system that becomes increasingly differentiated with age (Collin & van den Heuvel, 2013; Gao, Alcauter, Smith, Gilmore, & Lin, 2015).

The results from our working taxonomy of network activity during emotion regulation also have several potential applications for future research. For instance, understanding how networks are typically organized across childhood and adolescence may help to establish pediatric growth curves of brain development by providing an organizational standard against which individuals can be compared. Such efforts could ultimately facilitate deep phenotyping and precision medicine applications in pediatric neuroscience (Robinson, 2012).

Limitations and Future Directions. This study is accompanied by several limitations. First, we utilized just one machine learning approach, KNN, and thus it is unclear whether results would generalize to other techniques as well as to models using continuous measures of age

(KNN necessitated the use of age categories). Our rationale for using KNN was rooted in a desire to assume as little as possible given the lack of prior work on this topic as well as sample size. Relatedly, although our sample size is rather large for pediatric neuroimaging, and is twice the median cell size of human fMRI studies (Poldrack et al., 2017), it is nevertheless somewhat small in comparison to other network neuroscience studies (e.g., Baum et al., 2017; DuPre & Spreng, 2017). Further, although cross-sectional research is an excellent starting point for characterizing the neurodevelopment of emotion regulation, future research would benefit in having longitudinal data to examine how the trajectory of each network's activity change across time, helping us better parse age and emotion regulation activity. Additionally, amid growing concerns about reliability (Elliott et al., 2019; Enkavi et al., 2019), we feel compelled to note that the behavioral and neural test-retest reliability of the emotion regulation task used here has not yet been established, let alone in a sample of youth. Addressing these concerns under further empirical scrutiny will help lend further confidence in the take-away conclusions raised by our results.

In sum, this study is the first to our knowledge that examines how the developing brain supports the acquisition of emotion regulation—a cornerstone of mental health—independent of chronological age. We show that age and emotion regulation abilities are separately encoded in patterns of network activity, while also highlighting that these networks fit into a broader organizational structure that can be mapped out using a taxonomical approach.

STAR * Methods

Participants. Participants for the current study were drawn from a larger sample of individuals enrolled in a longitudinal neuroimaging study about childhood maltreatment and affective neurodevelopment. Youth in the current analyses were selected from a non-maltreated community control group who were screened with the following exclusion criteria: childhood maltreatment/violence, presence of a developmental disorder (e.g., autism), psychotropic medication use, and IQ < 75. To qualify for inclusion in the current report, participants were required to have anatomical images free of abnormalities or artifacts, low levels of motion during the scan (addressed below), and accompanying behavioral data from our scanning emotion regulation paradigm. Our final sample included 70 youth (34 female) ranging in age from 8.08 to 17.00 years (Mean age = 12.7). The study was conducted in a large, metropolitan area in the Pacific Northwest region of the United States. Youth and their parents provided written assent and consent, respectively, in accordance with the University of Washington's Institutional Review Board. Participants were remunerated $75 after completion of the scan. Data and analysis code can be accessed on the Open Science Framework (osf.io/bfg3d).

fMRI Emotion Regulation Paradigm. Participants completed a common and widely used emotion regulation task while undergoing fMRI (Silvers, Shu, Hubbard, Weber, & Ochsner, 2015). Participants passively viewed or effortfully regulated emotional responses to a series of developmentally-appropriate negative and neutral images (modeled after the International Affective Picture System (IAPS); Lang, Bradley, & Cuthbert, 2008). Each trial began with an instructional cue (2s duration), 'Look' or 'Far'. The former instructs participants to passively view a subsequent stimulus whereas the latter prompts participants to regulate it via a

psychological distancing variant of cognitive reappraisal. Following the cue, participants viewed a neutral or aversive stimulus. Aversive stimuli were either paired with 'Look' or 'Far'. Neutral images were always paired with the 'Look' cue; these images are not of interest to the current study and will not be discussed further. Stimuli were displayed for 6-10s (jittered). A 5-point Likert scale was presented thereafter to collect participants affective responses (4s). Lastly, a jittered ITI concluded each trial (1.5-6.5s). Prior to scanning, participants were trained extensively by experimenters to ensure they understood the task and were comfortable completing it, and afterwards completed a practice run. Stimuli were purchased from a stock photography website (https://shutterstock.com) and were normalized in a sample of youth (120, ages 6-16, 50% female) based on valence, arousal, and dominance. Stimuli and normed ratings are available online (osf.io/43hfq/). The most negatively rated images were selected for use in the current study. Participants completed four runs, each lasting approximately 220s (rungs ranged between 110-115 4D volumes). Each task condition ('Look' – Negative; 'Far' – Negative; 'Look' – Neutral) was presented 20 times over the four runs (5 per run). The task was programmed and presented in E-Prime (Psychology Software Tools, Inc., http://www.pstnet.com). A visual representation of the task is depicted in Figure S1.

fMRI Data Acquisition. If participants were younger than 12 years of age or exhibited any nervousness about scanning, they were taken to a mock MRI scanner in order to familiarize with them with scanning environment and to be trained on how to minimize head motion in the scanner. Additionally, in preparation for the scan, participants were packed into the head coil with an inflated, head-stabilizing pillow to restrict movement.

Images were acquired at the University of Washington's (Seattle) Integrated Brain Imaging Center on a 3T Phillips Achieva scanner. A 32-channel head coil was used in conjunction with parallel image acquisition. A T1-weighted, magnetization-prepared rapid acquisition gradient echo (MPRAGE) image was acquired for registration purposes (TR=2530ms, TE=1640-7040µs, 7º flip angle, 256 mm^2 FOV, 176 slices, 1 mm^3 isotropic voxels). Blood oxygenation level dependent (BOLD) signal during functional runs was recorded using a gradient-echo T2*-weighted echoplanar imaging (EPI) sequence. Thirty-two 3-mm thick slices were acquired parallel to the AC-PC line (TR=2000ms, TE=30ms, 90º flip angle, 256 x 256 FOV, 64 x 64 matrix size).

fMRI Data Pre-Processing. The FMRIB Software Library package (FSL, version 5.0.9; fsl.fmrib.ox.ac.uk) was used for pre-processing and later analysis. Prior to pre-processing, data were visually inspected for artifacts, anatomical abnormalities, and excessive head motion. Pre-processing began by using the brain extraction tool (BET) to remove non-brain tissue from images and estimating the extent of participant head motion by using FSL Motion Outliers to record volumes that exceed a 0.9 mm framewise displacement (FD) threshold (Siegel et al., 2014). At this stage, runs were discarded if they exceed this threshold more than ~ 20% of their duration (i.e., if 0.9 mm FD was surpassed on more than 25 volumes). Participants needed to have at least one run in order to be included in analysis. Six participants had at least one run discarded; more information about head motion in this sample is reported elsewhere (Guassi Moreira et al., 2019). Afterwards, we corrected for head motion by spatially realigning volumes

using MCFLIRT and then hi-pass filtered the data (100s cutoff). We pre-whitened the data to correct for temporally autocorrelated residuals.

Network Definition. As previously mentioned, we chose six networks that have been implicated in studies of emotion regulation and related perceptual, cognitive, affective, and social processes. We defined the ROIs of each network in a two-step process. We (i) acquired meta-analytic maps for each network and then (ii) extracted local peaks from the clusters in each image to serve as our ROIs. Meta-analytic maps from NeuroSynth (retrieved April, 2018) were used for all networks except the CRN, which obtained its meta-analytic map from a published meta-analysis of fMRI studies of cognitive reappraisal (Buhle et al., 2014). This was done intentionally in order to increase specificity of the CRN network, since the term 'emotion regulation' encompasses many different, and often disparate, strategies that would not have been consistent with our own experimental paradigm. Spheres (4 mm radius) were drawn across all peaks in order to extract data necessary to compute connectivity (described in the directly preceding section). A 4mm radius was used to minimize overlap between spheres (see Guassi Moreira, et al., 2019). The final ROI counts for each network follow: CRN – 32, DMN – 26, FPN – 25, SaN – 14, DAN – 27, and VAN – 21. Importantly, the 145 ROIs were maximally distinct – no possible pairs of spheres overlapped entirely and partial overlap was minimal. ROIs for each network are visualized in Figure 1 and their constituent ROI coordinates and labels are listed in the supplement (Table S1).

Estimating Network Connectivity Profiles via Beta-Series. In order to obtain metrics of network-level activity, we first modeled the task data within-subjects using a Least Squares Single (Mumford, Davis, & Poldrack, 2014) design to yield a beta-series (Rissman, Gazzaley, & D'Esposito, 2004). For each run within an individual subject, a fixed-effects GLM was created for each 'Far' (regulation of aversive stimuli) and 'Look-Negative' (passive observation of aversive stimuli) 'target trial'. Thus, a separate GLM was created for each individual target trial. Each GLM consisted of a single-trial regressor which modeled that GLM's target trial, a nuisance regressor modeling all other events in the target trial's condition, regressors for the other two conditions—each modeled as they normal would in a standard univariate analysis—and motion parameter regressors. For example, the first 'Far' trial for a participant would receive its own GLM wherein it was modeled in a single regressor; all other 'Far' trials were placed in a separate nuisance regressor while 'Look-Negative' and 'Look-Neutral' trials modeled in their own respective regressors. Afterwards, a new GLM was created for the *second* 'Far' trial, where it was the 'target trial' and modeled in a single-event regressor and the *first* trial is now modeled in the separate nuisance regressor with other trials of that type. FSL's extended motion parameters estimated from MCFLIRT (rotation+translation, derivatives, and squares of each) and additional volumes that exceeded our FD = 0.9 mm threshold were also added as regressors of non-interest to each first-level LSS GLM.

Following estimation of the LSS GLMs, we used parameter estimates from each trial-specific GLM to create linear contrast images comparing both trial types of interest ('Far' & 'Look-Negative'), respectively, to baseline. We then extracted estimates of activation for both trial types across a series of ROIs nested within six different networks (see below). This yielded two *t*

x *p* x *n* arrays (one for each task condition), where entries to each cell represent the average parameter estimate at the *t*-th trial for the *p*-th ROI in the *n*-th network. For every network, beta-series amongst all its ROIs were correlated amongst each other (Spearman's Rho), yielding two connectivity matrices (one for each trial type) per network. These matrices describe each network's connectivity for a given task condition (i.e., connectivity matrix), and we refer to them as network connectivity profiles (See Figure 2 for an overview).

Using Representational Similarity Analysis to Quantify Network Differentiation. The first feature of network activity we chose to examine was differentiation. In order to estimate network differentiation, we compared profiles of network connectivity when youth were engaging in emotion regulation versus when they were simply viewing emotional stimuli using representational similarity analysis (RSA) on our two task-specific connectivity matrices (Kriegeskorte, 2008; Kriegeskorte et al., 2008). Traditionally, RSA is a form of multi-voxel pattern analysis that relies on the presumption that multi-voxel activation patterns reliably contain information about a specific stimulus (Etzel, Zacks, & Braver, 2013; Kriegeskorte, Simmons, Bellgowan, & Baker, 2009). Comparison of two stimulus patterns (via correlation) in RSA can reveal the extent to which representations of stimuli encode unique or similar information. However, one is not bound to use RSA solely with multivariate patterns. In fact, RSA can be used to compare any two different representations, as information about pattern representations are abstracted away from the modality in which they were first acquired, forced into a common space, and then compared. To this point, prior research has used RSA to compare voxelwise patterns to behavioral patterns (Parkinson, Kleinbaum, & Wheatley, 2017), behavioral patterns to other behavioral patterns (Brooks & Freeman, 2018; Stolier, Hehman, Keller, Walker, & Freeman, 2018), and, most importantly for our purposes, representations of brain connectivity patterns (Lee, Miernicki, & Telzer, 2017). We built upon and extended these prior implementations of RSA for our purposes in comparing network connectivity states between task conditions.

For each network and task condition, a *r* x *r* connectivity matrix was constructed (noted in the preceding section), where *r* is the number of ROIs in the network. Representational similarity analyses involved vectorizing the off-diagonal elements of each task condition's various connectivity matrices, Fisher transforming them (for variance stabilization), and then correlating the elements (Spearman). This yielded a scalar value that describes the extent to which a given network's connectivity profile differs between task conditions (see Figure 2). An extreme high value indicates that the pattern of connectivity across the entire network remains consistent during Far and Look-Negative trials; An extreme low value (anticorrelation) indicates a differentiated connectivity profile between the two task states.

Network Community Structure.

Community Number Differences Between Network States. The second feature of network activity we examined was community structure between task conditions (e.g., when regulating versus when passively viewing emotional stimuli) (Sporns, 2010). That is, we examined the extent to which nodes within a network formed communities, or subclusters comprised of densely interconnected nodes within the network. We were specifically interested in quantifying

the number of communities within each network as a function of task condition. To achieve this, we first used the aforementioned connectivity matrices to create undirected, weighted adjacency matrices. An adjacency matrix summarizes a graph—each row and column represent the ROIs (nodes) that comprise the network, while an entry into the off-diagonal elements of the matrix denotes the presence or absence of a connection (edge) between two nodes. We thresholded our connectivity matrices at a reasonably high threshold of $\rho > .5000$ (Bolt et al., 2017), meaning that a pair of ROIs whose beta-series shared a correlation of .4999 or lower were indicated as not being linked by an edge. ROI pairs whose correlations surpassed the threshold received an edge with a weight corresponding to the pair's Spearman correlation coefficient. We tested the robustness of our findings by re-running our analyses with new thresholds (.4000 and .6000, respectively)—our results largely did not depend on the choice of edge-defining threshold (see Supplement for details and results).

We then estimated the number of communities in each task condition, for all networks, for all subjects using the R igraph package's walktrap clustering algorithm (Csardi & Nepusz, 2006; Pons & Latapy, 2005) in R (R Core Team, 2014). The walktrap clustering algorithm (Pons & Latapy, 2005) estimates community structure in a graph using random walks, step-wise paces from one node to another on a graph according to random chance. The algorithm operates on the principle that the probability of taking a walk (step) from one node to another depends on the number of shared connections between nodes. This ultimately means that walks are more likely to become 'trapped' inside a community than a non-community due to the dense interconnections between nodes that define community membership. The number of a given network's communities across both task conditions were extracted from the graphs and differenced ('Far' – 'Look'), resulting in six metrics of community structure, one per network, per participant. A higher value on this metric indicates a greater number of observed communities within a network during emotion regulation, compared to passively viewing emotional stimuli.

Modularity Differences Between Network States. Our final network metric was the difference in modularity between task conditions for a given network. Building upon the concept of community structure, a graph is said to be highly modular if it exhibits relatively dense interconnections within its communities and relatively sparse connections between them. On the other hand, a graph that is low in modularity has nodes that have relatively equal connections within and between their communities. In other words, modularity can be thought to measure the extent to which "cross-talk" exists between the communities within a network (i.e., a ratio between intra- versus inter-community ties). Modularity metrics across both task states in each network across all participants were obtained using the communities identified in the prior section. The same differencing was taken ('Far' – 'Look') to obtain a metric of modularity differences between task conditions across each participant's connectome. Notably, because our focus is to understand the degree to which information about development and emotion regulation ability is encoded across a set of brain networks, we opted to focus on *within*-network modularity, rather than *between*-network modularity.

Taxonomizing Brain Networks Via Hierarchical Clustering. Computing the aforementioned metrics left us with 3 measurements per network (differentiation of network connectivity profiles, number of network communities, and network modularity) across 6 networks (CRN, DMN, FPN, SaN, DAN, VAN) per participant. Our next step was to hierarchically taxonomize the brain networks studied here with respect to activity during emotion regulation (Tversky & Hutchinson, 1986). To this end, we used an agglomerative clustering algorithm with R's `hclust()` function. Agglomerative clustering is a 'bottom-up' approach in which the algorithm begins by assigning each item to its own cluster and then iteratively groups items into larger, hierarchical clusters based on similarity of a specified feature. Separate instances of clustering were run based on each of our metrics (similarity, community, modularity), yielding three separate dendrograms. Cophenetic correlation was used to evaluate the consistency of clustering across the three dendrograms (Liu, Zhu, Qiu, & Chen, 2012). Lastly, we used resampling methods (bootstrapping) and cophenetic correlation to help evaluate whether our clustering results were obtained simply due to chance.

K-Nearest Neighbor Classification. Finally, we used machine learning to predict a youth's emotion regulation ability and age based on (i) their individual metrics across all networks and (ii) separate models within each network. Specifically, we employed a K-Nearest Neighbor (KNN) algorithm, which involves comparing a set of datapoints—in our case, metrics of whole-brain network activity—against its *K* nearest neighbors in multidimensional space (Euclidean space here) (VanderPlas, 2016). After finding a datapoint's *K* nearest neighbors, the class membership of each neighbor is tallied and the datapoint is classified as belonging to the class represented by the majority of its neighbors.

KNN was selected for several reasons. First, KNN is a non-parametric classifier that makes minimal assumptions about the data its classifying. This point is critical given there have been virtually no prior network neuroscience investigations of emotion regulation in youth. Second, KNN's non-parametric nature allows for meaningful results with relatively smaller samples. Lastly, because our investigation is the very first of its kind, KNN's parsimony makes it a natural fit at this stage of the scientific process because it allows us to explaining the data in the simplest manner before subsequently building complexity.

Here we use KNN in conjunction with leave-one-out cross-validation (LOO-CV) to determine whether signatures of network activity (network differentiation, community structure, and modularity) contain meaningful information about age and emotion regulation ability. In LOO-CV, each individual case serves as a validation set. That is, *n*-1 cases are used to predict the class of the *n*th case, and the procedure is repeated *n* times until each case has been used as the validation case. We specifically opted for LOO-CV over *k*-fold cross validation given the size of our dataset: *k*-fold cross validation typically requires larger samples, whereas LOO-CV estimation is effective in sample sizes comparable to ours (Kohavi, 1995). Importantly, the use of cross-validation helps safeguard against overfitting (though does not eliminate said risks).

Due to sample size constraints we opted to classify categorically, instead of using regression-based classifiers. We tested two class labels of interest: age and emotion regulation ability. For age, we split age participants into 'children' and 'teens'. We used three age cut-offs.

First, we split participants based on lay conceptions of adolescence and childhood (e.g., 13 years old). Second, we split participants into even groups based on the mean (12.5 years). Third, we split participants based on the mean age (12.7 years). The mean split was the most predictive split, followed by the median split, followed by lay conceptions cut-off (likely due to the fact that this latter split had the most unequal group sizes). Results were approximately comparable between splits and we report the most predictive results (mean split). Results with other cut-offs are reported in the supplement (Table S11). For ability, we split individuals into a 'high' emotion regulation ability group and a low emotion regulation ability group based on a cut-off of 20% in negative affect, which was the sample's mean ability score and is also consistent with prior work (Silvers et al., 2012). Emotion regulation ability scores were calculated by taking the percent change between a given participant's average rating of negative during 'Far' compared to 'Look' (Silvers et al., 2012) (high ability > 20%, the approximate average score; low ability < 20%). We tested an alternative threshold corresponding to the median emotion regulation ability score (19%) in order to achieve equal group sizes and the results did not differ appreciably (Table S12).

We then ran two separate sets of KNN models, one predicting age and the other predicting ability. Importantly, for both sets of labels, membership was approximately evenly distributed across categories (Age: 37 children, 33 teens; Ability: 34 high, 36 low). We first began by classifying both sets of labels (e.g., Age—Teen/Child, Ability—High/Low) from all 18 metrics of network activity with K set equal to 3, 5, or 7. We only used odd numbers in order to have a tiebreaking nearest neighbor; We stopped at 7 because it would have been less meaningful to use any larger number of neighbors; additional neighbors added from $K = 9$ or $K = 11$ would have comprised more than 10% of our sample, would have been farther away in Euclidean distance and thus, by our reasoning, less meaningful. We selected the best K value using a composite F_1 ratio. Briefly, such ratios evaluate the overall accuracy of a model for predicting a given class (e.g., children) by considering information about *precision* (true positives / [true positives + false positives]) and *recall* (true positives / [true positives + false negatives]). Bound between 0 and 1, a greater F_1 value indicates more overall accuracy. The calculation of composite F_1 ratio is detailed in the results. Lastly, using the best K (according to F_1 ratios, described in the results), we tested whether we could recover the same amount of information with individual networks as opposed to looking across networks. To this end, we trained an additional six models per class (age and ability) using only metrics from each individual network as features and compared these models to the cross-network measure. Leave-one-out cross validated KNN classification was run in R using the class package's knn.cv()function. Ninety-five percent confidence intervals for classification accuracies were derived from the binomial distribution using the binom.test()function in R.

References

Aldao, A., Sheppes, G., & Gross, J. J. (2015). Emotion Regulation Flexibility. *Cognitive Therapy and Research*, *39*(3), 263–278. https://doi.org/10.1007/s10608-014-9662-4

Baum, G. L., Ciric, R., Roalf, D. R., Gur, R. C., Bassett, D. S., Satterthwaite, T. D., … Satterthwaite, T. D. (2017). Modular segregation of structural brain networks supports the development of executive function. *Current Biology*, *27*(11), 1561-1572.e8. https://doi.org/10.1016/j.cub.2017.04.051

Blakemore, S.-J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology*, *65*, 187–207. https://doi.org/10.1146/annurev-psych-010213-115202

Bolt, T., Nomi, J. S., Yeo, B. T. T., & Uddin, L. Q. (2017). Data-driven extraction of a nested model of human brain function. *Journal of Neuroscience*, *0323–17*. https://doi.org/10.1523/JNEUROSCI.0323-17.2017

Braunstein, L. M., Gross, J. J., & Ochsner, K. N. (2017). Explicit and implicit emotion regulation: A multi-level framework. *Social Cognitive and Affective Neuroscience*, *ePub ahead*, 1–13. https://doi.org/10.1080/02699931.2010.544160

Brody, G. H., Yu, T., Nusslock, R., Barton, A. W., Miller, G. E., Chen, E., … Sweet, L. H. (2019). The protective effects of supportive parenting on the relationship between adolescent poverty and resting-state functional brain connectivity during adulthood. *Psychological Science*, 1–9. https://doi.org/10.1177/0956797619847989

Brooks, J. A., & Freeman, J. B. (2018). Conceptual knowledge predicts the representational structure of facial emotion perception. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-018-0376-6

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*(2), 49–57. https://doi.org/10.1016/j.tics.2006.11.004

Buhle, J. T., Silvers, J. A., Wage, T. D., Lopez, R., Onyemekwu, C., Kober, H., … Ochsner, K. N. (2014). Cognitive reappraisal of emotion: A meta-analysis of human neuroimaging studies. *Cerebral Cortex*, *24*(11), 2981–2990. https://doi.org/10.1093/cercor/bht154

Calkins, S. (1994). Origins and outcomes of individual differences in emotion regulation. *Monographs of the Society for Research in Child Development*, *59*(2–3), 53–72.

Casey, B. J., Jones, R. M., & Hare, T. A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences*, *1124*, 111–126. https://doi.org/10.1196/annals.1440.010

Casey, B. J., Jones, R. M., Levita, L., Libby, V., Pattwell, S., Ruberry, E., … Somerville, L. H. (2010). The storm and stress of adolescence: Insights from human imaging and mouse genetics. *Developmental Psychobiology*, *52*(3), 225–235. https://doi.org/10.1002/dev.20447.The

Collin, G., & van den Heuvel, M. P. (2013). The ontogeny of the human connectome:

Development and dynamic changes of brain connectivity across the life span. *The Neuroscientist*, *19*(6), 616–628.

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nature Reviews Neuroscience*, *13*(9), 636–650. https://doi.org/10.1038/nrn3313

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, *Complex Sy*, 1695.

Demirtas, M., Burt, J. B., Helmer, M., Ji, J. L., Adkinson, B. D., Glasser, M. F., … Murray, J. D. (2019). Hierarchical heterogeneity across human cortex shapes large-scale neural dynamics. *Neuron*, *101*, 1181–1194. https://doi.org/10.1016/j.neuron.2019.01.017

Doremus-Fitzwater, T. L., Varlinskaya, E. I., & Spear, L. P. (2010). Motivational systems in adolescence: Possible implications for age differences in substance abuse and other risk-taking behaviors. *Brain and Cognition*, *72*(1), 1–20. https://doi.org/10.1016/j.bandc.2009.08.008.Motivational

DuPre, E., & Spreng, R. N. (2017). Structural covariance networks across the lifespan, from 6-94 years of age. *Network Neuroscience*, 1–38. https://doi.org/10.1162/NETN_a_00016

Eisenberg, I., Bissett, P., Enkavi, A. Z., Li, J., MacKinnon, D., Marsch, L., & Poldrack, R. (2018). Uncovering mental structure through data-driven ontology discovery. *PsyArXiv*, 1–10. https://doi.org/10.31234/OSF.IO/FVQEJ

Elliott, M. L., Knodt, A. R., Ireland, D., Morris, M. L., Ramrakha, S., Sison, M. L., … Hariri, A. R. (2019). Poor test-retest reliability of task-fMRI: New empirical evidence and a meta-analysis. *BioRxiv*, (919), 1–31.

Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., Mackinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Large-scale analysis of test – retest reliabilities of self-regulation measures. *Proceedings of the National Academy of Sciences*, *116*(12), 5472–5477. https://doi.org/10.1073/pnas.1818430116

Ernst, M., Pine, D. S., & Hardin, M. (2006). Triadic model of the neurobiology of motivated behavior in adolescence. *Psychological Medicine*, *36*(3), 299–312. https://doi.org/10.1017/S0033291705005891.Triadic

Etzel, J. A., Zacks, J. M., & Braver, T. S. (2013). Searchlight analysis: Promise, pitfalls, and potential. *NeuroImage*, *78*, 261–269. https://doi.org/10.1016/j.actbio.2009.04.013.Role

Feldman Barrett, L., & Satpute, A. (2013). Large-scale brain networks in affective and social neuroscience: Towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology*, *23*(3), 361–372. https://doi.org/10.1016/j.conb.2012.12.012.Large-scale

Fuligni, A. J. (2018). The need to contribute during adolescence. *Perspectives on Psychological Science*, 1–13. https://doi.org/10.1177/1745691618805437

Gao, W., Alcauter, S., Smith, K., Gilmore, J. H., & Lin, W. (2015). Development of human brain cortical network architecture during infancy. *Brain Structure and Function*, *220*(2), 1173–1186.

Gee, D. G., Gabard-Durnam, L., Telzer, E. H., Humphreys, K. L., Goff, B., Shapiro, M., … Tottenham, N. (2014). Maternal buffering of human amygdala-prefrontal circuitry during childhood but not during adolescence. *Psychological Science*, *25*(11), 2067–2078. https://doi.org/10.1177/0956797614550878

Gershman, S. J., Monfils, M. H., Norman, K. A., & Niv, Y. (2017). The computational nature of memory modification. *ELife*, *6*, 1–41. https://doi.org/10.7554/eLife.23763

Gratton, C., Laumann, T. O., Nielsen, A. N., Greene, D. J., Gordon, E. M., Gilmore, A. W., … Petersen, S. E. (2018). Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation. *Neuron*, *98*(2), 439-452.e5. https://doi.org/10.1016/j.neuron.2018.03.035

Greene, A. S., Gao, S., Scheinost, D., & Constable, R. T. (2018). Task-induced brain state manipulation improves prediction of individual traits. *Nature Communications*, *9*(1), 1–13. https://doi.org/10.1038/s41467-018-04920-3

Gross, J. J. (1998). The emerging field of emotion regulation: an integrative review. *Review of General Psychology*, *2*(5), 271–299. https://doi.org/10.1017.S0048577201393198

Guassi Moreira, J. F., McLaughlin, K. A., & Silvers, J. A. (2019). Spatial and temporal cortical variability track with age and affective experience during emotion regulation in youth. *Developmental Psychology*, 1–47.

Hauser, T. U., Iannaccone, R., Walitza, S., Brandeis, D., & Brem, S. (2015). Cognitive flexibility in adolescence: Neural and behavioral mechanisms of reward prediction error processing in adaptive decision making during development. *NeuroImage*, *104*, 347–354. https://doi.org/10.1016/j.neuroimage.2014.09.018

Kaufmann, T., Alnæs, D., Doan, N. T., Brandt, C. L., Andreassen, O. A., & Westlye, L. T. (2017). Delayed stabilization and individualization in connectome development are related to psychiatric disorders. *Nature Neuroscience*, *20*(4), 513–515. https://doi.org/10.1038/nn.4511

Kim, P., Evans, G. W., Angstadt, M., Ho, S. S., Sripada, C. S., Swain, J. E., … Phan, K. L. (2013). Effects of childhood poverty and chronic stress on emotion regulatory brain function in adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 18442–18447. https://doi.org/10.1073/pnas.1308240110

Kleckner, I. R., Zhang, J., Touroutoglou, A., Chanes, L., Xia, C., Simmons, W. K., … Feldman Barrett, L. (2017). Evidence for a large-scale brain system supporting allostasis and interoception in humans. *Nature Human Behaviour*, *1*(5). https://doi.org/10.1038/s41562-017-0069

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *International Joint Conference on Artificial Intelligence*.

Kriegeskorte, N. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 1–28. https://doi.org/10.3389/neuro.06.004.2008

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., … Bandettini, P. A.

(2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, *60*(6), 1126–1141. https://doi.org/10.1016/j.neuron.2008.10.043

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience*, *12*(5), 535–540. https://doi.org/10.1038/nn.2303

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8.* https://doi.org/10.1016/j.epsr.2006.03.016

Larson, R., & Lampman-Petraitis, C. (1989). Daily emotional states as reported by children and adolescents. *Child Development*, *60*(5), 1250–1260.

Lee, T. H., Miernicki, M. E., & Telzer, E. H. (2017). Families that fire together smile together: Resting state connectome similarity and daily emotional synchrony in parent-child dyads. *NeuroImage*, *152*, 31–37. https://doi.org/10.1016/j.neuroimage.2017.02.078

Linnæus, C. (1758). *Systema Naturae*. Gmelin.

Liu, X., Zhu, X.-H., Qiu, P., & Chen, W. (2012). A correlation-matrix-based hierarchical clustering method for functional connectivity analysis. *Journal of Neuroscience Methods*, *211*(1), 94–102. https://doi.org/10.1016/j.jneumeth.2012.08.016.A

Marek, S., & Dosenbach, N. U. F. (2018). The frontoparietal network: function, electrophysiology, and importance of individual precision mapping. *Dialogues in Clinical Neuroscience*, *20*, 133–140. Retrieved from www.dialogues-cns.org

Martin, L. J., Hathaway, G., Isbester, K., Mirali, S., Acland, E. L., Niederstrasser, N., … Mogil, J. S. (2015). Reducing social stress elicits emotional contagion of pain in mous and human strangers. *Current Biology*, *25*, 326–332. https://doi.org/10.1016/j.cub.2014.11.028

McRae, K., Gross, J. J., Weber, J., Robertson, E. R., Sokol-Hessner, P., Ray, R. D., … Ochsner, K. N. (2012). The development of emotion regulation: An fMRI study of cognitive reappraisal in children, adolescents and young adults. *Social Cognitive and Affective Neuroscience*, *7*(1), 11–22. https://doi.org/10.1093/scan/nsr093

Morgan, S. E., White, S. R., Bullmore, E. T., & Vértes, P. E. (2018). A network neuroscience approach to typical and atypical brain development. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 1–13. https://doi.org/10.1016/j.bpsc.2018.03.003

Mumford, J. A., Davis, T., & Poldrack, R. A. (2014). The impact of study design on pattern estimation for single-trial multivariate pattern analysis. *NeuroImage*, *103*, 130–138. https://doi.org/10.1016/j.neuroimage.2014.09.026

Ochsner, K. N., Silvers, J. A., & Buhle, J. T. (2012). Review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences*, (1251), E1–E24. https://doi.org/10.1111/j.1749-6632.2012.06751.x.Functional

Parkinson, C., Kleinbaum, A. M., & Wheatley, T. (2017). Spontaneous neural encoding of social network position. *Nature Human Behaviour*, *1*(5), 1–7. https://doi.org/10.1038/s41562-017-0072

Perino, M. T., Miernicki, M. E., & Telzer, E. H. (2016). Letting the good times roll: adolescence as a period of reduced inhibition to appetitive social cues. *Social Cognitive and Affective Neuroscience*, *11*(11), 1762–1771. https://doi.org/10.1093/scan/nsw096

Pfeifer, J. H., Masten, C. L., Borofsky, L. A., Dapretto, M., Fuligni, A. J., & Lieberman, M. D. (2009). Neural correlates of direct and reflected self-appraisals in adolescents and adults: When social perspective-taking informs self-perception. *Child Development*, *80*(4), 1016–1038. https://doi.org/10.1038/jid.2014.371

Pitskel, N. B., Bolling, D. Z., Kaiser, M. D., Crowley, M. J., & Pelphrey, K. A. (2011). How grossed out are you? The neural bases of emotion regulation from childhood to adolescence. *Developmental Cognitive Neuroscience*, *1*(3), 324–337. https://doi.org/10.1016/j.dcn.2011.03.004

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M., … Yarkoni, T. (2017). Scanning the horizon: Towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience*, *18*, 115–126. https://doi.org/http://dx.doi.org/10.1101/059188

Pollak, S. D., Messner, M., Kistler, D. J., & Cohn, J. F. (2009). Development of perceptual expertise in emotion regulation. *Cognition*, *110*(2), 242–247.

Pons, P., & Latapy, M. (2005). Computing communities in large networks using random walks. In *International Symposium on Computer and Information Sciences* (pp. 284–293). Berlin, Heidelberg: Springer.

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., … Petersen, S. E. (2011). Functional network rrganization of the human brain. *Neuron*, *72*(4), 665–678. https://doi.org/10.1016/j.neuron.2011.09.006

Richardson, H., Lisandrelli, G., Riobueno-Naylor, A., & Saxe, R. (2018). Development of the social brain from age three to twelve years. *Nature Communications*, *9*(1), 1027. https://doi.org/10.1038/s41467-018-03399-2

Rissman, J., Gazzaley, A., & D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage*, *23*(2), 752–763. https://doi.org/10.1016/j.neuroimage.2004.06.035

Robinson, P. N. (2012). Deep phenotyping for precision medicine. *Human Mutation*, *33*(5), 777–780. https://doi.org/10.1002/humu.22080

Rudolph, M. D., Miranda-Domínguez, O., Cohen, A. O., Breiner, K., Steinberg, L., Bonnie, R. J., … Fair, D. A. (2017). At risk of being risky: The relationship between "brain age" under emotional states and risk preference. *Developmental Cognitive Neuroscience*, *24*, 93–106. https://doi.org/10.1016/j.dcn.2017.01.010

Siegel, E. H., Wormwood, J. B., Quigley, K. S., & Feldman Barrett, L. (2018). Seeing what you feel: Affect drives visual perception of structurally neutral faces. *Psychological Science*, *29*(4), 496–503.

Silvers, J. A., Insel, C., Powers, A., Franz, P., Helion, C., Martin, R. E., … Ochsner, K. N. (2017). vlPFC – vmPFC – amygdala interactions underlie age-related fifferences in

cognitive regulation of emotion. *Cerebral Cortex*, *27*(7), 3502–3514. https://doi.org/10.1093/cercor/bhw073

Silvers, J. A., McRae, K., Gabrieli, J. D. E., Gross, J. J., Remy, K. A., & Ochsner, K. N. (2012). Age-related differences in emotional reactivity, regulation, and rejection sensitivity in adolescence. *Emotion*, *12*(6), 1235–1247. https://doi.org/10.1037/a0028297.Age-Related

Silvers, J. A., Shu, J., Hubbard, A. D., Weber, J., & Ochsner, K. N. (2015). Concurrent and lasting effects of emotion regulation on amygdala response in adolescence and young adulthood. *Developmental Science*, *18*(5), 771–784. https://doi.org/10.1111/desc.12260

Sizemore, A. E., Giusti, C., Kahn, A., Vettel, J. M., Betzel, R. F., & Bassett, D. S. (2018). Cliques and cavities in the human connectome. *Journal of Computational Neuroscience*, *44*, 115–145.

Spear, L. P. (2011). Rewards, aversions, and affect in adolescence: Emerging convergences across laboratory and human data. *Developmental Cognitive Neuroscience*, *1*, 390–403. https://doi.org/10.1016/j.dcn.2011.08.001

Sporns, O. (2010). *Networks of the Brain*. MIT Press.

Stolier, R. M., Hehman, E., Keller, M. D., Walker, M., & Freeman, J. B. (2018). The conceptual structure of face impressions. *Proceedings of the National Academy of Sciences*. https://doi.org/10.31234/osf.io/d9ha7

Tamietto, M., & De Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience*, *11*(10), 697.

Thompson, R., Lewis, M. D., & Calkins, S. (2008). Reassessing Emotion Regulation. *Child Development Perspectives*, *2*(3), 124–131.

Tversky, A., & Hutchinson, J. W. (1986). Nearest neighbor analysis of psychological spaces. *Psychological Review*, *93*(1), 3–22.

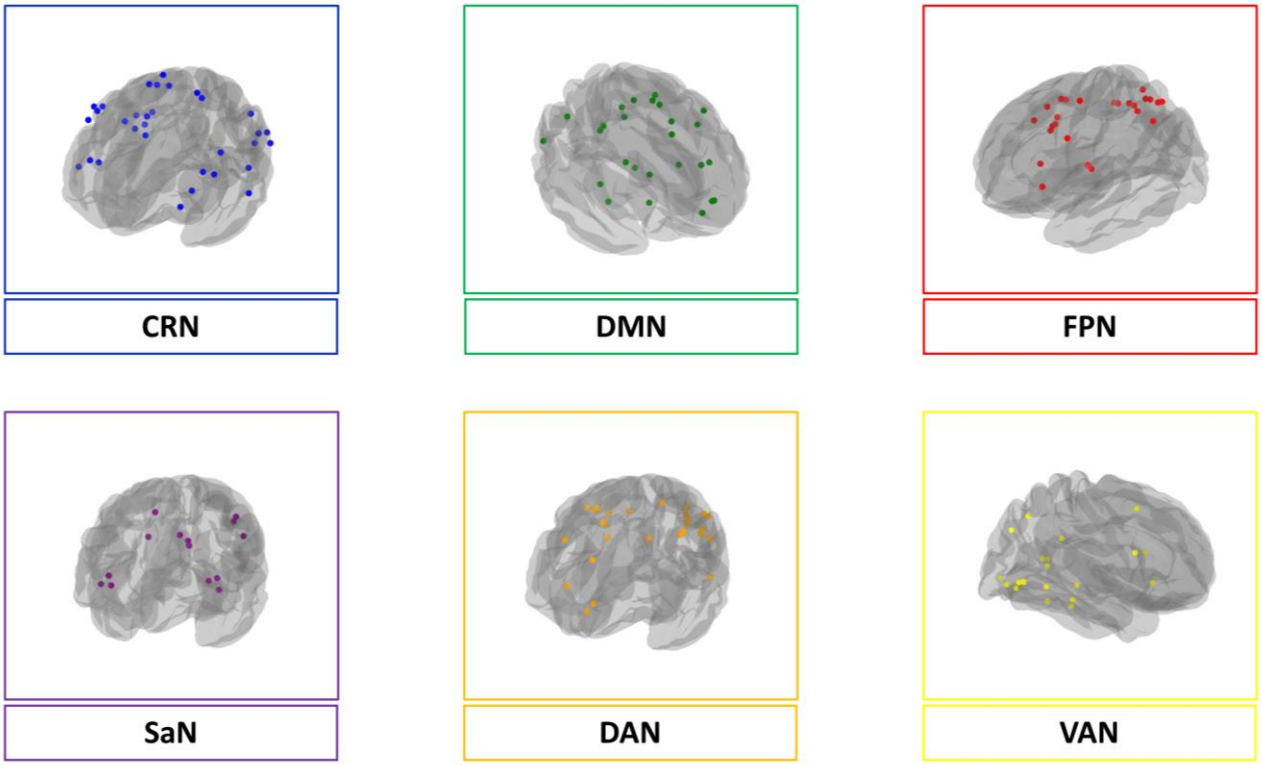
VanderPlas, J. (2016). *Python data science handbook: Essential tools for working with data*. O'Reilly Media, Inc.

Wagner, W. H. (1950). *The Fern Genus Diellia: Structure, Affinities, and Taxonomy*. University of California, Berkeley.

Yeo, B. T. T., Krienen, F. M., Eickhoff, S. B., Yaakub, S. N., Fox, P. T., & Buckner, R. L. (2015). Functional Specialization and Flexibility in Human Association Cortex. *Cerebral Cortex*, *25*(10), 3654–3672. https://doi.org/10.1093/cercor/bhu217

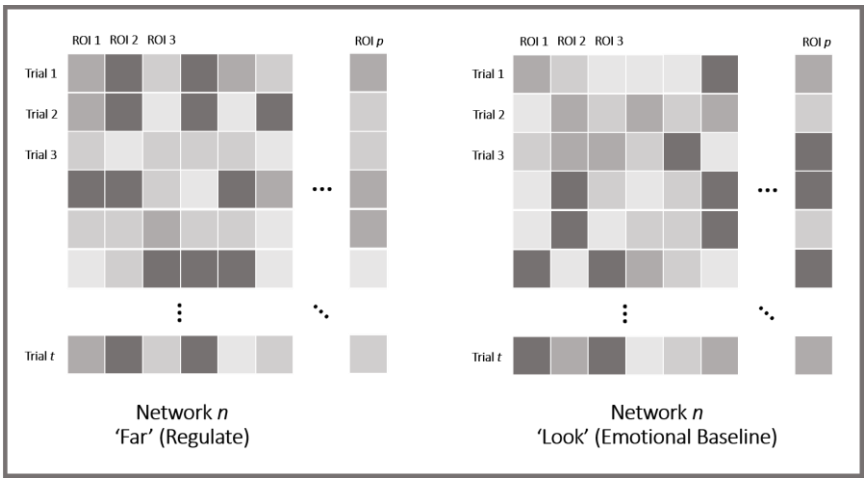
Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., … Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*, 1125–1165. https://doi.org/10.1152/jn.00338.2011.

Figure 1. Meta-analytic whole-brain networks and their parcellations.

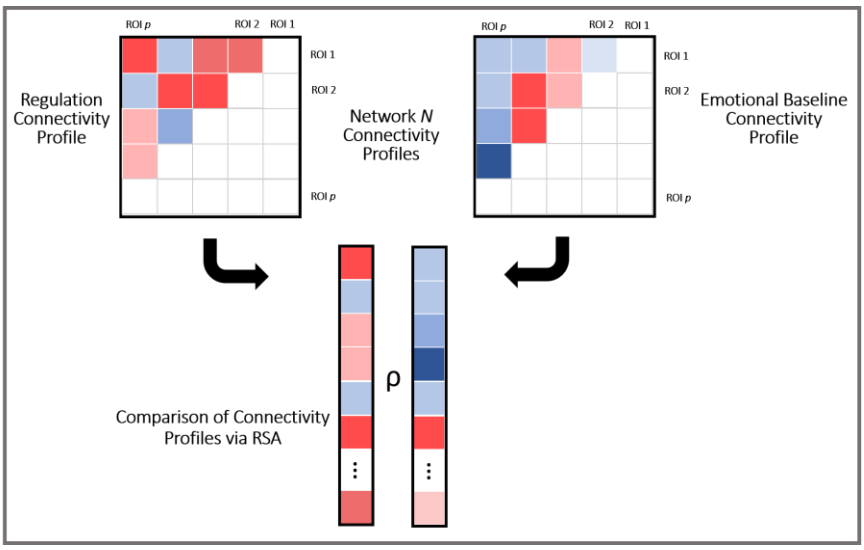


Note. CRN = Cognitive Reappraisal Network; DMN = Default Mode Network; FPN = Frontoparietal Network; SaN = Salience Network; DAN = Dorsal Attention Network; VAN = Ventral Attention Network.

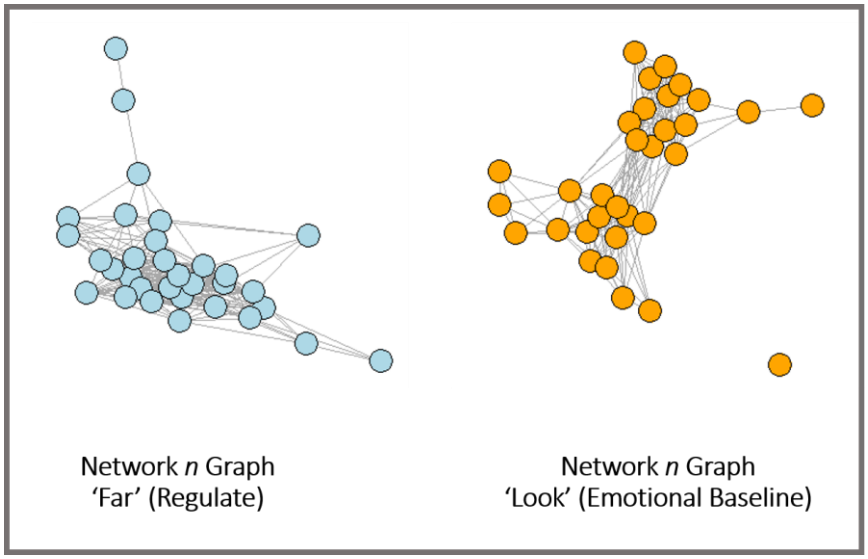
Figure 2. Overview of network activity metrics estimation.



Network Connectivity via Beta-Series



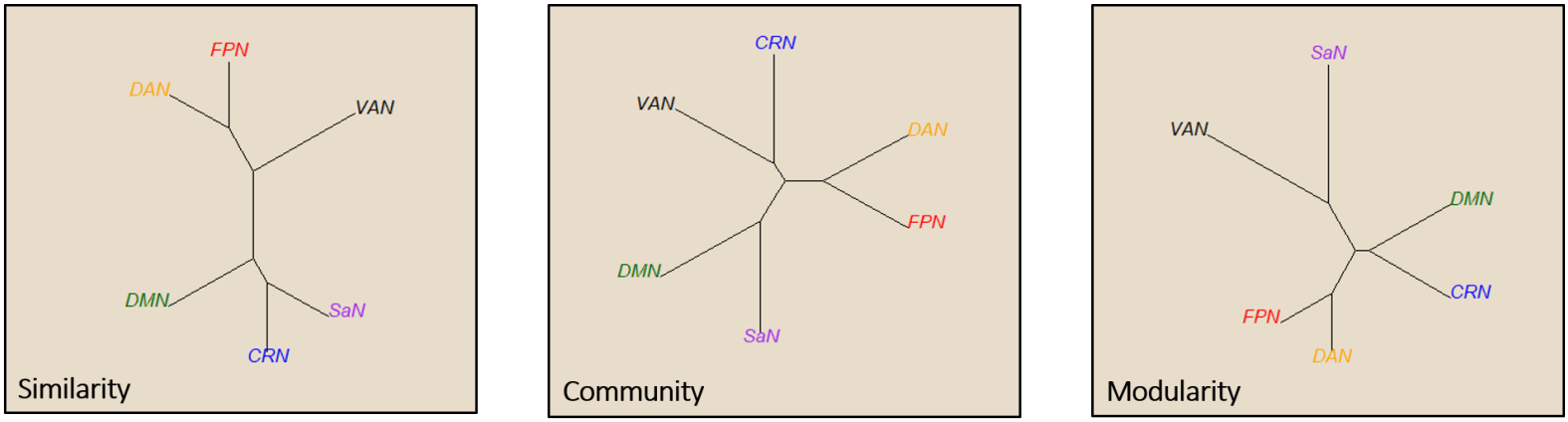
Network Differentiation via RSA



Network Community Structure via Graph Theory

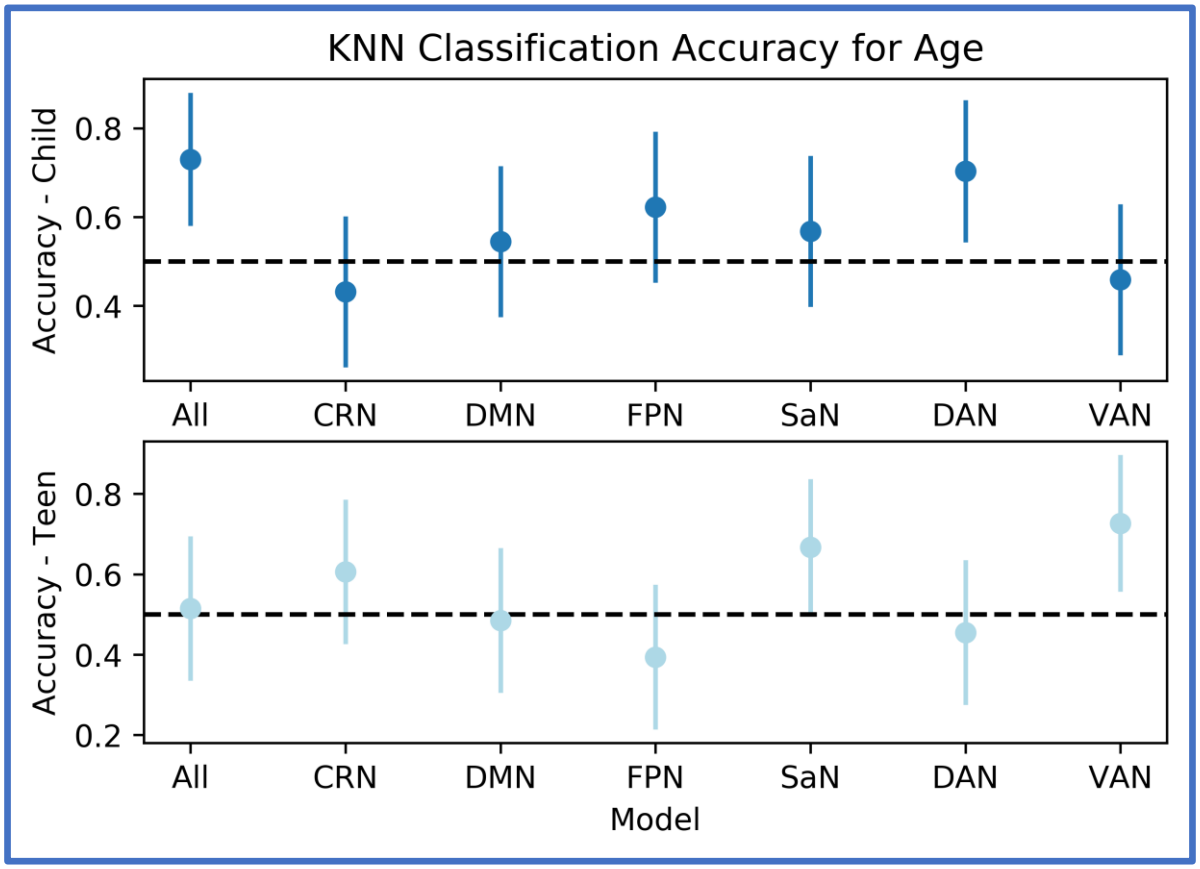
Note. Top panel (Network Connectivity via Beta-Series) depicts the beta-series estimated for each task condition (regulate ('Far'), emotional baseline ('Look')), for the *n*-th network. Each column corresponds to the *p*-th ROI, each row to the *t*-th trial, and each entry is the average beta-value for a given ROI at a given trial. Middle panel (Network Differentiation via RSA) depicts process of computing network differentiation. The two boxes represent connectivity profiles (matrices) for each condition for the *n*th network. Each entry to the matrix is the connectivity estimate between a given pair of ROIs. Unique off-diagonal elements were vectorized and then correlated (Spearman's Rho) to yield a single, scalar value of network connectivity similarity. Bottom panel (Network Community Structure via Graph Theory) depicts graphs for each task condition for the *n*-th network; each node represents an ROI within a given network, whereas each edge is the connectivity between a pair of ROIs. This step allowed us to compute the number of network communities (i.e., number of modules/communities in each network) and modularity. Network connectivity matrices were thresholded to create adjacency matrices and subsequently graphs. Estimates of community structure and modularity were compared across both task conditions for each network.

Figure 3. Network taxonomies defined by similarity, community number and modularity.



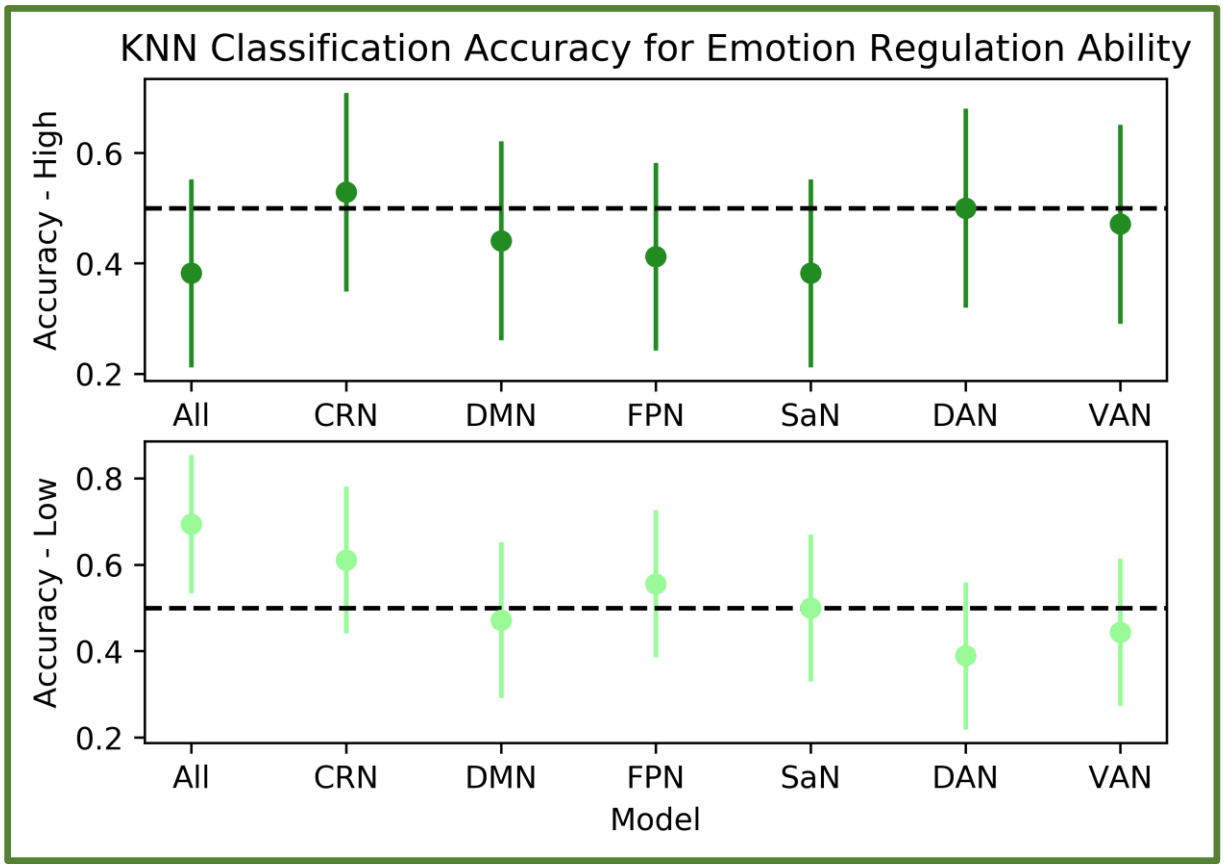
Note. The nearer two networks are to one another in the diagram, the more closely related they are along the metric listed in the bottom left corner of each figure (a la agglomerative clustering). Branches indicate splits in the taxonomy's nested hierarchy. CRN = Cognitive Reappraisal Network; DMN = Default Mode Network; SaN = Salience Network; DAN = Dorsal Attention Network; VAN = Ventral Attention Network. 'Community' refers to community number.

Figure 4. Classification accuracy among best fitting overall model and network-only models, broken down by age.



Note. All = Overall Model; CRN = Cognitive Reappraisal Network; DMN = Default Mode Network; SaN = Salience Network; DAN = Dorsal Attention Network; VAN = Ventral Attention Network. The dotted line indicates chance performance; 95% confidence interval depicted.

Figure 5. Classification accuracy among best fitting overall model and network-only models, broken down by low and high emotion regulation ability.



Note. All = Overall Model; CRN = Cognitive Reappraisal Network; DMN = Default Mode Network; SaN = Salience Network; DAN = Dorsal Attention Network; VAN = Ventral Attention Network. The dotted line indicates chance performance; 95% confidence interval depicted.

Table 1. Correlations between participant age and ability, and three network metrics between task states.

Network	*Age - Differentiation*	*Ability - Differentiation*	*Age - Modularity*	*Ability - Modularity*	*Age - Community*	*Ability - Community*
CRN	0.09 [-.14, .32]	-0.19 [-.46, .11]	0.02 [-.18, .22]	0.16 [-.12, .42]	-0.18 [-.44, .11]	**-0.30** **[-.55, .00]**
DMN	0.05 [-.22, .32]	0.05 [-.21, .30]	-0.01 [-.24, .23]	0.16 [-.14, .43]	0.12 [-.17, .39]	0.00 [-.20, .19]
FPN	-0.06 [-.33, .22]	-0.04 [-.31, .23]	0.00 [-.19, .20]	-0.09 [-.37, .21]	0.16 [-.14, .44]	-0.03 [-.28, .23]
SaN	0.09 [-.20, .36]	0.00 [-.20, .20]	-0.01 [-.25, .22]	0.02 [-.23, .28]	0.04 [-.23, .31]	0.16 [-.13, .43]
DAN	0.12 [-.14, .36]	-0.08 [-.31, .16]	-0.05 [-.25, .15]	-0.15 [-.41, .13]	**0.36** **[.07, .60]**	0.20 [-.09, .47]
VAN	-0.02 [-.28, .23]	0.02 [-.22, .25]	-0.03 [-.30, .23]	-0.17 [-.44, .14]	-0.01 [-.21, .19]	0.05 [-.23, .32]

Note. CRN = Cognitive Reappraisal Network; DMN = Default Mode Network; SaN = Salience Network; DAN = Dorsal Attention Network; VAN = Ventral Attention Network. Correlation coefficients are reported along with 90% confidence intervals. Correlations whose CIs do not include zero are bolded.

Table 2. Classification accuracy of KNN models predicting age from network level metrics of brain activity.

Model	*Precision - Child*	*Recall - Child*	*F_1 - Child*	*Precision - Teen*	*Recall - Teen*	*F_1 - Teen*	*F_1 - Composite*
A1. All Networks, *k* = 3	*0.757*	0.622	*0.683*	0.485	0.640	0.552	0.610
A2. All Networks, *k* = 5	0.730	0.623	0.675	*0.515*	0.630	*0.567*	**0.616**
A3. All Networks, *k* = 7	0.757	0.583	0.659	0.394	0.591	0.473	0.550
A4. CRN, *k* = 5	0.432	0.552	0.485	0.606	0.488	0.541	0.511
A5. DMN, *k* = 5	0.545	0.545	0.545	0.485	0.485	0.485	0.511
A6. FPN, *k* = 5	0.622	0.535	0.575	0.394	0.481	0.433	0.494
A7. SaN, *k* = 5	0.568	0.656	0.609	0.667	0.579	0.620	**0.614**
A8. DAN, *k* = 5	*0.703*	0.591	*0.642*	0.455	0.577	0.508	0.567
A9. VAN, *k* = 5	0.459	0.654	0.540	*0.727*	0.545	*0.623*	0.579

Note. Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores). The horizontal line subdivides two types of model. Those above the line use information from all networks; those below the line are network-specific.

Table 3. Classification accuracy of KNN models predicting emotion regulation ability from network level metrics of brain activity.

Model	*Precision – High*	*Recall - High*	*F_1 - High*	*Precision - Low*	*Recall - Low*	*F_1 - Low*	*F_1 - Composite*
C1. All Networks, *k* = 3	*0.382*	0.542	*0.448*	0.694	0.543	0.610	**0.517**
C2. All Networks, *k* = 5	0.265	0.500	0.346	0.750	0.519	0.614	0.443
C3. All Networks, *k* = 7	0.265	0.529	0.353	*0.778*	0.528	*0.629*	0.452
C4. CRN, *k* = 3	*0.529*	0.563	*0.545*	*0.611*	0.579	*0.595*	**0.569**
C5. DMN, *k* = 3	0.441	0.441	0.441	0.472	0.472	0.472	0.456
C6. FPN, *k* = 3	0.412	0.467	0.438	0.556	0.500	0.526	0.478
C7. SaN, *k* = 3	0.382	0.419	0.400	0.500	0.462	0.480	0.436
C8. DAN, *k* = 3	0.500	0.436	0.466	0.389	0.452	0.418	0.441
C9. VAN, *k* = 3	0.471	0.444	0.457	0.444	0.471	0.457	0.457

Note. Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores). The horizontal line subdivides two types of model. Those above the line use information from all networks; those below the line are network-specific.

Varying Edge Defining Thresholds. As noted in the main document, we aimed to replicate our results using different edge defining thresholds. The main results are reported using an edge defining threshold of r = .5000. Here we re-ran our analyses using two alternate thresholds, r = .4000 and .6000. First, we attempted to replicate our bivariate correlations between age, ability, and network activity metrics for the six networks. At a .4000 threshold, the association between emotion regulation ability and CRN community differences remained significant (Supplementary Table 2); this was not the case at .6000 threshold (Supplementary Table 7). At both the .4000 and.6000 threshold, the association between emotion regulation ability and DAN community were not significant, although the magnitude of their associations remained significant. Other non-significant bivariate correlations largely retained their effect sizes (i.e., the magnitudes of correlations were mostly the same across the different thresholds). This point is notable because this study is exploratory in nature, and small but reliable effects may exist but are difficult to detect with certain sample sizes.

Next, we examined whether our network taxonomizations were affected by choices in thresholds. At the .4000 edge defining threshold, the network taxonomy based on community was noticeably different from the analyses reported in the main document (See Supplementary Figure 3A, left panel). The bootstrapped cophenetic correlation coefficient is also quite different, but remains insignificant, just like in the main document (c = .376, CI [-.325, .980]). The taxonomy based on modularity had minor differences in the dendrogram (DAN, FPN, DMN, and VAN clustered together as opposed to DAN, FPN, DMN, and CRN), compared to the main results (See Supplementary Figure 3A, right panel), but the results were statistically consistent (c = .706, CI [.227, .991]). In terms of pairwise correspondence between dendrograms, again only community and modularity were significantly similar (c = .740, CI [.077, .994]; Similarity and community: c = -.069, CI [-.368, .512]; Similarity and modularity: c = .35, -.267, .919). At the .6000 edge defining threshold, the community dendrogram was more similar to the analyses in the main document, but remained not statistically significant (c = .299, CI [-.287, .943]) (See Supplementary Figure 3B, left panel). The dendrogram based on modularity was nearly identical to the one reported in the main document and was also statistically significant (c = .749, CI [.293, .988]) (See Supplementary Figure 3B, right panel). In terms of pairwise correspondence between dendrograms, again only community and modularity were significantly similar (c = .740, CI [.077, .995]; Similarity and community: c = .100, CI [-.355, .878]; Similarity and modularity: c = .282, CI [-.162, .886]).

Last, we examined to see how changing the edge defining threshold affected our classifier analyses. We first examined the analyses at a new edge defining threshold of .4000. In terms of age, we find that k = 5 still yielded the best model, although the composite F_1 composite score was much lower. Despite this, the model was still better than chance at predicting children than adolescents. The SaN remained the individual network with the best F_1 composite (Supplementary Tables 5A & 5B). When examining prediction of emotion regulation abilities, k = 3 again yielded the best model and the composite F_1 score was comparable to that in the main document. Once again, the classifier was better able to predict individuals with low scores rather than ones with high scores. The best individual network predictor was the FPN and not the CRN (although the CRN still faired relatively well compared to other networks) (Supplementary Table

6A & 6B). Afterwards we re-ran the results at a new edge defining threshold of .6000. In terms of age, we find that $k = 5$ still yielded the best model, although the composite F_1 composite score was much lower. Despite this, the model was again far better at predicting children than adolescents (the former well above chance, the latter well below). This time the FPN was the individual network with the best F_1 composite (Supplementary Tables 8A & 8B). In terms of ability, $k = 3$ again yielded the best model and exhibited a comparable F_1 score. The model was again better at classifying individuals with low emotion regulation abilities compared to those with high abilities. The CRN was again the best individual network predictor (with a F_1 composite score that was much higher than the one reported in the main document) (Supplementary Tables 9A & 9B).

Functional Connectivity Analysis and Network Verification. We extracted beta-series from every ROI across all six networks and subsequently correlated them as a metric of functional connectivity. An overall functional connectivity matrix and network specific connectivity matrices are displayed in Supplemental Figure 2. Glancing at the overall network connectivity matrix suggests that our networks were adequately defined, while also highlight that each network contained intriguing substructures within them as well as seemingly meaningful 'points of contact' with other ROIs belonging to other networks. To further ensure our network definition was sound, we took our overall functional connectivity matrix and submitted it to the walktrap clustering algorithm to see if we could faithfully re-capture our networks through community detection. Notably, we wouldn't expect this recovery to be perfect—emotion regulation is an integrative and holistic cognitive process, and any network parcellation scheme will contain noise. Thus, our hope to was to ensure that nodes from each network were consistently identified in the same community, without being too concerned about a perfect network recovery. Indeed, our network recovery analysis via clustering was successful insofar that our network distinctions were reasonably reconstructed (Supplementary Table 10).

Supplementary Table 1. List of network ROI labels and their coordinates

CRN	*DMN*	*FPN*	*SaN*	*DAN*	*VAN*
R dlPFC_1 [60, 24, 3]	ACC_1 [2, 22, 44]	R SPL_1 [18, -68, 52]	ACC_1 [6, 26, 28]	L LOC_1 [-46, -68, 28]	L LOC_1 [-28, -74, 26]
R dlPFC_2 [48, 24, 9]	ACC_2 [2, 6, 54]	R SPL_2 [26, -60, 42]	ACC_2 [2, 24, 44]	L LOC_2 [-26, -66, 48]	L LOC_2 [-28, -60, 44]
R dlPFC_3 [60, 24, 3]	ACC_3 [2, 0, 56]	R SPL_3 [44, -44, 48]	PCC_1 [-2, -56, 24]	L LOC_3 [-48, -66, -4]	L LOC_3 [-28, -74, 20]
L dlPFC_1 [-33, 3, 54]	ACC_4 [2, 2, 60]	R SPL_4 [38, -46, 48]	PCC_2 [-2, -52, 28]	L LOC_4 [-42, -62, 36]	L LOC_4 [-30, -80, -12]
L dlPFC_2 [-36, 15, 57]	ACC_5 [2, 46, 22]	R SPL_5 [36, -60, 46]	PCC_3 [0, -32, 32]	L LOC_5 [-40, -64, 32]	R LOC_1 [46, -62, -12]
L dlPFC_3 [-36, 21, -3]	PCC_1 [0, -50, 28]	R SPL_6 [20, -62, 60]	L AI_1 [-38, 14, -6]	R LOC_1 [24, -60, 52]	R LOC_2 [34, -70, 32]
L dlPFC_4 [-42, 18, 9]	PCC_2 [2, -26, 40]	R SPL_7 [30, -68, 52]	L AI_2 [-32, 18, 0]	R LOC_2 [48, -62, 28]	R LOC_3 [28, -60, 44]
L dlPFC_5 [-42, 45, -6]	PCC_3 [-8, -62 20]	R SPL_8 [26, 35, 60]	L AI_3 [-38, 22, 2]	L PCG_1 [-26, -6, 54]	L FFG_1 [-38, -54, -16]
L dlPFC_6 [-51, 12, 21]	mPFC_1 [-2, 50, -6]	L SPL_1 [-30, -54, 40]	R AI 1 [38, 18, 2]	R PCG_1 [44, 10, 28]	L FFG_2 [-28, -48, -12]
L dlPFC_7 [-51, 21, 9]	mPFC_2 [10, 54, 0]	L SPL_2 [-14, -66, 52]	R AI 2 [42, 6, -6]	R PCG_2 [32, -2, 52]	L FFG_3 [-28, -72, -12]
dmPFC_1 [9, 30, 39]	mPFC_3 [2, 44, -16]	L SPL_3 [-22, -60, 52]	R AI 3 [44, 18, -4]	R PCG_3 [28, -10, 52]	R FFG_1 [28, -42, -12]
dmPFC_2 [0, 15, 63]	L TPJ_1 [-46, -70, 32]	R dlPFC_1 [50, 10, 28]	L TPJ_1 [-48, -64, 32]	R PCG_4 [28, -6, 50]	R FFG_2 [32, -58, -8]
dmPFC_3 [0, 6, 63]	L TPJ_2 [-34, -46, 44]	R dlPFC_2 [46, 6, 36]	L TPJ_2 [-40, -58, 44]	L SPL_1 [-32, -50, 48]	R FFG_3 [26, -62, -6]
dmPFC_4 [0, -9, 63]	R TPJ_1 [50, -62, 26]	R dlPFC_3 [50, 12, 24]	L TPJ_3 [-44, -38, 46]	L SPL_2 [-34, -46, 44]	R FFG_4 [32, -62, -8]
dmPFC_5 [0, 18, 42]	R TPJ_2 [38, -48, 44]	R dlPFC_4 [46, 8, 30]	-	L SPL_3 [-30, -56, 54]	R FFG_5 [32, -64, -12]
dmPFC_6 [-3, 24, 30]	L Hip_1 [-26, -22, -16]	L dlPFC_1 [-44, 6, 32]	-	L SPL_4 [-44, -56, 48]	L IFG_1 [-46, 12, 28]
dmPFC_7 [-9, 12, 69]	L Hip_2 [-26, -38, -14]	ACC_1 [6, 16, 48]	-	L SPL_5 [-48 ,-54, 46]	L AI_1 [-32, 20, -2]
R vlPFC_1 [51, 15, 48]	L Hip_3 [26, -42, -16]	ACC_2 [6, 26, 38]	-	R SPL_1 [38, -46, 50]	R AI_1 [34, 24, 4]
R vlPFC_2 [51, 6, 48]	R Hip_1 [26, -22, -20]	ACC_3 [0, 6, 54]	-	R SPL_2 [44, -52, 40]	L OP_1 [-12, -94, -4]
R vlPFC_3 [42, 21, 45]	L dlPFC_1 [-24, 26, 48]	L AI _1 [-32, 22, -2]	-	ACC_1 [4, 24, 40]	ACC_1 [2, 18, 48]
R vlPFC_4 [42, 30, 39]	L dlPFC_2 [-46, 10, 30]	R AI_1 [34, 22, -2]	-	PCC_1 [2, -52, 28]	-
L SPL_1 [-42, -66, 42]	R dlPFC_1 [26, 32, 44]	Tha_1 [-10, -12, 6]	-	R AI_1 [34, 22, -4]	-
L SPL_2 [-45, -69, 18]	L AI_1 [-34, 22, 2]	Tha_2 [-14, -14, 4]	-	L dlPFC_1 [-46, 16, 32]	-
L SPL_3 [-51, -60, 27]	R AI_1 [34, 22, -2]	L PMC_1 [-28, -4, 54]	-	L MFG_2 [-46, 10, 34]	-
L SPL_4 [-54, -72, 27]	Tha_1 [10, -18, 8]	R PMC_1 [26, 0, 52]	-	mPFC_1 [-4, 50, -8]	-
L SPL_5 [-63, -51, 21]	L TS_1 [-60, -14, -16]	-	-	mPFC_2 [-6, 54, 18]	-
R SPL_1 [63, -51, 39]	-	-	-	mPFC_3 [-2, 54, -12]	-
R SPL_2 [60, -60, 30]	-	-	-	-	-
R SPL_3 [51, -60, 42]	-	-	-	-	-
R SPL_4 [60, -45, 27]	-	-	-	-	-
L MTG_1 [-51, -39, 3]	-	-	-	-	-
L MTG_2 [-57, -24, -12]	-	-	-	-	-

Note. CRN = Cognitive Reappraisal Network; DMN = Default Mode Network; SaN = Salience Network; DAN = Dorsal Attention Network; VAN = Ventral Attention Network

Supplementary Table 2. 95% Confidence intervals on classification accuracies of Models A1-A9 (age outcome)

Model	*Child*	*Teen*
A1. All Networks, *k* = 3	*[.59, .88]*	[.31, .66]
A2. All Networks, *k* = 5	**[.56, .86]**	***[.34, .69]***
A3. All Networks, *k* = 7	[.59, .88]	[.23, .58]
A4. CRN, *k* = 5	[.27, .61]	[.42, .77]
A5. DMN, *k* = 5	[.37, .71]	[.31, .66]
A6. FPN, *k* = 5	[.45, .78]	[.23, .58]
A7. SaN, *k* = 5	**[.39, .73]**	**[.48, .82]**
A8. DAN, *k* = 5	*[.53, .84]*	[.28, .64]
A9. VAN, *k* = 5	[.29, .63]	*[.54, .87]*

Note. Italics and bold indicate best models from Table 2 in the main document: Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores).

Supplementary Table 3. 95% Confidence intervals on classification accuracies of Models C1-C9 (ability outcome).

Model	*High*	*Low*
C1. All Networks, *k* = 3	***[.22, .56]***	***[.52, .84]***
C2. All Networks, *k* = 5	[.13, .44]	[.58, .88]
C3. All Networks, *k* = 7	[.13, .44]	[.61, .90]
C4. CRN, *k* = 3	***[.35, .70]***	***[.43, .77]***
C5. DMN, *k* = 3	[.27, .62]	[.30, .65]
C6. FPN, *k* = 3	[.25, .59]	[.38, .72]
C7. SaN, *k* = 3	[.22, .56]	[.33, .37]
C8. DAN, *k* = 3	[.32, .68]	[.23, .57]
C9. VAN, *k* = 3	[.30, .65]	[.28, .62]

Note. Italics and bold indicate best models from Table 3 in the main document. Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores).

Supplementary Table 4. Correlations between participant age and capacity, and three network metrics between task states (Far and Look-Negative) thresholding the network at r = .4000.

Network	*Age - Similarity*	*Ability - Similarity*	*Age - Modularity*	*Ability - Modularity*	*Age - Community*	*Ability - Community*
CRN	0.09 [-.11, .29]	-0.19 [-.37, .01]	0.05 [-.15, .25]	0.19 [-.01, .19]	-0.05 [-.24, .15]	-0.24 [-.41, -.04]
DMN	0.05 [-.15, .25]	0.05 [-.15, .25]	0.05 [-.15, .25]	0.14 [-.06, .33]	0.04 [-.16, .24]	0.04 [-.16, .24]
FPN	-0.06 [-.25, .14]	-0.04 [-.24, .16]	-0.07 [-.27, .13]	-0.14 [-.33, .06]	0.21 [.01, .39]	0.17 [-.03, .35]
SaN	0.09 [-.11, .28]	0.00 [-.20, .20]	0.01 [-.19, .21]	0.00 [-.20, .20]	-0.08 [-.28, .12]	0.01 [-.19, .21]
DAN	0.12 [-.08, .31]	-0.08 [-.27, .12]	0.05 [-.15, .24]	-0.11 [-.30, .09]	0.13 [-.07, .32]	0.00 [-.20, .20]
VAN	-0.02 [-.22, .18]	0.02 [-.18, .22]	-0.03 [-.23, .17]	-0.12 [-.31, .08]	0.20 [.00, .38]	0.10 [-.10, .29]

Note. CRN = Cognitive Reappraisal Network; DMN = Default Mode Network; SaN = Salience Network; DAN = Dorsal Attention Network; VAN = Ventral Attention Network. Correlation coefficients are reported along with 90% confidence intervals.

Supplementary Table 5.

A. Classification accuracy of KNN models predicting age from network level metrics of brain activity, adjacency matrix defining threshold of .4000

Model	*Precision - Child*	*Recall - Child*	*F_1 - Child*	*Precision - Teen*	*Recall - Teen*	*F_1 - Teen*	*F_1 - Composite*
A1. All Networks, *k* = 3	0.757	0.549	0.636	*0.303*	0.526	*0.385*	0.479
A2. All Networks, *k* = 5	0.834	0.564	*0.674*	0.273	0.600	0.375	**0.482**
A3. All Networks, *k* = 7	*0.838*	0.554	0.667	0.242	0.571	0.340	0.451
A4. CRN, *k* = 5	0.432	0.552	0.485	0.606	0.488	0.541	0.511
A5. DMN, *k* = 5	0.514	0.500	0.507	0.424	0.438	0.431	0.466
A6. FPN, *k* = 5	0.622	0.535	0.575	0.394	0.481	0.433	0.494
A7. SaN, *k* = 5	0.568	0.656	0.609	0.667	0.579	0.620	**0.614**
A8. DAN, *k* = 5	*0.703*	0.591	*0.642*	0.455	0.577	0.508	0.567
A9. VAN, *k* = 5	0.459	0.654	0.540	*0.727*	0.545	*0.623*	0.579

Note. Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores). The horizontal line subdivides two types of model. Those above the line use information from all networks; those below the line are network-specific.

B. Confidence intervals on classification accuracies of Models A1-A9 (age outcome), adjacency matrix defining threshold of .4000

Model	*Child*	*Teen*
A1. All Networks, *k* = 3	[.59, .88]	*[.13, .46]*
A2. All Networks, *k* = 5	**[.68, .94]**	**[.13, .46]**
A3. All Networks, *k* = 7	*[.68, .94]*	[.11, .42]
A4. CRN, *k* = 5	[.37, .71]	[.28, .64]
A5. DMN, *k* = 5	[.34, .68]	[.25, .61]
A6. FPN, *k* = 5	[.50, .82]	[.11, .42]
A7. SaN, *k* = 5	**[.22, .55]**	**[.11, .42]**
A8. DAN, *k* = 5	*[.32, .66]*	[.23, .58]
A9. VAN, *k* = 5	[.42, .75]	*[.34, .69]*

Note. Italics and bold indicate best models from Supplementary Table 5A. Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores).

Supplementary Table 6.

A. Classification accuracy of KNN models predicting emotion regulation ability from network level metrics of brain activity, adjacency matrix defining threshold of .4000

Model	*Precision – High*	*Recall - High*	*F_1 - High*	*Precision - Low*	*Recall - Low*	*F_1 - Low*	*F_1 - Composite*
C1. All Networks, *k* = 3	*0.412*	0.538	*0.467*	0.667	0.545	0.600	**0.525**
C2. All Networks, *k* = 5	0.294	0.455	0.357	0.667	0.500	0.571	0.440
C3. All Networks, *k* = 7	0.265	0.563	0.360	*0.806*	0.537	*0.644*	0.462
C4. CRN, *k* = 3	0.471	0.485	0.478	0.528	0.514	0.521	0.498
C5. DMN, *k* = 3	0.412	0.359	0.384	0.306	0.355	0.328	0.354
C6. FPN, *k* = 3	0.441	0.652	0.526	*0.778*	0.596	*0.675*	**0.591**
C7. SaN, *k* = 3	0.353	0.353	0.353	0.389	0.389	0.389	0.370
C8. DAN, *k* = 3	*0.618*	0.512	*0.560*	0.444	0.552	0.492	0.524
C9. VAN, *k* = 3	0.412	0.389	0.400	0.389	0.412	0.400	0.400

Note. Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores). The horizontal line subdivides two types of model. Those above the line use information from all networks; those below the line are network-specific.

B. Confidence intervals on classification accuracies of Models C1-C9 (ability outcome), adjacency matrix defining threshold of .4000

Model	*High*	*Low*
A1. All Networks, *k* = 3	***[.25, .59]***	**[.49, .81]**
A2. All Networks, *k* = 5	[.15, .47]	[.49, .81]
A3. All Networks, *k* = 7	[.13, .44]	*[.64, .92]*
A4. CRN, *k* = 3	[.30, .65]	[.35, .70]
A5. DMN, *k* = 3	[.25, .59]	[.16, .48]
A6. FPN, *k* = 3	**[.27, .62]**	***[.61, .90]***
A7. SaN, *k* = 3	[.20, .54]	[.23, .57]
A8. DAN, *k* = 3	*[.44, .78]*	[.28, .62]
A9. VAN, *k* = 3	[.25, .59]	[.23, .57]

Note. Italics and bold indicate best models from Supplementary Table 6A. Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores).

Supplementary Table 7. Correlations between participant age and capacity, and three network metrics between task states (Far and Look-Negative) thresholding the network at r = .6000.

Network	*Age - Similarity*	*Ability - Similarity*	*Age - Modularity*	*Ability - Modularity*	*Age - Community*	*Ability - Community*
CRN	0.09 [-.11, .29]	-0.19 [-.37, .01]	-0.05 [-.24, .15]	0.13 [-.07, .32]	0.05 [-.15, .24]	-0.09 [-.28, .11]
DMN	0.05 [-.15, .25]	0.05 [-.15, .25]	0.01 [-.19, .21]	0.04 [-.16, .24]	-0.07 [-.27, .13]	-0.12 [-.31, .08]
FPN	-0.06 [-.25, .14]	-0.04 [-.24, .16]	0.09 [-.11, .28]	-0.11 [-.30, .09]	0.17 [-.03, .36]	0.12 [-.08, .31]
SaN	0.09 [-.11, .28]	0.00 [-.20, .20]	-0.04 [-.24, .16]	0.09 [-.11, .28]	0.03 [-.17, .22]	0.12 [-.08, .31]
DAN	0.12 [-.08, .31]	-0.08 [-.27, .12]	-0.05 [-.25, .15]	-0.09 [-.28, .11]	0.16 [-.03, .35]	0.15 [-.05, .34]
VAN	-0.02 [-.22, .18]	0.02 [-.18, .22]	-0.03 [-.22, .17]	-0.14 [-.33, .06]	0.08 [-.12, .27]	0.26 [.07, .44]

Note. CRN = Cognitive Reappraisal Network; DMN = Default Mode Network; SaN = Salience Network; DAN = Dorsal Attention Network; VAN = Ventral Attention Network. Correlation coefficients are reported along with 90% confidence intervals.

Supplementary Table 8.

A. Classification accuracy of KNN models predicting age from network level metrics of brain activity, adjacency matrix defining threshold of .6000

Model	*Precision - Child*	*Recall - Child*	*F_1 - Child*	*Precision - Teen*	*Recall - Teen*	*F_1 - Teen*	*F_1 - Composite*
A1. All Networks, *k* = 3	0.676	0.510	0.581	0.273	0.429	0.333	0.424
A2. All Networks, *k* = 5	*0.834*	0.564	*0.674*	0.273	0.600	*0.375*	**0.482**
A3. All Networks, *k* = 7	0.412	0.489	0.537	*0.303*	0.400	0.345	0.420
A4. CRN, *k* = 5	0.324	0.316	0.320	0.212	0.219	0.215	0.257
A5. DMN, *k* = 5	0.351	0.351	0.351	0.273	0.273	0.273	0.307
A6. FPN, *k* = 5	*0.676*	0.625	*0.649*	*0.545*	0.600	*0.571*	**0.607**
A7. SaN, *k* = 5	0.459	0.447	0.453	0.363	0.375	0.369	0.407
A8. DAN, *k* = 5	0.649	0.558	0.600	0.424	0.519	0.467	0.525
A9. VAN, *k* = 5	0.486	0.450	0.468	0.333	0.367	0.349	0.400

Note. Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores). The horizontal line subdivides two types of model. Those above the line use information from all networks; those below the line are network-specific.

B. Confidence intervals on classification accuracies of Models A1-A9 (age outcome), adjacency matrix defining threshold of .6000.

Model	*Child*	*Teen*
A1. All Networks, *k* = 3	[.42, .75]	[.16, .49]
A2. All Networks, *k* = 5	***[.50, .82]***	***[.23, .58]***
A3. All Networks, *k* = 7	[.42, .75]	[.16, .49]
A4. CRN, *k* = 5	[.18, .50]	[.09, .39]
A5. DMN, *k* = 5	[.20, .53]	[.13, .46]
A6. FPN, *k* = 5	***[.50, .82]***	***[.36, .72]***
A7. SaN, *k* = 5	[.29, .63]	[.20, .55]
A8. DAN, *k* = 5	[.47, .80]	[.25, .61]
A9. VAN, *k* = 5	[.32, .66]	[.18, .52]

Note. Italics and bold indicate best models from Supplementary Table 8A. Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores).

Supplementary Table 9.

A. Classification accuracy of KNN models predicting emotion regulation ability from network level metrics of brain activity, adjacency matrix defining threshold of .6000

Model	*Precision – High*	*Recall - High*	*F_1 - High*	*Precision - Low*	*Recall - Low*	*F_1 - Low*	*F_1 - Composite*
C1. All Networks, *k* = 3	*0.412*	0.560	*0.475*	*0.694*	0.556	*0.617*	**0.536**
C2. All Networks, *k* = 5	0.265	0.375	0.310	0.583	0.457	0.512	0.387
C3. All Networks, *k* = 7	0.294	0.435	0.351	0.639	0.489	0.554	0.429
C4. CRN, *k* = 3	*0.706*	0.727	*0.716*	*0.750*	0.729	*0.740*	**0.728**
C5. DMN, *k* = 3	0.294	0.323	0.308	0.417	0.385	0.400	0.348
C6. FPN, *k* = 3	0.324	0.478	0.386	0.667	0.511	0.578	0.463
C7. SaN, *k* = 3	0.529	0.514	0.522	0.528	0.543	0.535	0.528
C8. DAN, *k* = 3	0.471	0.390	0.427	0.306	0.379	0.338	0.377
C9. VAN, *k* = 3	0.618	0.477	0.538	0.361	0.500	0.419	0.472

Note. Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores). The horizontal line subdivides two types of model. Those above the line use information from all networks; those below the line are network-specific.

B. Confidence intervals on classification accuracies of Models C1-C9 (ability outcome).

Model	*High*	*Low*
A1. All Networks, *k* = 3	***[.25, .59]***	***[.52, .84]***
A2. All Networks, *k* = 5	[.13, .44]	[.41, .74]
A3. All Networks, *k* = 7	[.15, .47]	[.46, .79]
A4. CRN, *k* = 3	***[.53, .85]***	***[.58, .88]***
A5. DMN, *k* = 3	[.15, .47]	[.26, .59]
A6. FPN, *k* = 3	[.17, .51]	[.49, .81]
A7. SaN, *k* = 3	[.35, .70]	[.35, .70]
A8. DAN, *k* = 3	[.30, .65]	[.16, .48]
A9. VAN, *k* = 3	[.44, .78]	[.21, .54]

Note. Italics and bold indicate best models from Supplementary Table 9A. Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores).

Supplementary Table 10. Data driven network recovery.

Recovered Community	CRN	DMN	FPN	SaN	DAN	VAN
1 (nodes = 20)	55%	15%	0%	10%	20%	0%
2 (nodes = 19)	37%	26%	16%	11%	5%	5%
3 (nodes = 9)	0%	22%	0%	0%	0%	78%
4 (nodes = 10)	0%	10%	50%	0%	30%	10%
5 (nodes = 30)	0%	7%	43%	3%	37%	10%
6 (nodes = 20)	35%	10%	10%	30%	5%	10%
7 (nodes = 6)	0%	0%	0%	0%	17%	83%
8 (nodes = 12)	58%	17%	0%	8%	17%	0%
9 (nodes = 2)	0%	50%	0%	0%	50%	0%
10 (nodes = 5)	0%	40%	0%	40%	10%	0%
11 (nodes = 2)	0%	100%	0%	0%	0%	0%
12 (nodes = 2)	0%	0%	0%	0%	0%	100%
13 (nodes = 3)	0%	67%	0%	0%	33%	0%
14 (nodes = 3)	0%	33%	67%	0%	0%	0%
15 (nodes = 1)	0%	100%	0%	0%	0%	0%
16 (nodes = 1)	0%	0%	0%	0%	100%	0%

Note. Rows represent each recovered network community from our network 'sanity check' analysis. Columns represent networks and entries indicate the percentage of nodes in a recovered community belonging to a given network. CRN = Cognitive Reappraisal Network; DMN = Default Mode Network; SaN = Salience Network; DAN = Dorsal Attention Network; VAN = Ventral Attention Network.

Supplementary Table 11.

A. Classification accuracy of KNN models predicting age from network level metrics of brain activity. Age cutoff = 13

Model	*Precision – High*	*Recall - High*	*F_1 - High*	*Precision - Low*	*Recall - Low*	*F_1 - Low*	*F_1 - Composite*
C1. All Networks, *k* = 3	0.717	0.571	*0.636*	*0.323*	0.476	*0.385*	**0.479**
C2. All Networks, *k* = 5	0.717	0.571	*0.636*	*0.323*	0.476	*0.385*	**0.479**
C3. All Networks, *k* = 7	*0.744*	0.547	0.630	0.226	0.412	0.292	0.399
C4. CRN, *k* = 5	0.436	0.472	0.453	0.387	0.353	0.369	0.407
C5. DMN, *k* = 5	0.513	0.513	0.513	0.387	0.387	0.387	0.441
C6. FPN, *k* = 5	0.667	0.565	0.612	0.355	0.458	0.400	0.484
C7. SaN, *k* = 5	0.641	0.694	*0.667*	0.645	0.588	*0.615*	**0.640**
C8. DAN, *k* = 5	*0.692*	0.574	0.628	0.355	0.478	0.407	0.494
C9. VAN, *k* = 5	0.487	0.679	0.567	*0.710*	0.524	0.603	0.584

Note. Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores). The horizontal line subdivides two types of model. Those above the line use information from all networks; those below the line are network-specific.

B. Classification accuracy of KNN models predicting age from network level metrics of brain activity. 50-50 split (corresponding to age of 12.5)

Model	*Precision – Child*	*Recall - Child*	*F_1 - Child*	*Precision - Teen*	*Recall - Teen*	*F_1 - Teen*	*F_1 - Composite*
C1. All Networks, *k* = 3	*0.686*	0.585	*0.632*	0.514	0.621	0.563	**0.595**
C2. All Networks, *k* = 5	0.629	0.579	0.603	*0.543*	0.594	*0.567*	0.584
C3. All Networks, *k* = 7	0.657	0.523	0.582	0.400	0.538	0.459	0.513
C4. CRN, *k* = 3	0.286	0.370	0.322	0.514	0.419	0.462	0.380
C5. DMN, *k* = 3	0.457	0.471	0.464	0.457	0.485	0.471	0.467
C6. FPN, *k* = 3	0.600	0.512	0.553	0.429	0.517	0.469	0.507
C7. SaN, *k* = 3	*0.629*	0.647	*0.638*	0.657	0.639	*0.648*	**0.642**
C8. DAN, *k* = 3	0.514	0.474	0.493	0.429	0.469	0.448	0.469
C9. VAN, *k* = 3	0.571	0.645	0.606	*0.686*	0.545	0.608	0.607

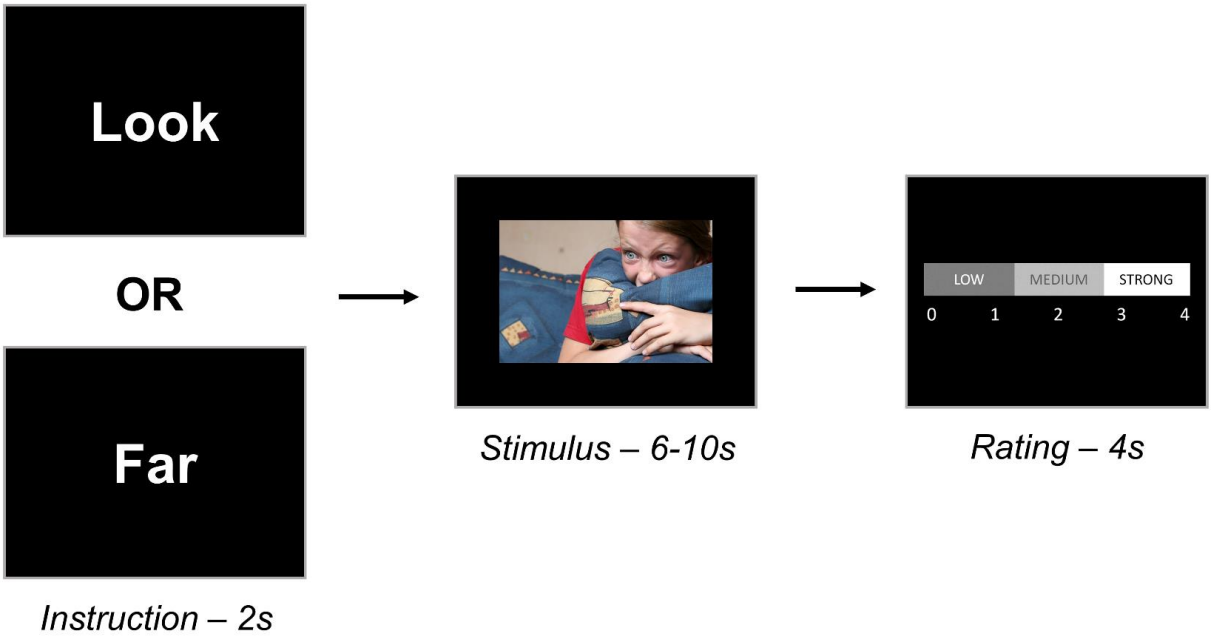
Note. Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores). The horizontal line subdivides two types of model. Those above the line use information from all networks; those below the line are network-specific.

Supplementary Table 12. Classification accuracy of KNN models predicting emotion regulation ability from network level metrics of brain activity. Median split cut-off.

Model	*Precision – High*	*Recall - High*	*F_1 - High*	*Precision - Low*	*Recall - Low*	*F_1 - Low*	*F_1 - Composite*
C1. All Networks, *k* = 3	*0.400*	0.583	*0.475*	0.714	0.543	0.617	**0.537**
C2. All Networks, *k* = 5	0.257	0.500	0.340	*0.743*	0.500	*0.598*	0.433
C3. All Networks, *k* = 7	0.257	0.500	0.340	*0.743*	0.500	*0.598*	0.433
C4. CRN, *k* = 3	*0.543*	0.594	*0.567*	*0.629*	0.579	*0.603*	**0.584**
C5. DMN, *k* = 3	0.429	0.405	0.417	0.371	0.394	0.382	0.399
C6. FPN, *k* = 3	0.457	0.471	0.464	0.486	0.472	0.479	0.471
C7. SaN, *k* = 3	0.400	0.452	0.424	0.514	0.462	0.486	0.453
C8. DAN, *k* = 3	0.514	0.461	0.486	0.400	0.452	0.424	0.453
C9. VAN, *k* = 3	0.457	0.444	0.451	0.429	0.441	0.434	0.443

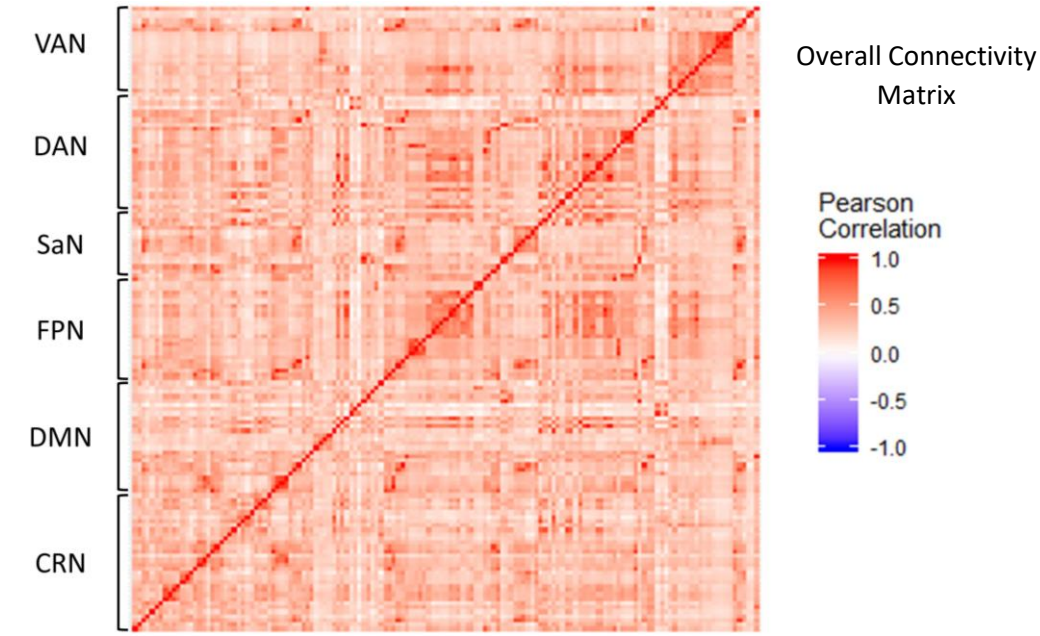
Note. Entries in italics signify best accuracy for class-specific metrics (precision and recall). Entries in bolded font indicate best overall model (judged by composite F_1 scores). The horizontal line subdivides two types of model. Those above the line use information from all networks; those below the line are network-specific.

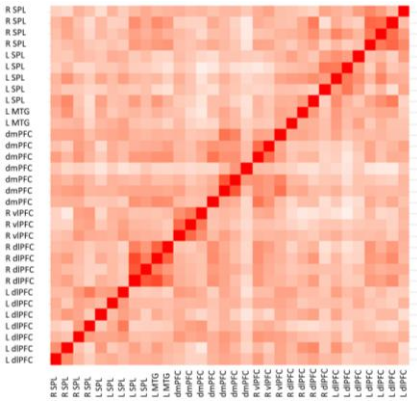
Supplemental Figure 1. Emotion Regulation Paradigm



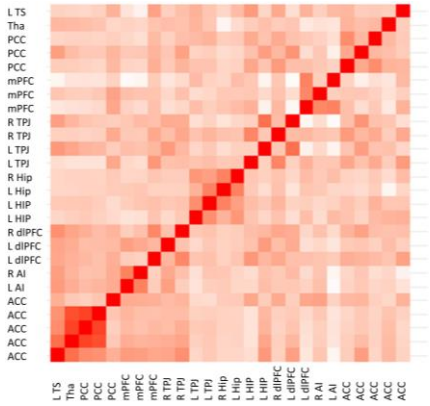
Note. 'Far' denotes the instructional cue participants received to engage in emotion regulation (via reappraisal). 'Look' is the cue shown when participants were to passively observe the subsequent stimuli. Participants then regulated aversive stimuli or passively viewed aversive or neutral (not pictured) stimuli. A five-point Likert scale was used to affective responses. A jittered ITI (1.5-6.5s) concluded each trial.

Supplemental Figure 2. Functional connectivity matrices.

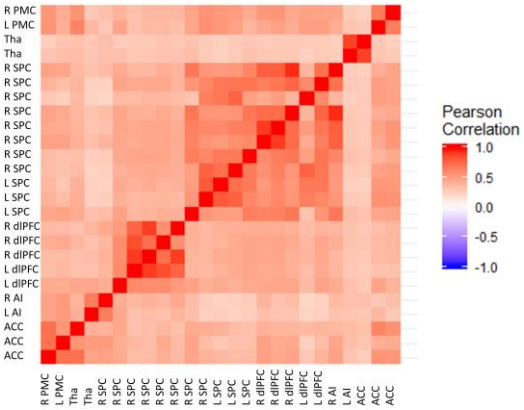




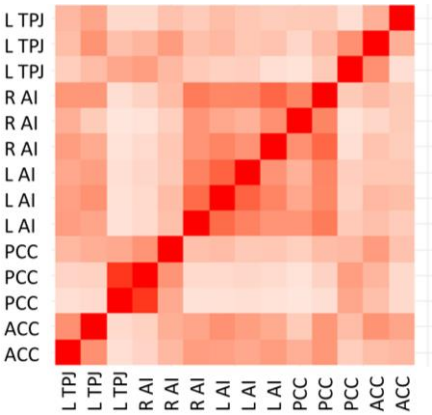
CRN Connectivity Matrix



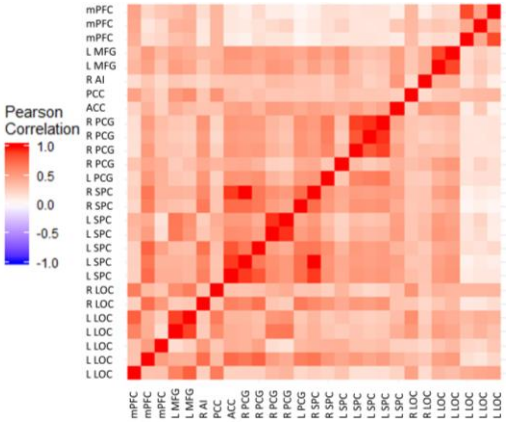
DMN Connectivity Matrix



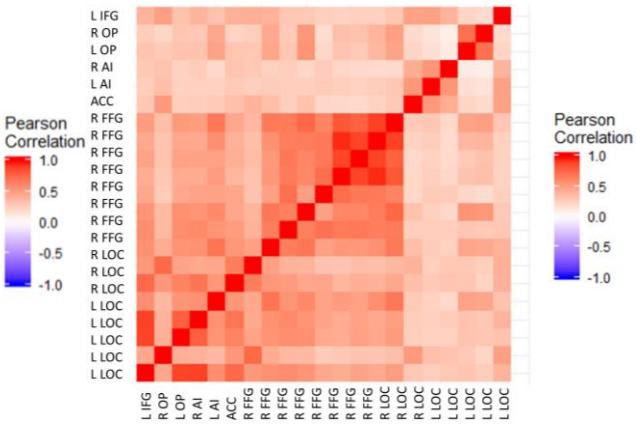
FPN Connectivity Matrix



SaN Connectivity Matrix



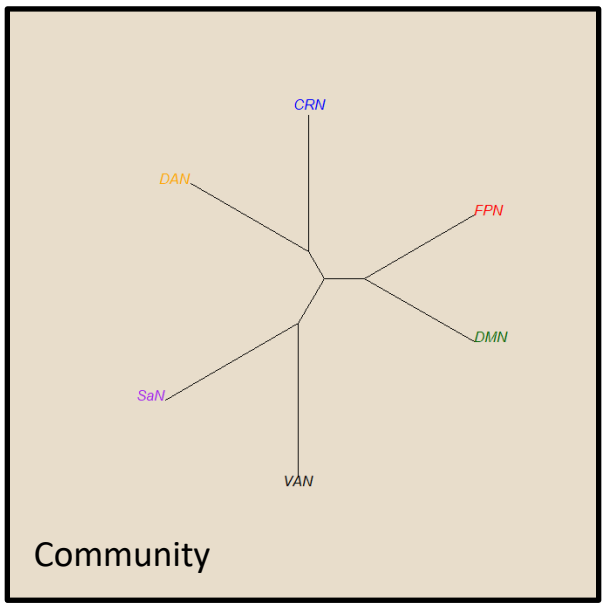
DAN Connectivity Matrix

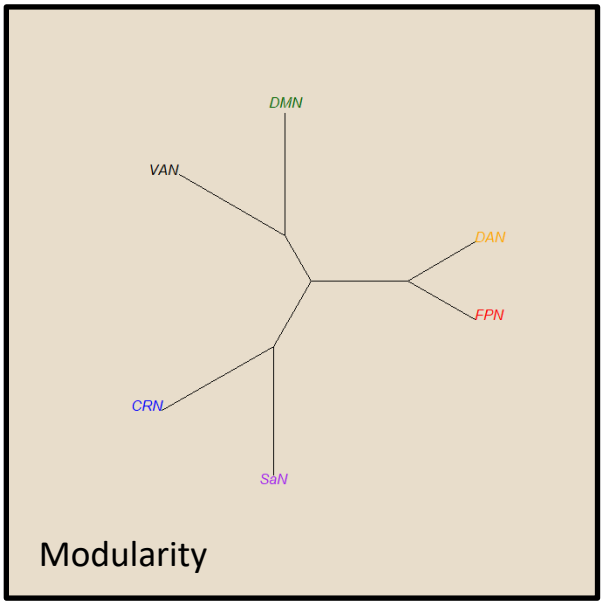
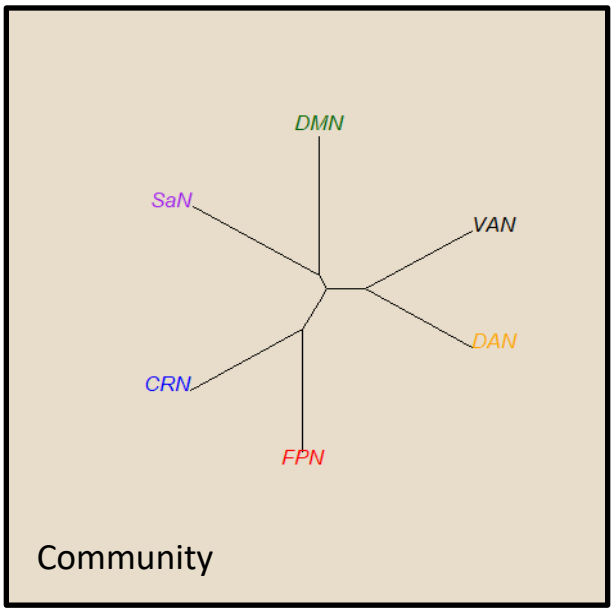
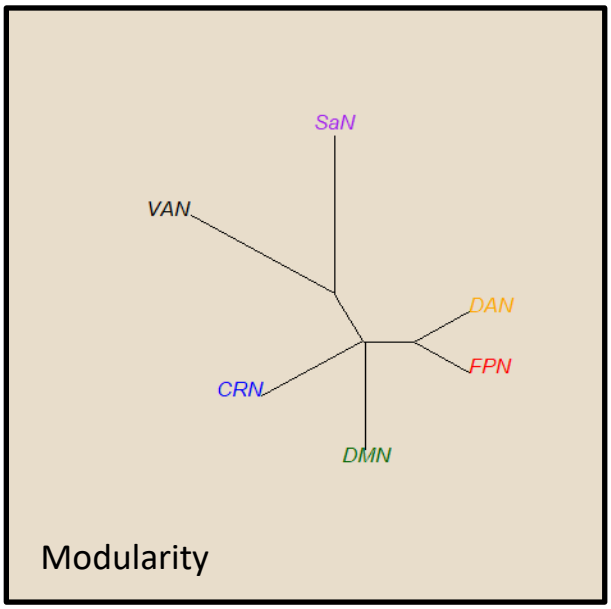


VAN Connectivity Matrix

Supplemental Figure 3.

A. Network taxonomies based on community and modularity at an edge defining threshold of .4000





B. Network taxonomies based on community and modularity at an edge defining threshold of .6000