NeuroImage 229 (2021) 117630

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Functional connectivity patterns predict naturalistic viewing versus rest across development

Sara Sanchez-Alonso[a,*], Monica D. Rosenberg[b], Richard N. Aslin[a,c,d]

[a] Haskins Laboratories, New Haven, CT, USA
[b] Department of Psychology, University of Chicago, Chicago, IL, USA
[c] Department of Psychology, Yale University, New Haven, CT, USA
[d] Child Study Center, Yale University, New Haven, CT, USA

2046

A R T I C L E I N F O

Keywords:
Resting-state
Naturalistic imaging
Movie-watching
Brain state
Development
Functional connectivity

A B S T R A C T

Cognitive states, such as rest and task engagement, share an 'intrinsic' functional network organization that is subject to minimal variation over time and yields stable signatures within an individual. Importantly, there are also transient state-specific functional connectivity (FC) patterns that vary across neural states. Here, we examine functional brain organization differences that underlie distinct states in a cross-sectional developmental sample. We compare FC fMRI data acquired during naturalistic viewing (i.e., movie-watching) and resting-state paradigms in a large cohort of 157 children and young adults aged 6–20. Naturalistic paradigms are commonly implemented in pediatric research because they maintain the child's attention and contribute to reduced head motion. It remains unknown, however, to what extent the brain-wide functional network organization is comparable during movie-watching and rest across development. Here, we identify a widespread FC pattern that predicts whether individuals are watching a movie or resting. Specifically, we develop a model for prediction of multilevel neural effects (termed PrimeNet), which can with high reliability distinguish between movie-watching and rest irrespective of age and that generalizes across movies. In turn, we characterize FC patterns in the most predictive functional networks for movie-watching versus rest and show that these patterns can indeed vary as a function of development. Collectively, these effects highlight a 'core' FC pattern that is robustly associated with naturalistic viewing, which also exhibits change across age. These results, focused here on naturalistic viewing, provide a roadmap for quantifying state-specific functional neural organization across development, which may reveal key variation in neurodevelopmental trajectories associated with behavioral phenotypes.

1. Introduction

Cognitive processes, including crucial abilities such as attention, fluid intelligence, and language, arise from complex neural interactions that are reflected in the brain's large-scale functional organization. In the adult brain, these interactions mature to a 'trait-like' pattern that is relatively stable over time and subject to minimal day-to-day variability (Gratton et al., 2018). This 'intrinsic' functional network organization is indeed shared across brain states (e.g., rest and tasks) (Cole et al., 2014; Fox et al., 2007). Although these functional connectivity (FC) patterns yield stable signatures within an individual (Finn et al., 2015; Miranda-Dominguez et al., 2014), there is also variation in FC patterns across cognitive states such as attentional engagement (Rosenberg et al., 2020), semantic judgment (Gratton et al., 2016) or motor tasks (Cole et al., 2014). Thus, despite the existence of an intrinsic functional network organization that extends across brain states, these studies also support the existence of transient state-specific FC patterns that differentiate between rest and task and vary across cognitive states.

Cognitive states and the functional brain architecture that underlies them are not fixed across the lifespan, but rather are shaped by neurodevelopment. A number of studies have examined the developmental trajectory of the intrinsic functional organization of the brain at 'rest' during early and late childhood (Fair et al., 2009; Gao et al., 2015; Smyser et al., 2010; Supekar et al., 2009). Nevertheless, direct comparisons of resting-state and task-based FC in childhood or comparisons of resting-state FC between children and adults are limited due to the challenges of collecting fMRI data with children. Specifically, younger children are often unwilling or unable to adequately comply with the demands of resting-state scans, which require them to remain awake and motionless during 6–10 min runs of data-collection. This is especially problematic for resting-state paradigms because correlation-based analyses such as FC measures are more sensitive to head motion

* Corresponding author.
E-mail address: sara.sanchez.alonso@yale.edu (S. Sanchez-Alonso).

https://doi.org/10.1016/j.neuroimage.2020.117630
Received 12 September 2019; Received in revised form 16 October 2020; Accepted 7 December 2020
Available online 2 January 2021

than task activation analyses (Power et al., 2012). As a result, children younger than school age are often scanned while performing attention-grabbing tasks, asleep, or sedated rather than resting (Damaraju et al., 2014; Dinstein et al., 2011; C. T. Ellis et al., 2020; Konishi et al., 2002; Uematsu et al., 2012) (see reviews by (Cameron T. Ellis & Turk-Browne, 2018; Snyder, 2016)).

Consequently, a main challenge in pediatric neuroimaging has been to examine the brain's functional network organization using paradigms that are better suited for young children and that allow direct comparison with the adult brain unconfounded by differences in state. To avoid developmental comparisons that involve sleep, sedation, or cognitive tasks (which can be confounded by performance differences), a common strategy has been to scan young children using 'naturalistic paradigms,' which consist of stimuli that necessitate rapid integration of continuous real-time information (Bottenhorn et al., 2019; Sonkusare et al., 2019). The most commonly implemented type of naturalistic paradigm in pediatric research has, to date, been movies which provide a compelling mixture of visual and auditory events to maintain the child's cooperation (e.g., language, face processing, social interactions, different camera angles, music) and are amenable to FC analyses (i.e., identification of spatially separate brain regions or voxels that have correlated BOLD signal time-courses). Furthermore, movie-watching paradigms in the scanner have been demonstrated as an efficient method of reducing motion in children under the age of 10 (Greene et al., 2018; Vanderwal, Eilbott, and Castellanos, 2019). Therefore, it is becoming a common method of assessing the brain's functional network organization in younger populations. Importantly, the use of these two different paradigms, rest as a predominant method in adult research and movie-watching in children, requires an understanding of the similarities and differences in functional network organization between these two contexts.

Despite the growing role of movie-watching paradigms in neurodevelopmental research, it is still unclear to what extent the functional network organization during movie-watching and rest are comparable in the developing brain and how they differ across age. Prior work suggests that the two states are associated with distinct functional network organizations. There are FC changes that are specific to resting-state, as well as task-specific hubs (i.e., highly interconnected regions) in the frontoparietal, attention and default-mode networks, which are not observed during rest (Cole et al., 2014; Dixon et al., 2017; Gilson et al., 2018). Resting-state scans often exhibit a more distributed and stronger FC in comparison to movie-watching paradigms. These differences are more clearly observed in the visual, sensorimotor, default mode and dorsal-attention networks (Lynch et al., 2018). Other studies have shown that within-network connectivity in both visual and auditory networks increases during rest, whereas connectivity between visual and language networks increases during naturalistic-viewing (Betti et al., 2013). Furthermore, movie-watching evokes inter-subject correlations (ISCs), or time-locked responses that are shared across subjects, but reduced during rest (Baldassano et al., 2017; Hasson et al., 2004, 2010; Pajula et al., 2012; Simony et al., 2016). A recent study also showed that movie-watching FC overlaps with ISC patterns, but only in temporal and occipital regions, and that this core pattern could be explained by a single principal component (Demirtaş et al., 2019). Overall these studies indicate that movie-watching is associated with FC patterns specific to this attentive and engaged brain state, which are not observed during resting-state. Resting-state, on the other hand, often reveals more widespread FC that modulates state-specific networks. Collectively, these studies show that despite a common 'intrinsic' functional network organization shared by both movie-watching and resting states, these two neural states exhibit differences in FC patterns that remain uncharacterized across neurodevelopment.

Research on developmental changes in FC using resting-state paradigms have shown both whole-brain and network-specific effects. The neonatal brain already shows early formation of resting-state networks and a strong overlap with the mature adult brain (81%) (De Asis-Cruz et al., 2015; van den Heuvel et al., 2015). Developmental changes in brain network organization during the first two years of life lead to the establishment of sensory networks well before the formation of higher-level cognitive networks (Eggebrecht et al., 2017; Gao et al., 2015). Functional network maturation may extend until late childhood and seems to be ultimately dependent on the network's functional roles, thus displaying complex within- and between-network time-dependent changes (see Grayson and Fair (2017) for a recent review on developmental changes in large-scale functional networks). Finally, anatomical data (i.e., cortical thickness) between early childhood and late adolescence (5-18 years of age) further indicates that structural brain networks serving basic functions (e.g., primary sensorimotor regions) mature the earliest relative to networks supporting higher-order cognitive functions (e.g., higher-order association and paralimbic regions) (Khundrakpam et al., 2013).

Studies investigating age-related differences in FC during movie-watching are, however, much more limited. One of the first studies that examined the relationship between movie-related FC and development identified a strong relationship between FC and age in a group of 44 2–6-year-old children (Long et al., 2017). Specifically, some networks (e.g., the frontoparietal network) exhibited increased FC change with age, whereas other networks exhibited more complex FC changes as a function of age, such as local-to-distributed FC shifts (e.g., temporal networks). Another set of studies in a group of girls aged 4-7 found that age is positively associated with FC in sensory and higher-order cognitive networks, further suggesting that attention engagement during movie-watching in early childhood involves a distributed functional network organization (Rohr et al., 2017, 2018). One of the few studies comparing movie-watching and resting-state FC across development focused on 6-year-olds (Emerson et al., 2015). This study showed that movie-watching in children elicits uncoupling between the visual and dorsal-attention networks relative to rest. Furthermore, children exhibit an adult-like pattern of network interactions among the dorsal-attention, the default-mode and the frontal control networks. In contrast to adults, however, children show marginally significant changes in the magnitude of these network-level interactions between movie-watching and rest, which the authors characterize as an 'immature' pattern. Collectively, these studies reveal a strong relationship between FC and age across both resting and movie-watching states. Network-specific effects seem to be characterized by i) an increase in FC with age, ii) a local-to-distributed shift in brain functional organization and iii) modulation of state-specific (movie versus rest) brain networks.

Crucially, to our knowledge, three core knowledge gaps remain unaddressed in our understanding of the brain's functional network organization during distinct cognitive states and associated developmental changes. 1) Although there is evidence for both intrinsic and, to a lesser degree, state-specific FC patterns in the adult brain, the extent to which a child's unique FC signature is affected by cognitive state remains relatively unexplored. 2) Despite initial evidence for movie versus rest differences, no study has characterized whether FC patterns consistently distinguish movie-watching from rest irrespective of age and whether these patterns generalize across movies. 3) No study has examined whether such state-specific FC patterns are associated with large-scale functional networks across the brain, and how these patterns vary as a function of age.

To address these knowledge gaps, here we compare movie-watching and resting-state fMRI data acquired from the same subjects. We hypothesize that these two states will differentially engage the default mode and dorsal-attention networks, networks associated with internal and external attention, in addition to networks that have been associated with processing auditory stimuli, such as the auditory and language networks. Furthermore, we expect that FC, particularly between networks, will be strengthened as a function of age. To test these hypotheses, we examine FC differences across movie-watching and resting states and how these differences may vary as a function of age. Specifically, we characterize movie-watching versus resting-state FC fMRI data in a group of 157 participants aged 6–20 from the publicly-available

Healthy Brain Network (HBN) dataset (Alexander et al., 2017). We quantify BOLD fluctuations during movie-watching and resting-state conditions using a whole-brain data-driven FC approach. To identify patterns that are predictive across both brain states and developmental trajectories, we developed the **Pr**ed **i**ction of **M**ulti-L **e**vel **N**eural **E**ffec **t**s (PrimeNet) framework. Specifically, the PrimeNet modeling framework was developed to concurrently test hypotheses related to within-subject (i.e., movie versus rest) and across-subject state changes (i.e., age). This multi-level modeling framework is specifically designed to address the key knowledge gaps highlighted above and is implemented here to assess how the effects of movie-watching map onto the developing functional connectome.

Using this analytic framework, we identify a robust set of FC patterns that are predictive of movie-watching relative to rest. We also explicitly test whether we can train a predictive model based on the identified FC features and irrespective of age. Furthermore, we use the Human Connectome Project parcellation (Glasser et al., 2016) and the recently developed brain-wide network partition (Ji et al., 2019) to establish whether the predictive patterns are localized to specific networks across the brain. Finally, building from this predictive set of FC features, we test whether these effects differ as a function of age. Collectively, this study highlights that there are robust and predictive movie-watching versus rest (that is, cognitive-state-specific) widespread FC patterns that are age-independent and show high overlap across different movies. In addition, a subset of the most predictive movie versus rest FC patterns is associated with specific large-scale functional networks. Finally, developmental status (i.e., age) interacts with the core predictive movie versus rest FC pattern. These effects highlight a ‘core’ movie-rest FC pattern, which serves as a basis for understanding the functional neural organization of naturalistic viewing. In turn, this predictive core FC pattern provides an entry point for characterizing similarities and differences in naturalistic viewing across development.

2. Materials and methods

2.1. Dataset

Data were obtained from the Healthy Brain Network (HBN) dataset, which is openly available in 1,000 Functional Connectomes Project and its International Neuroimaging Data-sharing Initiative (FCP/INDI) and can be accessed at http://fcon_1000.projects.nitrc.org/indi/cmi_healthy_brain_network/. The HBN study was approved by the Chesapeake Institutional Review Board (https://www.chesapeakeirb.com/). Written informed consent was obtained from participants ages 18 or older prior to enrolling in the study. Written assent was obtained from participants younger than 18 and written consent was obtained from their legal guardians. Data included in the study were collected across two different sites in the United States. Data from the first six releases (Data Release 6.0, 02/16/2019) were considered for analysis. Subjects were included if the following criteria were met: i) BOLD data met the requirements of the HCP minimum processing pipelines (Glasser et al., 2013), ii) BOLD data passed visual quality control (QC), iii) each BOLD run had more than 40% volumes remaining after scrubbing, iv) subjects had participated in at least a 5-min resting scan and a 3.47-min movie scan and v) subjects received no diagnosis or, if diagnosed, participants had no history of serious medical or neurological disorder that would likely affect cognitive functioning (i.e., a neurocognitive or depressive disorder, trauma, schizophrenia, or obsessive-compulsive disorder). After implementing these criteria, the original dataset with 1168 subjects was reduced to data from 157 individuals ranging in age from 6-20 years (mean: 11, SD: 3.5) (Table 1) (Demographics by study site can be found in **Supplementary Table S1** and subject IDs in **Table S6**.) We conducted secondary analyses on a subset of participants (n = 78) who had also participated in a second 10 min movie scan and a second 5 min resting-state scan. The HBN initiative focuses on the diagnosis of mental health and learning disorders in the developing human brain, thus providing a comprehensive overview of the clinical psychopathology status of each participant. In the selected dataset, a total of 63 individuals received no diagnosis, whereas 75 individuals were classified as having a neurodevelopmental disorder (attention deficit hyperactivity disorder [ADHD] or a specific learning disorder) and 19 as having an anxiety disorder (generalized anxiety disorder, social/separation anxiety or specific phobia).

2.2. Neuroimaging data acquisition

102 subjects underwent data collection at the Rutgers University Brain Imaging Center (RUBIC) and 55 subjects at the CitiGroup Cornell Brain Imaging Center (CBIC). Neuroimaging data acquired at RUBIC were obtained using a 3.0 T Siemens Tim Trio, with transversal slices parallel to the anterior-posterior commissure (AC-PC) using a T2*-weighted gradient-echo, multi-slice echo-planar sequence [time repetition (TR)/time echo (TE)=800/30, flip angle=31, field of view=216 × 216 mm, acquisition matrix = 90 × 90, voxel size = 2.4 × 2.4 × 2.4 mm]. Data acquisition produced 370 volumetric images per subject (66 slices/volume). Structural images were acquired using a T1-weighted image (TR/TE/time to inversion = 2500/2.88, flip angle = 8, field of view = 256 × 256 mm, acquisition matrix = 256 × 256, voxel size = 1 × 1 × 1 mm), with sagittal slices parallel to the AC-PC line. Neuroimaging data at the CBIC were obtained using a 3 T Siemens Prisma following the same data acquisition protocol implemented at RUBIC. The acquisition of the two resting-state scans lasted 5 min each, during which participants viewed a fixation cross located at the center of the computer screen. Throughout the scan, participants were instructed to open or close their eyes at various points. The acquisition of the first movie scan consisted of a 10 min clip of the movie ‘Despicable Me’ (Coffin and Renaud, 2010). The second movie lasted 3.47 min, during which participants viewed the animated short film ‘The Present’ (Frey, 2014). Both movies were played with English audio. The protocol included three additional scans, which consisted of calibration scans to monitor gaze direction via predictive eye estimation regression (*PEER* scans) (Alexander et al., 2017). The scans were presented in the following order across the two sites: *rest 1* (5 min), *peer 1* (1.9 min), *rest 2* (5 min), *peer 2* (1.9 min), *movie* ‘Despicable Me’ (10 min), *peer 3* (1.9 min), *movie* ‘The Present’ (3.47 min). For all subjects, we used the first rest (*rest 1*) and the second movie (*‘The Present’*) to conduct the main analyses in order to have a comparable number of frames per scan session. We conducted secondary analyses with data from the first movie (‘Despicable Me’) and the second resting-state run (*rest 2*) with the subset of subjects (n = 78) who participated in these sessions and whose BOLD data passed the established preprocessing criteria.

2.3. Data Preprocessing

The dataset was preprocessed in accordance with the Human Connectome Project (HCP) minimal preprocessing pipelines (Glasser et al., 2013). After HCP preprocessing and prior to parcellation, all BOLD images were mildly smoothed on the cortical surface and in the volume at a kernel of 2 mm isotropic to deal with subtle misalignment differences across development. Additionally, we implemented movement scrubbing following the guidelines described in (Power et al., 2012). Specifically, we computed the frame-to-frame signal intensity change between each two timeframes and calculated their normalized root mean square (RMS), divided by their mean intensity and multiplied by 100. This index reflects mean % change of the BOLD signal from frame to frame. We excluded frames when RMS>3, as well as the frame preceding and the 2 frames following the flagged frame. We also calculated an estimate of frame-to-frame displacement per subject by summing the displacement across all 6 rigid body movement correction parameters per frame. We removed volumes where mean frame-to-frame displacement (FD) estimate >0.5 mm across all parameters. Additionally, we

Table 1

Demographics by Age Group. *MovieTP*=movie 'The Present'. *Parental education* level is the average of the mother and father's education and was determined based on the following scale: 3=less than Grade 7; 6=Grade 8-9; 9=Grade 10-11; 12=High School graduate; 15=partial college (at least one year); 18=4-year college; 21=graduate degree. *tSNR* (temporal signal-to-noise ratio) was determined as the ratio of the mean signal for a given slice to the standard deviation across the relevant BOLD run, while excluding all non-brain voxels across all frames. *Percent (%) frames* refers to the percentage of frames that were included in the final analysis. *FD*=Framewise Displacement. *FD* and *DVARS* mean values were calculated on the raw data. Only scores for behavioral surveys with at least 90% of the data (n=141) are shown, namely: the Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Scale (*SWAN*), the Autism Spectrum Screening Questionnaire (*ASSQ*), the Social Communication Questionnaire (*SCQ*), The Columbia Impairment Scale (*CIS*), the WHO Disability Assessment Schedule (*WHODAS*) and the Inventory of Callous-Unemotional Traits (*ICU*).

Age Group	6		7		8		9		10		11		12		13		14		15		16		17-20	
Characteristic	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Age years	6.3	0.6	7.5	0.3	8.6	0.3	9.5	0.3	10.5	0.3	11.5	0.3	12.5	0.3	13.5	0.3	14.5	0.2	15.4	0.3	16.4	0.2	18.3	1.3
Sex (% female)	44	—	41	—	60	—	44	—	56	—	25	—	22	—	54	—	25	—	63	—	50	—	20	—
Handedness % right	89	—	88	—	83	—	94	—	94	—	95	—	78	—	92	—	100	—	88	—	67	—	100	—
Parental education	17.8	3.3	16.8	4.7	16.2	4.3	16.2	4.2	16.8	4.4	15.8	4.5	14.8	5.6	14.4	5.2	16.9	3.1	14.6	5.5	16	4.3	15.2	3.6
Sample size	9	—	17	—	25	—	16	—	16	—	20	—	9	—	13	—	8	—	8	—	6	—	10	—
Diagnosis (% healthy)	25	—	42	—	41	—	40	—	35	—	47	—	33	—	50	—	40	—	30	—	43	—	40	—
tSNR - rest	59.8	19	64.4	14.3	66.7	16.8	70.4	13.8	71	10.8	69.4	13.7	66.1	16.7	73.7	12.5	78	5.9	70.5	17.6	68.6	9.6	70.5	9.1
tSNR - movie	70.4	17.6	63.4	16.4	70.5	15.5	67.3	16.8	72.5	19.2	67.2	17.4	70.4	13.2	75.8	14	73.2	11.9	72.3	9.8	62	7.9	65.9	17.2
Frames % (rest)	86	15.8	90.5	10	90.8	12.2	94.4	8.2	96.8	4	94.8	7	94.6	9	95.6	7.3	99.1	2	99.5	0.5	97	3.6	97	7.8
Frames % (movie TP)	86.7	17.8	86.2	11.8	92.2	8.5	90	15.3	92.3	10.3	93.2	8	95	5.8	92.7	10.1	94.2	7.6	98.7	0.8	92.9	6.3	94.9	6.7
FD (rest)	0.19	0.02	0.48	0.41	0.24	0.1	0.29	0.28	0.24	0.22	0.19	0.12	0.22	0.13	0.19	0.09	0.15	0.08	0.13	0.02	0.12	0.05	0.14	0.07
FD (movie TP)	0.52	0.6	0.5	0.46	0.29	0.12	0.28	0.21	0.33	0.38	0.22	0.09	0.24	0.13	0.21	0.11	0.2	0.17	0.2	0.12	0.17	0.06	0.2	0.16
DVARS (rest)	1.03	0.03	1.17	0.17	1.05	0.05	1.06	0.09	1.06	0.1	1.04	0.06	1.06	0.07	1.03	0.03	1.03	0.06	1.01	0.02	1.02	0.02	1.02	0.02
DVARS (movie TP)	1.09	0.13	1.19	0.22	1.09	0.08	1.07	0.12	1.1	0.19	1.06	0.05	1.07	0.08	1.04	0.04	1.06	0.09	1.04	0.06	1.05	0.04	1.06	0.08
SWAN (total score)	0.5	0.7	0.1	0.9	0.3	0.9	-0.1	1.3	0.2	0.6	-0.2	1.1	0.5	1.5	0.03	0.4	-0.1	1.2	-0.2	1.1	-0.2	1.2	0.9	0.6
ASSQ (total score)	4.3	5.9	2.1	2.4	4.5	6.8	3.5	4.3	3.1	3.3	1.6	2.4	6.3	6.6	3.1	2.6	7.2	2.1	2.4	2.7	1.9	1.8	1.6	1.9
SCQ (total score)	2	1.4	4.3	3.3	4.5	2.8	6.7	5.6	6.3	5.4	5.3	4.1	7	5.1	4.5	2.9	8.3	5	5.5	2.3	5.6	2.8	7.1	3.3
CIS (parent report)	7	4.2	8	5.7	9.1	6.5	7.7	7.2	6.8	5.2	4.7	6.1	10.1	6.7	10.2	8.1	5.7	7.2	12.5	6.3	9.4	9.1	13.5	12.5
WHODAS (parent report)	5.2	6.9	11.1	19.8	11.3	10.7	8.1	8.6	4.8	8.5	5.1	6.7	13.5	15.2	4.6	5.9	12.1	18.5	1.3	2.6	8.9	7.4	7	9.8
ICU (parent report)	21.2	14.1	22.8	5.4	19.3	7.4	22	8.7	19.6	8.9	22.4	7.7	19.6	9	18.6	6.7	23.4	12.7	25.6	7.4	24.1	10.5	29.8	12.4

conducted supplementary analyses using more stringent FD values of 0.2 mm and 0.18 mm to examine the impact of FD on the predictive model. These cutoffs were chosen after examining the effects of age on head motion statistics, which include FD and DVARS measures (D refers to the temporal derivative of timecourses and VARS to the root mean square movement variance over voxels) (Power et al., 2012) (see **Supplementary Fig. S1** for visualizations of age-related effects on head motion). Next, the censored volumes were interpolated using a cubic spline function to preserve the overall BOLD frame number for each participant (Power et al., 2012, 2014). Participants were excluded if they had fewer than 40% frames remaining after scrubbing in any of the conditions (movie or rest). This criterion was determined separately for each condition to ensure that every BOLD run included in the study had at least 40% of frames retained. Next, we removed additional BOLD artifact using 'nuisance' regressors defined via subject-specific segmentation derived from FreeSurfer (Fischl, 2012). Specifically, we isolated subject-specific nuisance signals from the masks centered on the ventricles, deep white matter and mean gray matter. We computed a regression model such that the nuisance signal was used as an independent measure per each frame. The de-noised residual BOLD signal (i.e., after accounting for the nuisance covariation) was further band-pass filtered to retain frequencies in the range of 0.008-0.09Hz. In addition, we performed visual QC on BOLD images to ensure that all functional data were correctly mapped onto the CIFTI-generated cortical surfaces. BOLD temporal signal-to-noise (tSNR) ratio was calculated by computing the mean signal for all gray matter voxels across all BOLD frames divided by the standard deviation across all frames in a given BOLD run, but explicitly excluding all non-brain voxels (mean, SD and statistical tests across conditions for FD, DVARS, frames retained and tSNR can be found in **Supplementary Tables S2 and S3**). To isolate parcel-level and network-level signals we used the whole-brain CAB-NP parcellation (Ji et al., 2019) derived from the HCP atlas (Glasser et al., 2016), which includes 360 cortical parcels and 358 subcortical parcels assigned to 12 functional networks in CIFTI space. Parcel size and shape vary as a function of the alignment between functional and anatomical borders across imaging modalities and can be found in Glasser et al. (2016).

2.4. *Predictive modeling of state-specific multi-level neural effects (PrimeNet)*

Multi-level modeling and feature selection is outlined in Fig. 1. The model was iteratively trained on 80% of the sample (n = 125), while 20% of the sample was held out (n = 32). The prediction of the multi-level neural effects (PrimeNet) framework followed 3 sequential steps: i) Inference, ii) Training, and iii) Prediction. The **Inference** step identifies those edges whose mean FC significantly differs between movie-watching and rest FC matrices via a multi-level linear model (MLM). Selection of the final set of edges is performed by comparing the performance of a range of p-value thresholds via a support-vector machine (SVM) binary classifier with leave-one-out cross-validation (LOOCV). The **Training** step applies the top-performing p-value threshold to movie/rest FC matrices and trains an SVM classifier across iterations via LOOCV. The **Prediction** step applies the trained SVM model to the held-out dataset and evaluates its performance using receiver operating characteristic (ROC) curves. Additionally, this step calculates the area under the curve (AUC) and accuracy (ACC) for each classifier output. The default parameter in PrimeNet is to run 1000 iterations of the training and prediction step to assess the predictive value of the model. This section will discuss each of these steps in detail.

In the present study, we first input the FC matrices for 80% of the sample into an MLM as part of the **Inference** step. As explanatory independent variables we included state (movie or rest), age, diagnosis (no diagnosis or presence of any DSM [Diagnostic and Statistical Manual of Mental Disorders] diagnosis), temporal SNR and percentage of frames analyzed per subject. We additionally ran supplementary analyses with FD and DVARS as covariates to comprehensively account for the effects of head motion. As random effects, we included intercepts for subjects, as well as by-subject random slopes for the effect of state to account for within-subject variation between rest and movie FC matrices. In order to establish the cutoff for selection of edges that are predictive of movie vs rest, we computed the MLM model sweeping across a range of p-values from Bonferroni corrected (1.9×10^{-7}) to 0.05 (from 1.9×10^{-7} to 3×10^{-7} in increments of 1×10^{-9}, from 3×10^{-7} to 0.0001 in increments of 1×10^{-7}, and from 0.0001 to 0.05 in increments of 0.0001). To assess each p-value performance on the training dataset, we used a single-feature SVM binary classifier with a linear kernel. Specifically, for each subject, Fischer's r-to-z values (Fz) for the selected edges were summed into a single value per scan, which resulted in two values per subject (one for the movie scan and one for the rest scan) (a schematic of PrimeNet step for identification of predictive edges is provided in **Supplementary Fig. S2**). These summary values served as input into the SVM classifier. For each classifier run, we performed ten-fold cross-validation across a range of values of the cost parameter (0.001,0.01,0.1,1,5,10,100) and selected the cost with the lowest cross-validation error rate following the recommendations outlined in (James et al., 2013) (see **Supplementary Fig. S3** for a schematic of the SVM workflow). The top performing p-value was retained and used to select the predictive edges of movie versus rest.

Next, in the **Training** step, we used an SVM binary classifier to test whether whole-brain FC patterns can be used to distinguish between rest and movie scans. First, we ran 1,000 iterations of the MLM. For each iteration of the model, a subset of subjects (80%) was randomly selected and the corresponding movie and rest FC matrices were used as input to the MLM. After running the MLM, we masked each subject's FC matrices corresponding to the rest and movie scans with the edges that survived after applying the top-performing p-value threshold. Next, for each subject, Fz values corresponding to the predictive edges were summed into a single summary value for each movie/rest scan. These summary values were in turn used to train an SVM classifier to assess predictive accuracy of the selected edges. Specifically, we computed 1,000 cross-validation runs of the SVM classifier on the training dataset using LOOCV and following the kernel and cost parameters noted above. In each repetition of the LOOCV, summary values from a single subject (one rest scan and one movie scan) were used as the validation data and the remaining data were used as training. The classifier output consisted of the state-specific label (rest or movie) for each of the observations in the training dataset.

This SVM model was then tested on the held-out dataset (20% of the sample, randomly selected in each iteration) as part of the **Prediction** step. We employed ROC curves to quantify the performance of the SVM classifier. The ROC curve plots the true positive rate (TPR or sensitivity) against the false positive rate (FPR or 1-specificity) at various thresholds. The TPR or sensitivity is calculated as the proportion of scans correctly classified as movie-watching scans. The FPR (1-specificity) is calculated as the proportion of resting-state scans incorrectly classified as movie-watching scans. Additionally, we calculated the AUC, the area between the ROC curve and the x-axis, which can be interpreted as the probability that the classifier will accurately classify a randomly drawn pair of scans. We also calculated accuracy of the classifier for all 1,000 iterations of the classifier, which corresponds to the ratio of the number of movie/rest correct predictions to the total number of input samples. Finally, we identified the predictive edges of brain state (movie or rest) by visualizing those edges that survived across 1,000 iterations of the model. We refer to this set of edges as *state-specific predictive edges*.

2.5. *Network-level differences between movie & rest*

We calculated mean FC (Fz) values for all parcels predictive of state (movie vs. rest) and associated networks using the CAB-NP parcellation (Glasser et al., 2016; Ji et al., 2019). Next, to qualitatively assess differences across these predictive network clusters between movie and rest FC matrices, we computed one-way analyses of variance (ANOVA)

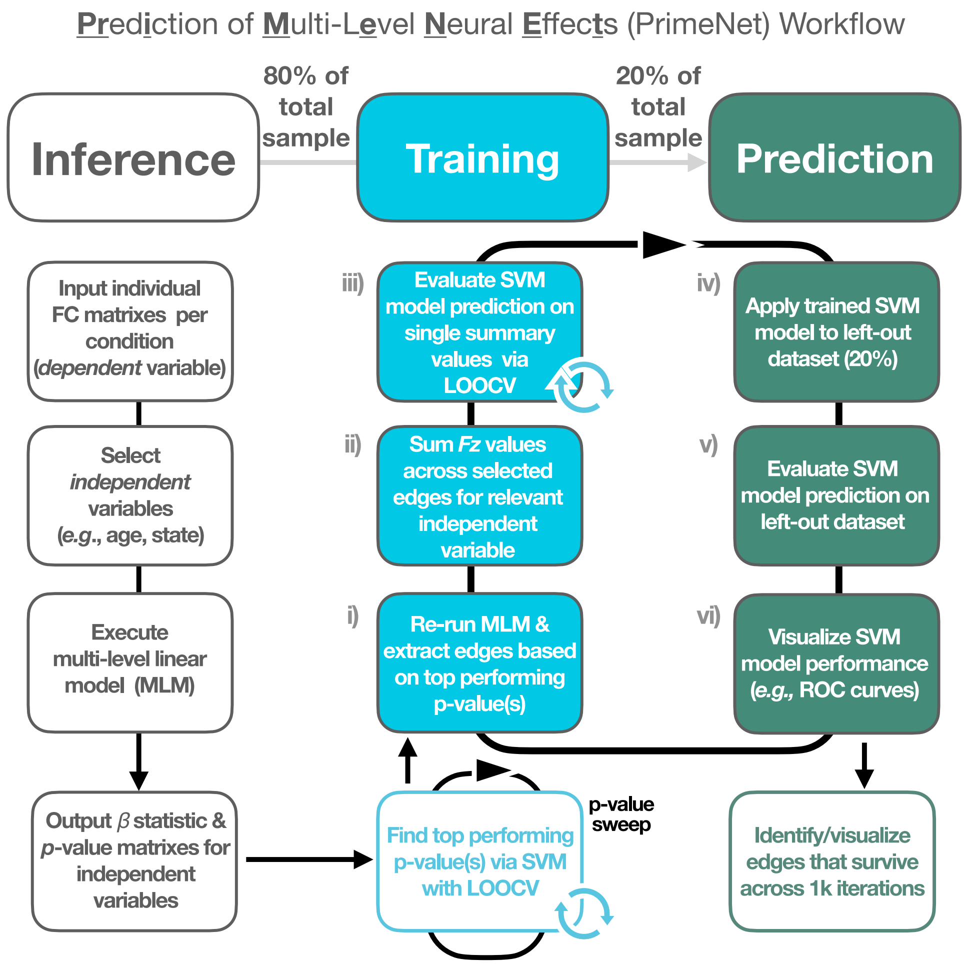

Fig. 1. PrimeNet Workflow. **Inference.** FC=Functional connectivity, LOOCV=Leave-one-out cross-validation, MLM=Multi-level model, SVM=Support-vector machine, ROC curve=Receiver Operating Characteristic curve. FC matrices for the training dataset (80% of the sample) are input into a multi-level model (MLM) for each condition (e.g., movie-watching and resting-state). Next, the selected explanatory independent variables (predictors) and covariates are input into the model. Two FC matrices are input per subject: movie-watching and resting-state FC matrices. The output of the MLM is two matrices per independent variable with either a ß statistic or p-value per edge in the FC matrices. Identification of the p-value cutoff for selection of predictive edges is performed by first thresholding the matrices iteratively across p-values ranging from Bonferroni corrected 1.9e-07 – 0.05 (p-value sweep). For each p-value threshold, a predictive Support-Vector Machine (SVM) model with Leave-One-Out Cross-Validation (LOOCV) is built using the selected edges on the training dataset. The predictive value of each p-value threshold is assessed via ROC curves, model accuracy and area under the curve (AUC). **Training.** The top performing p-value(s) is retained and the MLM is re-run across 1K iterations. In each iteration, the selected p-value threshold is used to identify the predictive edges. Next, FC values for the predictive edges are summed into a single summary value for each rest/movie scan. A binary predictive model is then built via SVM for the selected edges and cross-validated using LOOCV. The output is the predictive state category (movie or rest) for each observation in the training dataset. **Prediction.** In each iteration, the trained SVM model is applied to the left-out dataset (20% of the original sample) after calculating single summary values per scan in the test dataset. The output is the predictive state category for each scan in the test dataset. The SVM model prediction is evaluated via ROC curves. Mean accuracy, mean AUC and corresponding visualizations are created. The final set of maximally predictive edges is computed via identification of edges that survive across 1K iterations of the model.

for the effect of state (movie versus rest) on mean FC for each of the predictive networks.

2.6. Age-related effects: movie versus rest prediction accuracy and network-level FC differences

To examine accuracy of the predictive model across age, we first calculated mean predictive accuracy for all age groups in both training and test datasets across 1,000 iterations of PrimeNet. Second, to examine the relationship between age and mean FC in the predictive networks, we computed linear, quadratic and cubic regressions with age as predictor and mean FC as dependent variable for each predictive network cluster (Madhyastha et al., 2018). As covariates, we added diagnosis status (no diagnosis or diagnosed), site, FD and DVARS. Finally, we visualized mean FC for each of the predictive networks in the youngest and oldest groups of subjects to highlight FC differences across age for the predictive parcels and associated networks.

The data that support the findings of this study are openly available in 1,000 Functional Connectomes Project and its International Neuroimaging Data-sharing Initiative (FCP/INDI), which can be found at http://fcon_1000.projects.nitrc.org/indi/cmi_healthy_brain_network/. We are currently developing a manuscript describing the PrimeNet framework and anticipate that the code will be released to the wider community shortly after publication. In the meantime, the code is available from the corresponding authors upon reasonable request.

3. Results

3.1. Predictive modeling of state-specific (movie vs rest) multi-level neural effects

First, we examine whether there is a set of FC features that differentiate between naturalistic viewing (i.e., movie) and rest. The presence of such a feature set would provide evidence for state-specific FC patterns in a developmental sample.

Fig. 2A shows the unthresholded FC matrix from the subtraction movie-rest with Fz values shown for all edges for the training dataset using 80% of the original sample ($n = 125$). The matrix shows that movie-watching modulates brain-wide FC, particularly by reducing within-network FC (e.g., visual 1 [V1]), whereas it seems to elevate coupling of specific networks (e.g., default network with visual 2 [V2]) (FC matrices for movie-watching and rest BOLD acquisitions separately are given in **Supplementary Fig. S4**). Additionally, we visualized FC for a single V1 parcel in both the unthresholded and Bonferroni thresholded FC matrices to corroborate these results (Figs. 2B & C) (FC visualizations with FD motion cutoffs of 0.2 mm and 0.18 mm can be found in **Supplementary Figs. S5 & S6**). To identify FC predictive features of movie versus rest, we ran an MLM per edge of the FC matrix using FC values from both rest and movie scans as dependent variable and state (rest or movie) as the primary independent variable. As covariates, we included age, diagnosis status (no diagnosis or diagnosed) and

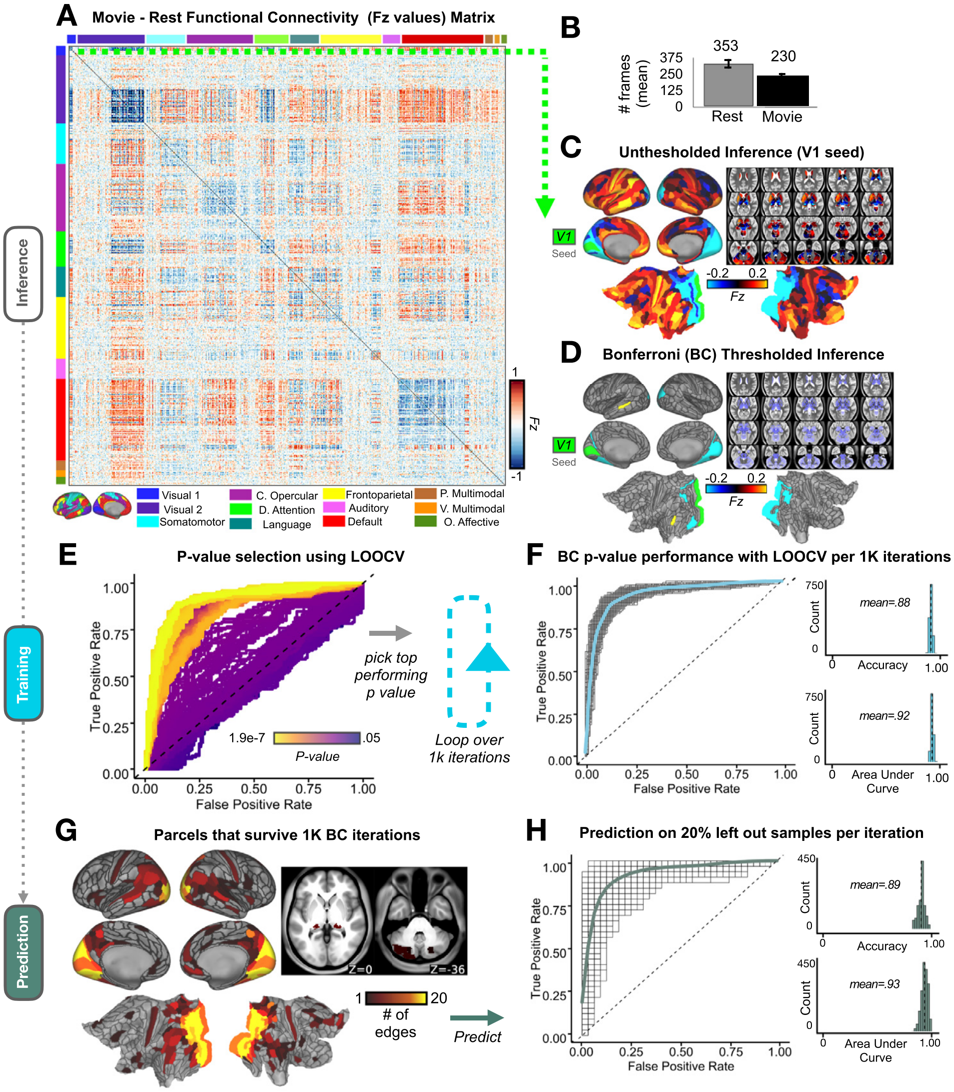

Fig. 2. Brain-wide Differences in Functional Connectivity Between Movie and Rest Across Development with the Movie 'The Present'. BC=Bonferroni corrected, LOOCV=Leave-one-out cross-validation. **A.** Unthresholded functional connectivity (FC) matrix (movie-rest difference) derived from the CAB-NP parcellation with Fz values shown for all edges. The matrix highlights the differences between movie-watching and rest in the training dataset (n=125, 80% of the original sample). **B.** Number of frames retained after scrubbing for resting-state and movie-watching conditions. Error bars reflect standard deviation. **C.** FC brain map seeded from the visual 1 network (V1) with Fz values shown for all parcels. **D.** FC map seeded from V1 showing only those parcels that survive at Bonferroni corrected (BC) p-value (p-value=1.9×10^{-7}). **E.** ROC curves quantifying performance of the support-vector machine (SVM) binary classifier after sweeping across a range of p-values from 1.9×10^{-7} – 0.05 to identify the top-performing threshold for edge selection. **F.** ROC curves quantifying the performance of the training classifier using a BC p-value threshold of p=1.9×10^{-7}. Each ROC curve represents a single SVM iteration using LOOCV for a total of 1K iterations. The solid blue line represents the average across the 1K iterations. We calculated the area under the curve (AUC) (i.e., the area between the ROC curve and the x-axis) and the accuracy (ACC) of the SVM classifier for all 1K iterations of the classifier. **G.** FC parcels selected by the training classifier as being predictive of movie-watching versus rest after applying BC p-value threshold. Brain areas displayed survived across 1K iterations of the model. Color gradient represents the number of edges per parcel (from 1 to 20 edges). **H.** ROC curves, AUC and ACC distributions quantifying the performance of the classifier on the left-out datasets (n=32, 20% of the original sample) using the selected p-value threshold (p=1.9e7). Each ROC curve represents a single SVM iteration using LOOCV for a total of 1K iterations. The solid green line represents the average across the 1K iterations. Note that the training and left out datasets were always randomly selected for each iteration of the model.

head motion statistics (FD, DVARS, % frames, tSNR; see Methods for a complete description), along with corresponding interactions. Next, threshold for edge selection was established by running an SVM binary classifier with a linear kernel on the training dataset and sweeping across a range of p-value cutoffs from Bonferroni corrected 1.9×10^{-7} to 0.05. The classifier trained on the Bonferroni corrected 1.9×10^{-7} p-value yielded the highest performance (accuracy = 87%, sensitivity = 88%, specificity = 86.4%, AUC = 0.91) (Fig. 2D).

Having established a p-value threshold for edge selection, we ran 1,000 iterations of the MLM and used Bonferroni corrected p-value 1.9×10^{-7} to select the relevant edges for each iteration. Next, we used an SVM classifier to test if movie-watching and rest states can be predicted from FC features. Specifically, the features selected in each of the 1,000 iterations were used to train single-feature SVM binary classifiers using LOOCV (mean accuracy=88%, mean sensitivity = 87.6%, mean specificity = 87.8%, AUC = 0.92, Fig. 2E). Finally, we identified the edges that survived across all 1,000 iterations of the model from the main effect of state (movie versus rest) to qualitatively assess the differences between movie and rest (Fig. 2F). The final set of predictive edges consisted of 154 parcels associated with ten functional networks (auditory [5 parcels], cingulo-opercular [13 parcels], default-mode [33 parcels], dorsal attention [9 parcels], frontoparietal [9 parcels], language [11 parcels], posterior-multimodal [8 parcels], somatomotor [8 parcels], V1 [6 parcels] and V2 [52 parcels]) (Fig. 2G).

We applied this cross-validated trained model to the held-out test datasets (20% of the sample, n = 32) across the 1,000 iterations (mean accuracy = 89%, mean sensitivity = 88.6%, mean specificity = 89%, AUC = 0.93, Fig. 2H). Importantly, training datasets (80% of the sample) and test datasets were randomly selected for each iteration of the model. These results indicate that the classifiers yielded highly comparable performance in both the training and test datasets across iterations. In addition, and to evaluate the effect of FD cutoff on the predictive model, we ran 1,000 iterations of PrimeNet with the same data and FD cutoff <0.2 mm. After excluding subjects with <40% of frames retained, the final dataset consisted of 126 subjects. We observed a slightly higher predictive accuracy in both training (mean accuracy=91.7%, mean sensitivity=92.3%, mean specificity = 91%, AUC = 0.95, **Fig. S6**) and testing (mean accuracy = 90.3%, mean sensitivity = 89.9%, mean specificity = 90.4%, AUC = 0.94) datasets relative to the results obtained with FD<0.5 mm. The final set of predictive edges mapped onto a more restricted set of parcels that overlapped with the set of FC features identified with a scrubbing criterion of FD <0.5 mm. This set of parcels consisted of 30 parcels, which were associated with five functional networks

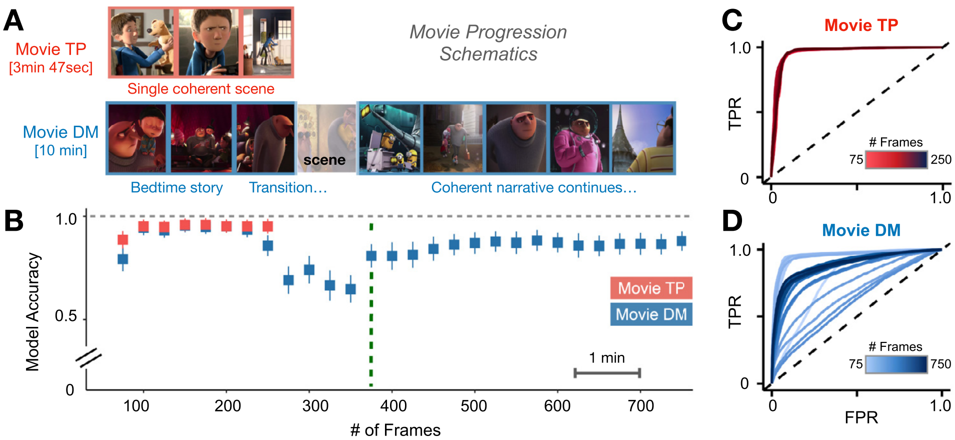

Fig. 3. Predictive Accuracy Across Movies as a Function of Movie Scan Length. Movie TP=Movie "The Present", Movie DM=Movie "Despicable Me", TPR=True Positive Rate, FPR=False Positive Rate. **A.** Screenshots of the first scene in each truncated BOLD run included in the analyses starting at 75 frames (1 min) for each of the movies. BOLD runs were truncated every 25 frames (20 sec). **B.** Predictive accuracy across PrimeNet runs as a function of the number of BOLD frames from each of the movies included in the analysis. Each data point represents mean predictive accuracy after running PrimeNet across 500 iterations for the corresponding truncated BOLD run and the full resting-state run (375 frames, 5 min). Error bars reflect standard deviation. The vertical green dashed line indicates the point at which movie and rest conditions have equal number of frames. **C.** ROC curves quantifying the performance of the classifier with data from the movie TP. Each ROC curve represents the average accuracy across 500 iterations for a given truncated movie BOLD run and the full resting-state run. **D.** ROC curves quantifying the performance of the classifier with data from the movie DM. Each ROC curve represents the average accuracy across 500 iterations for a given truncated movie BOLD run and the full resting-state run.

(frontoparietal [1 parcel], language [1 parcel], posterior-multimodal [1 parcel], V1 [3 parcels] and V2 [24 parcels]).

We conducted additional analyses with a second longer movie ('Despicable Me', 10 min) to examine i) generalizability of predictive accuracy across movies and ii) the effect of the number of BOLD frames (i.e., scan length) on state prediction. For these analyses, we used a subset of the data (n = 78) from the initial sample, which consisted of data from those participants who had also completed a movie-watching session with the movie "Despicable Me" and whose BOLD data passed the established preprocessing criteria.

To assess whether predictive accuracy of FC features generalizes across movies, we employed the model trained on the shorter 3.47 min movie 'The Present' (movie TP) to predict brain state on data from the longer movie 'Despicable Me' (movie DM). We first ran PrimeNet with the movie TP and the same resting-state data to confirm no differences in the results due to the smaller sample size (training dataset: mean accuracy = 94.8%, mean sensitivity = 97.1%, mean specificity = 92.4%, AUC = 0.965; testing dataset: mean accuracy = 94.9%, mean sensitivity = 96.4%, mean specificity = 93.4%, AUC = 0.97). The final set of predictive edges mapped onto 57 parcels associated with seven functional networks (default-mode [2 parcels], dorsal attention [3 parcels], frontoparietal [2 parcels], language [2 parcels], posterior-multimodal [2 parcels], V1 [3 parcels] and V2 [42 parcels]). Next, we applied the trained model to data from the movie DM. The results showed similar, albeit slightly lower, predictive accuracy (mean accuracy = 90%, mean sensitivity = 87%, mean specificity = 91%, AUC = 0.93) for this dataset.

Additionally, we ran 1,000 iterations of PrimeNet using only the movie DM and the same 5 min resting-state session. All three steps (Inference, Training and Testing) were implemented with the exact same parameters. We observed a slightly lower mean predictive accuracy with this movie for both the training (80% of sample, n = 63) (mean accuracy = 87%, mean sensitivity = 83%, mean specificity = 89%, AUC = 0.91) and test datasets (20% of sample, n = 15) (mean accuracy = 88%, mean sensitivity = 85%, mean specificity = 91%, AUC = 0.91). Relative to the set of predictive FC features previously identified with the movie TP, the final set of features showed 82% overlap. Specifically, the set of predictive features across the two movies shared 46 parcels, which were associated with six functional networks (shared predictive parcels across movies: default-mode [2 parcels], dorsal attention [1 parcel], frontoparietal [1 parcel], language [1 parcel], posterior-multimodal [1 parcel], V1 [1 parcel] and V2 [39 parcels]) (FC visualizations can be found in **Supplementary Fig. S7**).

Finally, to examine whether the number of frames input into the model impacts predictability of brain state, we truncated the movie BOLD runs iteratively in increments of 25 frames (20 sec) by starting at 75 frames (1 min) and extending the movie to its full length. For each truncated BOLD run, we computed 500 iterations of PrimeNet using the corresponding truncated movie scan and the full resting-state scan (375 frames). This process was implemented separately for each movie. We report here the results of applying the models trained with the truncated BOLD runs to the test datasets across the two movies. For the shorter movie (movie TP), we observe a predictive accuracy of 89% at 75 frames, which increases to 95% at 100 frames and remains stable across the remaining BOLD run (mean accuracy: 95.3) (Fig. 3A-B). For the longer movie (movie DM), we observe a lower predictive accuracy at 75 frames (79%), which increases to 94.3% at 100 frames and remains relatively stable until 225 frames (mean accuracy: 94%) (Fig. 3A-B). Crucially, in the longer movie, predictive accuracy starts to decrease at 250 frames with the lowest accuracy achieved at 350 frames (65%), after which there is a relatively steady increase in predictive accuracy until the end of the run as a function of the number of frames included in the model. This drop in accuracy between 250–350 frames coincides with two critical moments in the movie: i) the bedtime story that started the movie clip ends at 250 frames and ii) there is a one-minute scene between frames 275–350 in which a new character appears and the conversation appears decontextualized with references to earlier scenes of the movie (not shown in the clip) that the participants need to infer. Mean ROC curves across iterations for each truncated movie BOLD run are shown in Figs. 3C-D for movie TP and movie DM respectively. Additionally, we conducted the same analysis with the truncated movie BOLD runs, but this time using a longer resting-state run, which was the result of concatenating the two resting-state runs (10 min total, 750 frames). Specifically, we used the same dataset (n = 78), but also included data corresponding to the second resting-state session for the subset of participants who participated in this session and whose resting-state BOLD run passed the established preprocessing criteria (n = 58). The results showed a similar trend for both movies and significantly lower mean predictive accuracy for the movie TP when the concatenated resting-state scan was used for state prediction (p < 0.001). No differences in predictive accuracy were observed for the movie DM when the second resting-state was added to the model (p = 0.11) (**Supplementary Figs. S8-S9**).

3.2. Network-level differences between movie & rest

Having shown that FC features reliably distinguish between movie and rest and generalize across movies, we now turn to the predictive network areas (i.e., network clusters) that are associated with this effect.

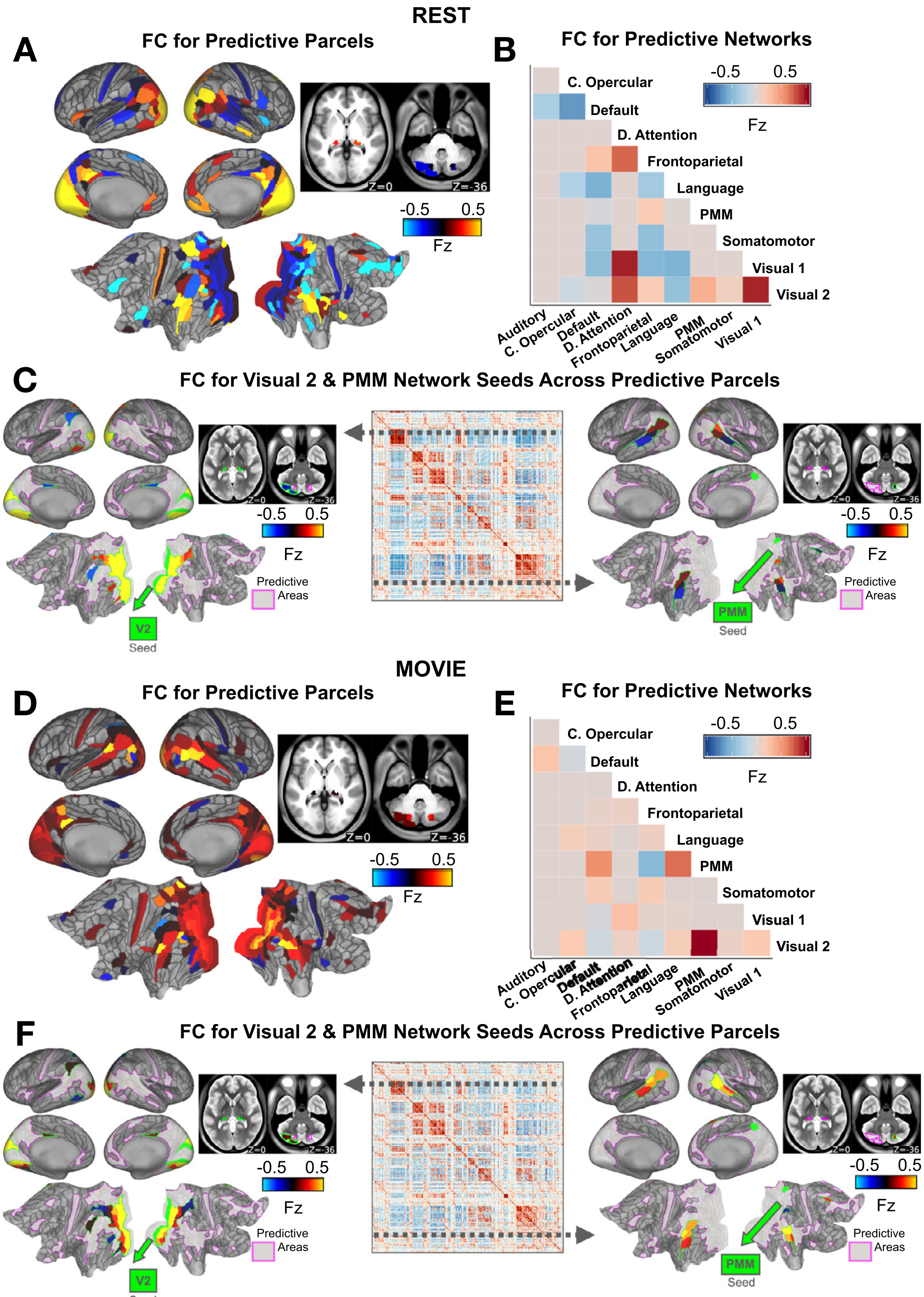

Fig. 4. Characterizing Network-Level Functional Connectivity During Movie & Rest. **Rest**. FC=Functional connectivity, V2=Visual 2 network, PMM=Posterior-multimodal network. **A.** Resting-state FC brain map with mean Fz values shown for all predictive parcels. **B.** Thresholded FC matrix with Fz values for all predictive networks. **C.** *Left:* FC brain map seeded from a V2 parcel with Fz values for all predictive edges. *Right*: FC map seeded from a PMM parcel with Fz values for all predictive edges. *Middle*: Unthresholded FC matrix with Fz values shown for all edges. The gray dotted lines indicate parcel selection for the V2 and PMM networks. In the brain maps, the total number of predictive parcels is outlined in purple. **Movie. D.** Movie-watching FC brain map with mean Fz values shown for all predictive parcels. **E.** Thresholded FC matrix with Fz values for all predictive networks. **F.** *Left*: FC brain map seeded from a V2 parcel with Fz values for all predictive edges. *Right*: FC map seeded from a PMM parcel with Fz values for all predictive edges. *Middle*: Unthresholded FC matrix with Fz values shown for all edges. The gray dotted lines indicate parcel selection for the V2 and PMM networks. In the brain maps, total number of predictive parcels is outlined in purple.

Specifically, we characterize mean FC across parcels and networks associated with the edges (i.e., brain connections) identified as predictive of state (movie versus rest).

We first examine mean FC across predictive parcels/networks during rest and movie-watching independently. Fig. 4A shows an FC map for the resting-state condition with mean Fz values for each of the predictive parcels, which reveals strong positive FC of visual and default networks. Network-level FC visualization (Fig. 4B) reveals positive coupling of visual networks (V1 and V2) and of V1 and dorsal-attention networks. Seed-based FC brain maps allow exploration of mean FC for a given parcel with all other predictive parcels. For visualization purposes, we focus on parcels from a sensory network (V2) and a higher-order cognitive network (posterior-multimodal [PMM]), which can be visualized on the unthresholded FC matrix in Fig. 4C. The V2-seeded brain map reveals strong FC between V1 and V2, whereas the PMM-seeded brain map shows primarily low FC with areas of the language network. As for the movie-watching condition, an FC brain map with mean Fz values for each of the predictive parcels (Fig. 4D) shows that movie-watching modulates FC of networks associated with higher-order cognitive functioning (language and PMM). Between-network FC reveals strong coupling between PMM and language/V2 networks (Fig. 4E). Coupling between PMM and language/V2 parcels can also be observed in the PMM-seeded brain map shown in Fig. 4F.

Next, we focus on differences in FC between movie and rest across networks. We analyzed mean FC for the movie-rest difference across all parcels/networks associated with the identified state-specific predictive edges. In Fig. 5A, an FC brain map with mean Fz values for each predictive parcel reveals stronger positive FC in language network areas during movie-watching relative to rest. In contrast, resting-state is characterized by increased FC in areas of the visual and default-mode networks.

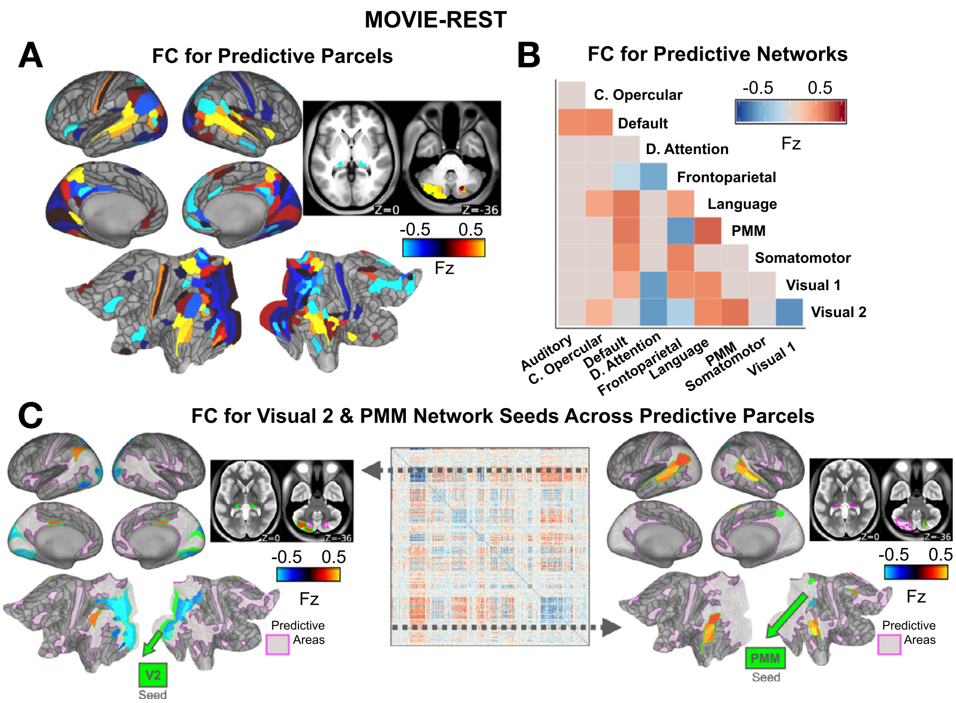

Fig. 5. Characterizing Network-Level Functional Connectivity for the Movie-Rest Difference. FC=Functional connectivity, V2=Visual 2 network, PMM=Posterior-multimodal network. **A.** Movie-Rest FC brain map with mean Fz values shown for all predictive parcels. **B.** Thresholded FC matrix with Fz values for all predictive networks **.C.** *Left*: FC brain map seeded from a V2 parcel with Fz values for all predictive edges. *Right*: FC map seeded from a PMM parcel with Fz values for all predictive edges. *Middle*: Unthresholded FC matrix with movie-rest Fz values for all edges. The gray dotted lines indicate parcel selection for the visual (V2) and posterior-multimodal (PMM) networks. In the brain maps, the total number of predictive parcels is outlined in purple.

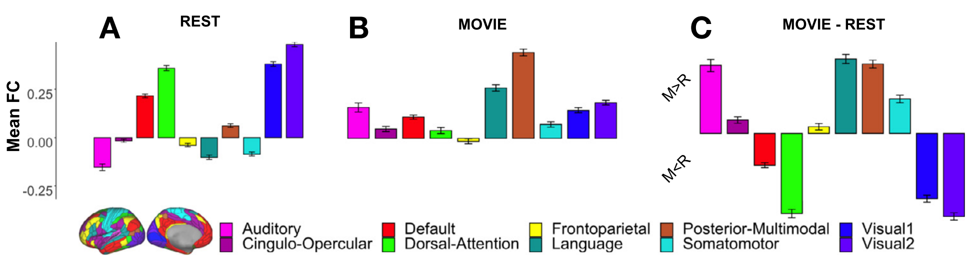

Fig. 6. Within-Network Functional Connectivity Differences Between Movie & Rest. FC=Functional connectivity, M=Movie, R=Rest. Error bars reflect standard error of the mean. **A.** Mean FC (Fz) across predictive network clusters during rest. **B.** Mean FC (Fz) across predictive network clusters during movie. **C.** Mean FC (Fz) across predictive network clusters for the movie-rest difference.

A between-network FC matrix visualization shows a stronger coupling of language and PMM networks during movie-watching relative to rest (Fig. 5B). There is also a stronger coupling of dorsal-attention network clusters with visual networks (V1 and V2) during resting-state relative to movie-watching. In addition, seed-based maps indicate stronger within-network FC in V2 during rest relative to movie-watching (Fig. 5C). In contrast, PMM and language networks show stronger between-network FC during movie-watching relative to rest (Fig. 5C). Seed-based FC maps for additional movie-rest predictive network clusters are provided in **Supplementary Fig. S10**.

An overall ANOVA on the effect of brain state (movie versus rest) on FC for each of the ten predictive network clusters showed that movie-watching is characterized by increased FC in auditory, language and PMM networks relative to rest (auditory: $F(1) = 127.7$, $p < 0.001$; language: $F(1) = 280.1$, $p < 0.001$; PMM: $F(1) = 339.2$, $p < 0.001$), whereas resting-state is characterized by increased visual and dorsal-attention FC relative to movie-watching (V1: $F(1) = 348.5$, $p < 0.001$; V2: $F(1) = 522.8$, $p < 0.001$); dorsal-attention: $F(1) = 355.2$, $p < 0.001$. These results are illustrated in Fig. 6 (ANOVA results for each of the predictive network clusters can be found in **Supplementary Table S4**).

3.3. Age-related effects: movie versus rest prediction accuracy and network-level differences

The results described so far reveal consistent differences in FC during movie-watching versus rest across childhood, adolescence, and early adulthood. Additionally, the results indicate that distinct functional network clusters exhibit robust differences in widespread coupling across the two states. Next, we characterize performance of the predictive model by age group and examine whether there are differences in resting-state and movie-watching FC across predictive parcels/networks as a function of age.

First, we visualize mean predictive accuracy across 1000 iterations of PrimeNet for both training and test datasets as a function of age (Fig. 7). State-specific predictive accuracy increases with age for both movie-watching and rest states. The results of a mixed model ANOVA with a between-subject factor of age group (i.e., 6–9,10–15,16–20) and within-subject brain state (i.e., rest versus movie) showed a trend for an interaction between age group and brain state ($F(2,154 = 2.78$, $p = 0.065$). Individual ANOVAs by age group showed that brain state predictions were more accurate for the youngest children (6 to 9-year-olds) when using movie-watching data relative to resting-state data ($p = 0.019$). In contrast, model predictions showed no differences in predictive accuracy between rest and movie-watching data in early adolescents (10 to 15-year-olds, $p = 0.79$) and late adolescents/young adults (16 to 20-year-olds, $p = 0.23$).

To examine age-related changes in mean state-specific (movie versus rest) FC across predicted parcels and network clusters, we computed multiple linear regressions for each of the ten network clusters predictive of brain state. We included age as a predictor and mean movie-rest FC as the dependent variable. As covariates, we included diagnosis status (no diagnosis or diagnosed), site, mean FD and mean DVARS. We ran two additional multiple regression models with age either as a quadratic or cubic term. The results revealed age-effects across four

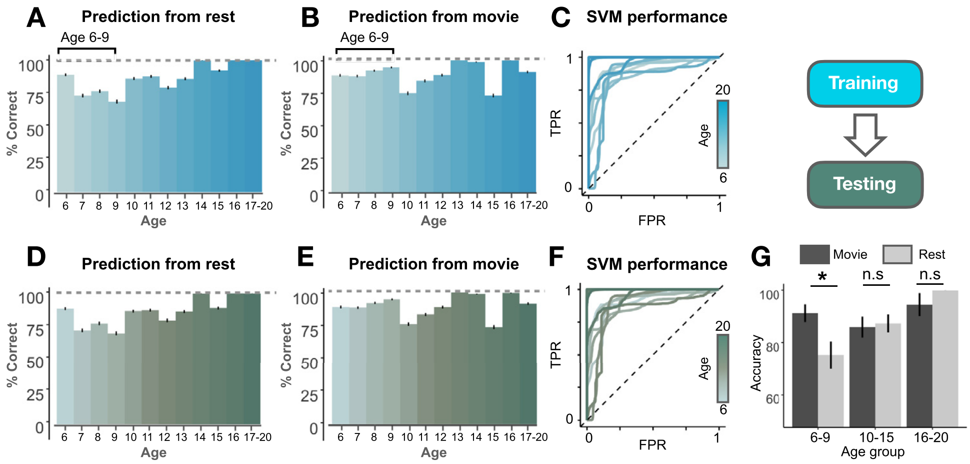

Fig. 7. Prediction Accuracy of Brain State Across Age. A, B. Prediction accuracy for movie-watching and resting-state scans in the training dataset across 1K iterations of PrimeNet. Error bars reflect standard error of the mean. **D, E** Prediction accuracy for both movie-watching and resting-state scans in the testing dataset across 1K iterations of PrimeNet**.** Error bars reflect standard error of the mean. **C, F.** ROC curves quantifying the performance of the classifier on the training (C) and test datasets (F). Each ROC curve represents an age group from 6 to 20 years of age. Note that participants older than 17 are collapsed into a single group to achieve comparable sample sizes across age groups. **G.** Mean predictive accuracy across age groups for movie and rest scans for the test dataset across 1K iterations of PrimeNet. Error bars reflect standard error of the mean.

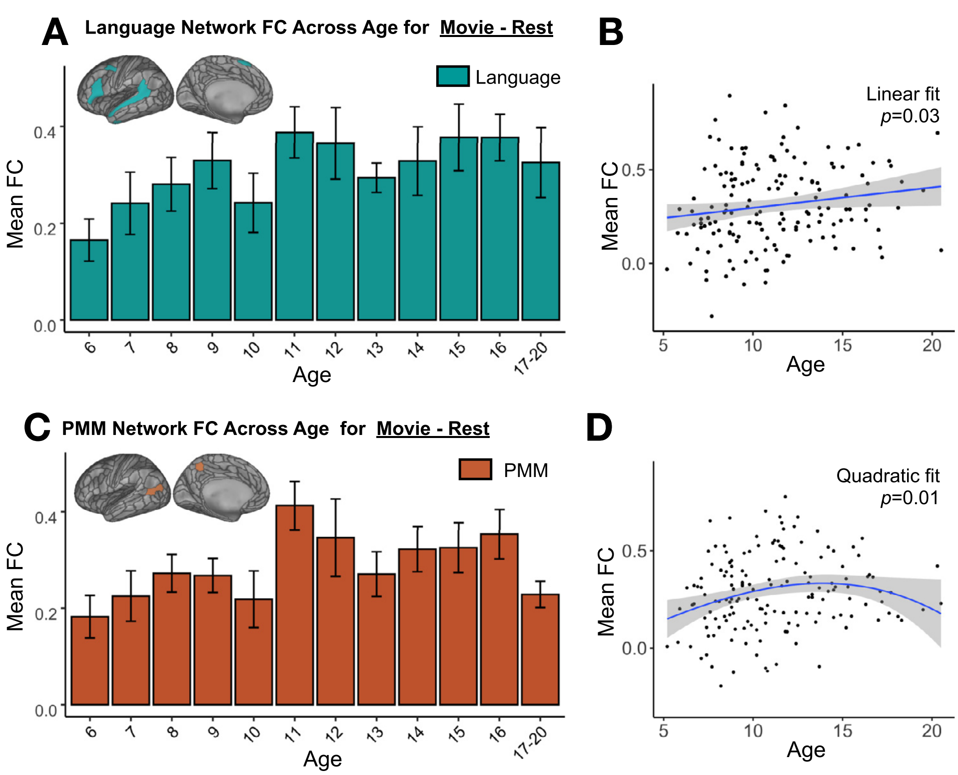

Fig. 8. Network-Level Functional Connectivity Across Age. FC=Functional connectivity, PMM=Posterior-multimodal. **A, C.** Mean movie-rest FC across age groups for all predictive parcels of the language (A) and PMM network clusters (C). Note that participants older than 17 are collapsed into a single group to achieve comparable sample sizes across age groups. Error bars reflect standard error of the mean. **B, D.** Mean movie-rest FC per subject for the language (B) and PMM (D) network clusters with lines of best fit and corresponding p-values.

network clusters: auditory, language, PMM and V2. To illustrate these effects, we focus on two networks that show a strong coupling between them: the language and PMM networks. A linear regression model provided the best fit for the language network ($p = 0.03$) (Fig. 8B) whereas a quadratic model yielded the best fit for the PMM network ($p = 0.01$) (Fig. 8D) (results for all networks are provided in **Supplementary Table S5;** visualizations of best fits for auditory and V2 networks are provided in **Supplementary Fig. S11**). Importantly, these effects hold after controlling for potential covariates (i.e., diagnosis status, site, FD and DVARS). Changes in mean FC across age for the language and PMM network clusters are visualized in Figs. 8A & 8C respectively. The lowest mean FC values were observed in the groups of 6- and 7-year-olds. Qualitatively, state-specific (movie-rest) FC across the language and PMM network clusters exhibits a steady increase between 6 and 9 years of age and reaches a peak in the group of 11-year-olds (mean FC for all predictive networks across age can be found in **Supplementary Fig. S12**).

Finally, to assess FC changes in naturalistic viewing across age, we visualized mean movie-rest FC in the youngest and oldest participants: 6–7-year-olds ($n = 26$) and 16-20-year-olds ($n = 16$) (FC brain maps for all age groups are provided in **Supplementary Fig. S13**). We observe key differences between these two groups with respect to their within- and between-network mean FC. First, overall FC across predictive parcels is stronger in late adolescents/young adults relative to children. This is observed particularly in areas of the language, V1 and V2 networks. However, a few areas in the somatomotor network reveal stronger mean FC in children (Figs. 9A & 9D). Second, adolescents/young adults show a stronger coupling of the language network cluster with PMM and default networks relative to mean FC in children (Figs. 9B & 9E). Brain maps seeded from the language and PMM networks provide further visualization of the coupling between these two networks (Figs. 9C-D & F-G). In contrast, children exhibit a stronger coupling between areas located in the frontoparietal and dorsal-attention networks relative to adolescents/young adults.

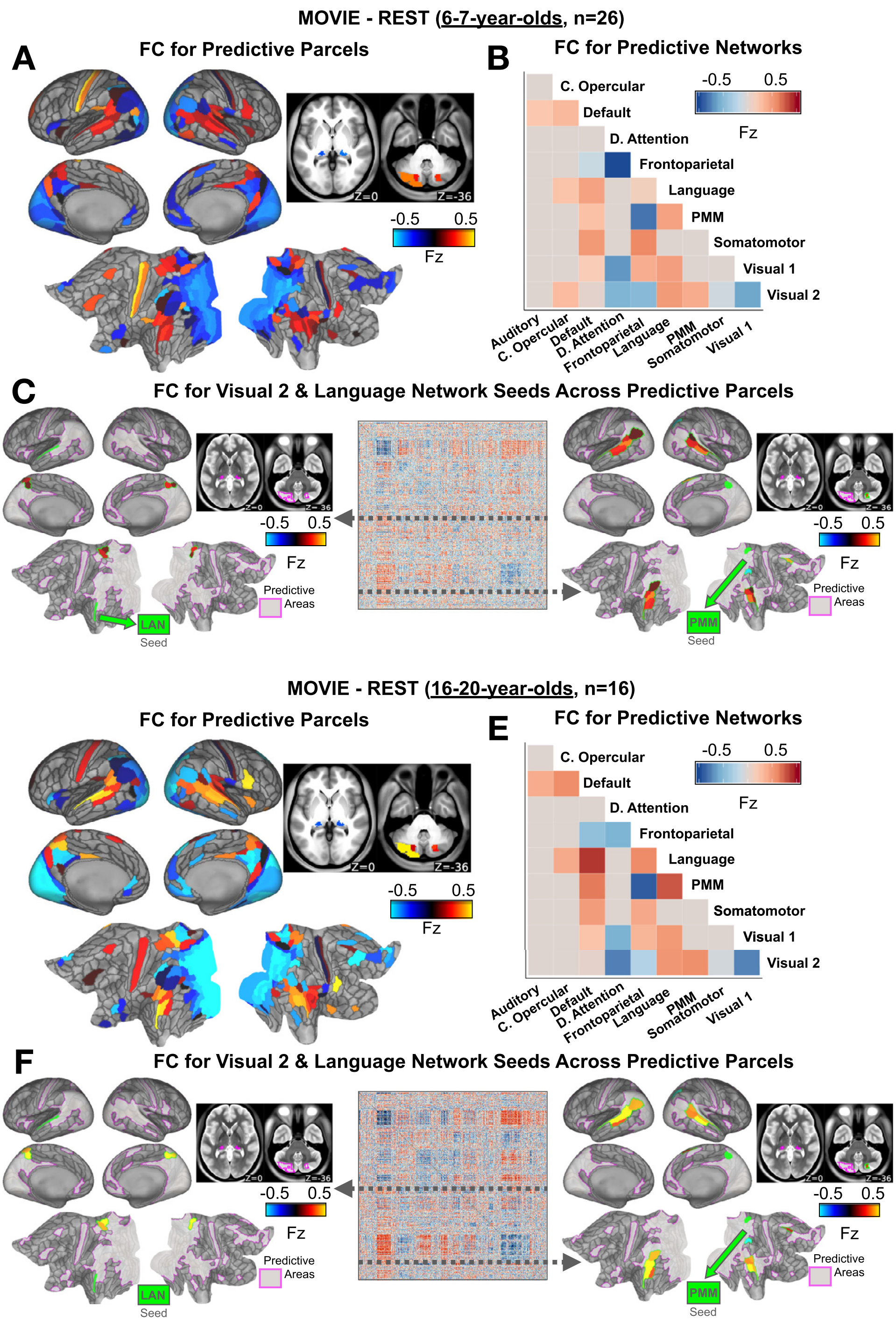

Fig. 9. Characterizing FC Differences Between Movie & Rest Across Age. FC=Functional connectivity, LAN=Language, PMM=Posterior multimodal. **A, D.** FC brain maps with movie-rest Fz values shown for all predictive parcels for the youngest (6-7-year-olds) and oldest (16-20-year-olds) subjects. **B, E.** FC matrices with movie-rest Fz values shown for all predictive network clusters. **C, F.** *Left*: FC brain maps seeded from the LAN network with Fz values for all predictive edges. *Right*: FC maps seeded from the PMM network with Fz values for all predictive edges. Middle: Unthresholded FC matrices with movie-rest Fz values for all edges. Gray dotted lines represent predictive parcels corresponding to the language and PMM networks. In the brain maps, the final set of predictive parcels is outlined in purple.

4. Discussion

Understanding the brain's functional network organization during distinct cognitive states and associated developmental changes remains a key knowledge gap in characterizing the neural bases of higher cognitive functioning in the developing brain. Here we use a novel data-driven approach, **Pr**ed **i**ction of **M**ulti-L **e**vel **N**eural **E**ffec **t**s (PrimeNet), to distinguish resting-state from movie-watching FC patterns in a cross-sectional developmental sample from 6 to 20 years of age. We demonstrate that whole-brain FC can reliably distinguish between movie-watching and rest irrespective of age. The differences between the two states are associated with a robust and predictive set of FC features that generalizes across two different movies. Despite cross-movie generalizability, we also observe that predictive accuracy of state (movie vs rest) is affected by movie-specific content. Furthermore, we demonstrate that distinct functional network clusters exhibit differences in FC and widespread coupling across movie-watching and resting-state BOLD acquisitions. Finally, we show that there are FC patterns that differ across movie-watching and rest as a function of age.

In addition to its relevance for predicting state-specific FC patterns, the relationship between movie-watching and rest states is relevant for applications in developmental neuroimaging. Although resting-state studies are commonplace in adult research, movie-watching fMRI paradigms are better suited for pediatric research since movies engage the subjects' attention and robustly reduce head motion in children under the age of 10 (Greene et al., 2018; Vanderwal, Eilbott, and Castellanos, 2019). Thus, comparisons across age are often confounded by the use of resting-state in adults and movie-watching in young children. Furthermore, the study of the functional network organization underlying movie-watching versus rest has been challenging due to the lack of large-scale developmental datasets that allow comparisons with the adult brain. Finally, there is limited work on the associated neurodevelopmental trajectories of these two states. Resolving this question is particularly important to improve fMRI paradigms in pediatric research, thereby allowing across-age comparisons and establishing a direct relationship of state differences with those seen in the adult brain. Our results address these knowledge gaps and inform three parallel questions, which we summarize below.

4.1. Whole-brain FC features predict naturalistic viewing versus rest

We demonstrate that movie-watching and rest are characterized by reliable, state-specific FC patterns across age. This finding is consistent with prior studies showing that movie-watching and resting-state paradigms are associated with distinct FC patterns in adults (Demirtaş et al., 2019; Vanderwal et al., 2015) (see review by (Vanderwal, Eilbott, and Castellanos, 2019)). We go beyond prior findings, however, by identifying a set of FC patterns that consistently and reliably distinguish movie-watching versus rest across subjects irrespective of age. Specifically, we use a predictive multi-level FC-based model (PrimeNet) to predict brain state (movie versus rest). In turn, we apply the model to a withheld subset of the data with very high average accuracy (FD<0.5 mm, mean accuracy: 89%) across 1,000 iterations. Similar results are observed with a more stringent head motion criterion (FD<0.2 mm, mean accuracy: 90%). These results provide robust evidence of state-specific FC patterns that differentiate between movie-watching and rest in a developmental population. Furthermore, our results indicate that FC patterns generalize across movies. Specifically, we observe similar predictive accuracy when the model trained with an animated short film ('The Present', 3.47 min) is applied to a longer movie clip ('Despicable Me', 10 min) using a subset of the initial data (mean predictive accuracy: 95% on the same movie versus 90% on a different movie). Furthermore, state-specific predictive FC features for models run on each of the movies independently show 82% overlap. These findings indicate that there seems to be a 'core' stable functional network organization across development, which is reliably dissociated by cognitive state (i.e., movie versus rest). We focus here on state-specific differences, but it is unknown whether both movie-watching and rest states are supported by a 'core' intrinsic FC network organization as found in the adult brain, in line with prior findings on state-trait variation (Gratton et al., 2018).

Importantly, our results shed light on the impact of scan length (i.e., number of BOLD frames) on prediction of brain state. First, our results indicate that more data are not necessarily better for identification of state-specific FC features. Indeed, data from the 3.47-min movie (movie TP, "The Present") achieves a higher mean predictive accuracy in comparison to the longer 10 min movie (movie DM, "Despicable Me") irrespective of the number of frames from the longer movie included in the model. Relatedly, state-specific predictive accuracy fluctuates within a given BOLD run and it is independent of the number of frames included in the movie. Specifically, we observe a drop in predictive accuracy in the movie DM at 4.5 min into the movie (mean accuracy: 64%). We argue that these fluctuations in predictive accuracy may be related to the type of stimuli shown in the movie, in line with recent studies on the adult brain (van der Meer et al., 2020). For example, decontextualized scenes that require rapid inferences and a constant engagement may lead to FC states during movie-watching that are more difficult to differentiate from rest. We believe this was the case for the 10 min movie clip that we used, which is extracted from a longer movie. Indeed, the drop in predictive accuracy coincides with a scene in which the characters have a conversation that assumes the audience has information from earlier scenes in the movie. Thus, we hypothesize that presentation of a complete narrative that is properly contextualized engages participants' attention to a greater extent, therefore leading to more consistent FC patterns over time across subjects.

Furthermore, the predictive model implemented in the current study focuses on a 5 min resting-state session and shows robust state-specific predictive accuracy even when using more stringent motion thresholds. In the context of prior studies on test-retest individual-level reliability, gains in reliability of resting-state networks and therefore stability of the resting-state neural signal across sessions are greater in scans of 9–12 min in adults (Birn et al., 2013; Noble et al., 2017). In developmental samples, there is moderate-to-high test-retest reliability of resting-state FC in children and adolescents (9-17-year-olds) in 6 min scans (Somandepalli et al., 2015; Thomason et al., 2011), although reliability seems to be lower than in adults (Kaufmann et al., 2017; Vanderwal, Eilbott, Kelly, et al., 2019). In our study, the addition of a second 5 min resting-state BOLD run to the model yielded significantly lower predictive accuracy for the movie TP, but not for the movie DM. Additional work is needed to characterize how the reliability of FC patterns changes as a function of amount of data and scan condition in children, and how FC reliability impacts predictive power. Finally, although we did not explore this possibility, it may be that some edges across subjects are less stable/reliable during longer versus shorter scan times since it is possible that the test-retest reliability of edges is higher in some networks/regions than in others across subjects.

4.2. Naturalistic viewing and rest involve distinct functional networks

We show that movie-watching versus rest FC differences are associated with distinct patterns within specific large-scale functional networks. The predictive FC patterns identified by the binary classifier are associated with brain areas that belong to ten functional networks. Our findings confirm prior results indicating that movie-watching modulates FC, particularly in the default mode and dorsal-attention networks (Betti et al., 2013; Gao & Lin, 2012; Lv et al., 2013; Vanderwal et al., 2015). In addition, we found that movie-watching modulates FC in auditory, language and posterior-multimodal (PMM) networks. Indeed, we observe a strong coupling of language and PMM networks during movie-watching relative to rest. It is likely that state-specific stimuli during movie-watching modulate additional network-level FC that is absent during rest and that were not captured in our study. Collectively, these results provide evidence of a common functional brain network organization that differentiates movie-watching from rest.

Importantly, here we isolated parcel-level and network-level signals using the whole-brain CAB-NP parcellation (Ji et al., 2019), derived from the HCP atlas (Glasser et al., 2016). Although further studies are needed to determine whether the results generalize to other parcellations, we would also expect that brain coverage, granularity (i.e., number of parcels and networks), topology (volume/surface) and age group representative of the parcellation (e.g., age-specific versus adult-based parcellations) would have an effect on the final set of predictive edges (Gonzalez-Castillo et al., 2015; Vanderwal et al., 2017) (see (Eickhoff et al., 2018) for a review of differences/similarities across imaging-based parcellations). Furthermore, network-level signals are expected to be more generalizable than those at the parcel level given their stability across scrubbing criteria (i.e., network signals get averaged across multiple parcels). Indeed, implementation of a more stringent scrubbing criterion with an FD<0.2 mm replicates the observed FC effects in frontoparietal, language, posterior-multimodal, V1 and V2 networks.

4.3. FC during naturalistic viewing versus rest varies across age

Most relevant for studies of neurodevelopment, we show that FC patterns in the most predictive functional network clusters for movie-watching versus rest also exhibit notable variation as a function of age. Put differently, finding the most predictive movie-rest features across age (i.e., main effect of state) does not guarantee that there will not be an interaction with age (i.e., state x age interaction). Our results are consistent with there being such an interaction that involves auditory, language, posterior-multimodal and V2 network clusters. Across network clusters and relative to older subjects, the youngest children (6 and 7-year-olds) tend to show the lowest average within-network FC in clusters of the default, visual, language and PMM networks, as well as in the coupling of these clusters with other brain areas. Some exceptions are areas of the somatomotor network and FC between areas of the frontoparietal and dorsolateral attention networks, which show stronger FC in children relative to adolescents/young adults. Additionally, FC in language and PMM networks exhibits an age-related steady increase in children between 6-9 years of age and reaches a peak in early adolescence. Finally, our results align with prior work indicating that movie-watching paradigms are better suited in children under the age of 10, which may be due to reduction of head motion during movie-watching relative to rest (Alexander et al., 2017; Greene et al., 2018; Vanderwal, Eilbott, and Castellanos, 2019). Specifically, we found that accuracy of state prediction increases in children 6–9 years of age for movie-watching relative to resting-state scans, while no differences in predictive accuracy between the two states are observed in adolescents and young adults.

Overall, these results strongly suggest that neurodevelopmental comparisons across age should be performed using the same paradigms in order to fully understand changes in functional network organization across the lifespan. These data also imply that movie-watching paradigms in younger children may not be a good proxy for FC comparisons using resting-state paradigms in older children and adults. Nonetheless, a potential application of the observed FC patterns is to inform computational models that allow transformation of movie-based FC patterns into "pseudo-rest" patterns, which may be beneficial for comparisons with resting-state data from adults and older children. Although further analyses are needed to examine the feasibility of this approach, we hypothesize that such a transformation would be particularly useful for comparing resting-state FC estimates even for the youngest children or those diagnosed with ADHD, who may have difficulty remaining still during resting-state scans. This would be particularly relevant when movie-to-movie comparisons between children and adults are not possible and also in the context of large-scale developmental datasets that include resting-state, but not movie-watching paradigms, such as the ABCD (Casey et al., 2018) or HCP Lifespan Development (Somerville et al., 2018) studies.

4.4. Limitations and future directions

While these results provide strong evidence of state-specific FC patterns that differentiate between movie-watching and rest in a developmental population, it is unclear which specific behavioral variation, such as attention or abstract thinking, may relate to this reliable state-specific effect. Addressing this question will require the use of task designs that manipulate different behavioral measures during movie-watching paradigms. For example, recent work has examined how stable patterns of brain activity can be used to identify 'event' structures that are associated with changes in causal structure or goal-setting during movies (Baldassano et al., 2017, 2018). Furthermore, studies examining the nature of dynamic FC during development have shown that networks reconfigure over time during a single task (Hutchison and Morton, 2015). Importantly, while network reconfigurations during a task show similarities across development, transitions between these reconfigurations are faster in adults. Similarly, during rest, networks reconfigure into independent 'states' (van der Meer et al., 2020), which vary as a function of age (Marusak et al., 2017). These whole-brain dynamic FC patterns have been associated with distinct cognitive phenotypes or cognitive-task-like processes (Diaz et al., 2013; Gonzalez-Castillo et al., 2019; Gonzalez-Castillo and Bandettini, 2018). Identification of moment-to-moment fluctuations in FC during a single task or resting-state scan is particularly relevant for neurodevelopmental research since it can provide age-related measures of variability in behaviorally relevant FC patterns over time. These measures can then be used in conjunction with FC static measures to examine the development of brain networks that support higher-level cognitive functions and how they are modulated by environmental stimuli and cognitive demands.

A further limitation is that the data-driven analytic approach implemented in the present study, while robust and successful in capturing differences between states, takes into account only FC at the level of single connections (i.e., edges). Therefore, it does not consider interconnections across edges, which may potentially reveal differences within clusters of adjacent edges. Future work would therefore benefit from graph-based models, or component-level analyses more generally, that focus on examining the organization, dynamics and topology of brain networks. To illustrate this point, one such analysis employs the network-based statistic (NBS) (Zalesky et al., 2010), which relies on cluster statistics to identify structure exhibited by single connections. The NBS approach has been successfully implemented to characterize network components across a range of psychopathologies (Korgaonkar et al., 2014; Lai et al., 2017; Z. Long et al., 2015; Lopes et al., 2017; Pua et al., 2018; Zajac et al., 2017). A key feature of the NBS approach is that it provides greater statistical power to control the family-wise error rate relative to other procedures, such as implementation of the false discovery rate (FDR). This type of approaches can be complementary to PrimeNet-like analyses as it can provide an additional measure of connectivity (i.e., brain network components) that may reveal crucial subnetworks not captured by analyses based on single connections.

Another outstanding question relates to the developmental sample used in the present study which, despite spanning a broad age range (6-20 years of age), does not include younger developmental populations (e.g., toddlers or pre-schoolers). Studies with younger populations are crucial since they may elucidate how development of specific cognitive abilities, such as language or attention, affects state-specific FC patterns. Such studies may also benefit from longitudinal designs that allow for within-subject comparisons over time. Longitudinal designs may be able to identify developmental trajectories that are key to understanding changes in state-specific FC patterns across age. In addition, one possible direction for future studies would be to examine differences between females and males in their patterns of neural responses, especially since prior studies have shown that neural responses to naturalistic stimuli vary as a function of sex (Petroni et al., 2018).

Due to the nature of the HBN dataset, which focuses on the diagnosis of mental health and learning disorders, a subset of participants was diagnosed with a learning or anxiety disorder or ADHD. Although we ensured a balanced inclusion of non-diagnosed/diagnosed subjects across age, our final sample included a mix of non-diagnosed and diagnosed participants, which may have had some impact on the observed effects. Specifically, mental health problems have been associated with delays in stabilization and individualization of the brain's functional connections or 'connectome' (Kaufmann et al., 2017; Vanderwal, Eilbott, Kelly, et al., 2019), which suggests that psychiatric disorders affect the maturation of FC patterns. To minimize the impact of psychopathology on the model results, we implemented the following measures: i) ensured a balanced inclusion of non-diagnosed/diagnosed subjects across age, ii) implemented a randomized design across the model iterations in all three steps (inference, training and prediction), iii) employed a within-subject design by including only subjects that had data for both resting-state scans and movie-watching scans, iv) included diagnostic status as a covariate in the model and v) resctricted the types of diagnosis to learning or anxiety disorders and ADHD. These steps limited the amount of data that we were able to include but were necessary to reduce the impact of psychopathology on the findings. Even with these precautions, the final dataset may have included subjects with a more variable connectome than expected in a healthy population. Although further studies are needed to examine the extent to which the observed FC patterns generalize to a typically-developing sample, we can conclude that the predictive model identifies FC patterns that are common to participants with and without clinical disorders.

Finally, some methodological considerations are worth noting. First, we aimed to limit the effects of motion as much as possible by imposing temporal SNR and movement scrubbing criteria that ensured a consistently high threshold for participant inclusion. However, it is possible that there may have been subtle motion differences that we were not able to capture. Second, it is unclear if there are subtle but important neuro-vascular and/or global signal respiratory changes/differences across age, which we cannot capture here. Further studies that examine age-specific changes in these physiological measures may be able to shed light on neurodevelopmental differences and improve available fMRI data processing pipelines.

5. Conclusion

Prior studies on the adult brain have identified an 'intrinsic' functional network organization that extends across brain states. There is also evidence of transient state-specific FC patterns that differentiate among specific cognitive states (e.g., rest and tasks). Here, we focus on how state-specific FC patterns differentiate between movie-watching and rest in a developmental sample. We demonstrate that whole-brain FC can reliably distinguish between these two states irrespective of age. Notably, the results generalize to a withheld subset of the data with very high average accuracy across 1000 iterations and also generalize across movies. The identified FC patterns are associated with functional networks that exhibit differences in FC and widespread coupling across movie-watching and rest. Importantly, we identify FC variation in the most predictive functional network clusters for movie-watching versus resting-state as a function of age. The results have implications for our understanding of state-specific network organization across development and for comparisons of state-specific FC patterns across age.

Declaration of Competing Interest

None.

Acknowledgments

Research reported in this publication was supported by National Institutes of Health Grant HD-037082 (R.N.A).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2020.117630.

References

Alexander, L.M., Escalera, J., Ai, L., Andreotti, C., Febre, K., Mangone, A., Vega-Potler, N., Langer, N., Alexander, A., Kovacs, M., Litke, S., O'Hagan, B., Andersen, J., Bronstein, B., Bui, A., Bushey, M., Butler, H., Castagna, V., Camacho, N., … Milham, M.P., 2017. Data descriptor: an open resource for transdiagnostic research in pediatric mental health and learning disorders. Sci. Data 4, 1–26. doi:10.1038/sdata.2017.181.

Baldassano, C., Chen, J., Zadbood, A., Pillow, J.W., Hasson, U., Norman, K.A., 2017. Discovering event structure in continuous narrative perception and memory. Neuron 95 (3), 709–721. doi:10.1016/j.neuron.2017.06.041.

Baldassano, C., Hasson, U., Norman, K.A., 2018. Representation of real-world event schemas during narrative perception. J. Neurosci. 38 (45), 9689–9699. doi:10.1523/JNEUROSCI.0251-18.2018.

Betti, V., DellaPenna, S., de Pasquale, F., Mantini, D., Marzetti, L., Romani, G.L., Corbetta, M., 2013. Natural scenes viewing alters the dynamics of functional connectivity in the human brain. Neuron 79 (4), 782–797. doi:10.1016/j.neuron.2013.06.022.

Birn, R.M., Molloy, E.K., Patriat, R., Parker, T., Meier, T.B., Kirk, G.R., Nair, V.A., Meyerand, M.E., Prabhakaran, V., 2013. The effect of scan length on the reliability of resting-state fMRI connectivity estimates. Neuroimage 83, 550–558. doi:10.1016/j.neuroimage.2013.05.099.

Bottenhorn, K.L., Flannery, J.S., Boeving, E.R., Riedel, M.C., Eickhoff, S.B., Sutherland, M.T., Laird, A.R., 2019. Cooperating yet distinct brain networks engaged during naturalistic paradigms: a meta-analysis of functional MRI results. Netw. Neurosci. 3 (1), 27–48. doi:10.1162/netn_a_00050.

Casey, B.J., Cannonier, T., Conley, M.I., Cohen, A.O., Barch, D.M., Heitzeg, M.M., Soules, M.E., Teslovich, T., Dellarco, D.V., Garavan, H., Orr, C.A., Wager, T.D., Banich, M.T., Speer, N.K., Sutherland, M.T., Riedel, M.C., Dick, A.S., Bjork, J.M., Thomas, K.M., … Dale, A.M., 2018. The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. Develop. Cogn. Neurosci. 32, 43–54. doi:10.1016/j.dcn.2018.03.001.

Coffin, P., Renaud, C., 2010. Despicable Me. Universal Pictures.

Cole, M.W., Bassett, D.S., Power, J.D., Braver, T.S., Petersen, S.E., 2014. Intrinsic and task-evoked network architectures of the human brain. Neuron 83 (1), 238–251. doi:10.1016/j.neuron.2014.05.014.

Damaraju, E., Caprihan, A., Lowe, J.R., Allen, E.A., Calhoun, V.D., Phillips, J.P., 2014. Functional connectivity in the developing brain: a longitudinal study from 4 to 9months of age. Neuroimage 84, 169–180. doi:10.1016/j.neuroimage.2013.08.038.

De Asis-Cruz, J., Bouyssi-Kobar, M., Evangelou, I., Vezina, G., Limperopoulos, C., 2015. Functional properties of resting state networks in healthy full-term newborns. Sci. Rep. 5, 1–15. doi:10.1038/srep17755.

Demirtaş, M., Ponce-Alvarez, A., Gilson, M., Hagmann, P., Mantini, D., Betti, V., Romani, G.L., Friston, K., Corbetta, M., Deco, G., 2019. Distinct modes of functional connectivity induced by movie-watching. Neuroimage 184 (September 2018), 335–348. doi:10.1016/j.neuroimage.2018.09.042.

Diaz, B.A., van der Sluis, S., Moens, S., Benjamins, J.S., Migliorati, F., Stoffers, D., den Braber, A., Poil, S.S., Hardstone, R., Van't Ent, D.V., Boomsma, D.I., de Geus, E., Mansvelder, H.D., Van Someren, E.J.W., Linkenkaer-Hansen, K, 2013. The amsterdam resting-state questionnaire reveals multiple phenotypes of resting-state cognition. Front. Human Neurosci. 7 (JUL), 1–15. doi:10.3389/fnhum.2013.00446.

Dinstein, I., Pierce, K., Eyler, L., Solso, S., Malach, R., Behrmann, M., Courchesne, E., 2011. Disrupted neural synchronization in toddlers with Autism. Neuron 70 (6), 1218–1225. doi:10.1016/j.neuron.2011.04.018.

Dixon, M.L., Andrews-Hanna, J.R., Spreng, R.N., Irving, Z.C., Mills, C., Girn, M., Christoff, K., 2017. Interactions between the default network and dorsal attention network vary across default subsystems, time, and cognitive states. Neuroimage 147 (December 2016), 632–649. doi:10.1016/j.neuroimage.2016.12.073.

Eggebrecht, A.T., Elison, J.T., Feczko, E., Todorov, A., Wolff, J.J., Kandala, S., Adams, C.M., Snyder, A.Z., Lewis, J.D., Estes, A.M., Zwaigenbaum, L., Botteron, K.N., McKinstry, R.C., Constantino, J.N., Evans, A., Hazlett, H.C., Dager, S., Paterson, S.J., Schultz, R.T., … Pruett, J.R., 2017. Joint attention and brain functional connectivity in infants and toddlers. Cerebral Cortex (New York, N.Y. : 1991) 27 (3), 1709–1720. doi:10.1093/cercor/bhw403.

Eickhoff, S.B., Yeo, B.T.T., Genon, S, 2018. Imaging-based parcellations of the human brain. Nat. Rev. Neurosci. 19 (11), 672–686. doi:10.1038/s41583-018-0071-7.

Ellis, C.T., Skalaban, L.J., Yates, T.S., Bejjanki, V.R., Córdova, N.I., Turk-Browne, N.B., 2020. Re-imagining fMRI for awake behaving infants. Nat. Commun. 11 (1), 1–12. doi:10.1038/s41467-020-18286-y.

Ellis, Cameron T., Turk-Browne, N.B., 2018. Infant fMRI: a model system for cognitive neuroscience. Trends Cogn. Sci. 22 (5), 375–387. doi:10.1016/j.tics.2018.01.005.

Emerson, R.W., Short, S.J., Lin, W., Gilmore, J.H., Gao, W., 2015. Network-level connectivity dynamics of movie watching in 6-year-old children. Front. Human Neurosci. 9 (NOV), 1–8. doi:10.3389/fnhum.2015.00631.

Fair, D.A., Cohen, A.L., Power, J.D., Dosenbach, N.U.F., Church, J.A., Miezin, F.M., Schlaggar, B.L., Petersen, S.E, 2009. Functional brain networks develop from a "local to distributed" organization. PLoS Comput. Biol. 5 (5), 14–23. doi:10.1371/journal.pcbi.1000381.

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T., 2015. Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity. Nat. Neurosci. 18 (11), 1664–1671. doi:10.1038/nn.4135.

Fischl, B., 2012. FreeSurfer. Neuroimage 62 (2), 774–781. doi:10.1016/j.neuroimage.2012.01.021.

Fox, M.D., Snyder, A.Z., Vincent, J.L., Raichle, M.E., 2007. Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior. Neuron 56 (1), 171–184. doi:10.1016/j.neuron.2007.08.023.

Frey, J., 2014. The Present. Independent short film.

Gao, W., Alcauter, S., Smith, K.J., Gilmore, J., Lin, W., 2015. Development of human brain cortical network architecture during infancy. Brain Struct. Function 344 (6188), 1173–1178. doi:10.1126/science.1249098.Sleep.

Gao, W., Lin, W., 2012. Frontal parietal control network regulates the anti-correlated default and dorsal attention networks. Hum. Brain Mapp. 33 (1), 192–202. doi:10.1002/hbm.21204.

Gilson, M., Deco, G., Friston, K.J., Hagmann, P., Mantini, D., Betti, V., Romani, G.L., Corbetta, M., 2018. Effective connectivity inferred from fMRI transition dynamics during movie viewing points to a balanced reconfiguration of cortical interactions. Neuroimage 180 (January 2017), 534–546. doi:10.1016/j.neuroimage.2017.09.061.

Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., Smith, S.M., Essen, D.C.Van, 2016. A multi-modal parcellation of human cerebral cortex. Nature 536 (7615), 171–178. doi:10.1038/nature18933.A.

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., Van Essen, D.C., Jenkinson, M., for the WU-Minn HCP Consortium (2013). The minimal preprocessing pipelines for the Human Connectome Project. NeuroImage 80 (October), 105–124, doi:10.1016/j.neuroimage.2013.04.127.

Gonzalez-Castillo, J., Bandettini, P.A., 2018. Task-based dynamic functional connectivity: recent findings and open questions. Neuroimage 180 (May 2017), 526–533. doi:10.1016/j.neuroimage.2017.08.006.

Gonzalez-Castillo, J., Caballero-Gaudes, C., Topolski, N., Handwerker, D.A., Pereira, F., Bandettini, P.A., 2019. Imaging the spontaneous flow of thought: distinct periods of

cognition contribute to dynamic functional connectivity during rest. Neuroimage 202 (February), 116129. doi:10.1016/j.neuroimage.2019.116129.
Gonzalez-Castillo, J., Hoy, C.W., Handwerker, D.A., Robinson, M.E., Buchanan, L.C., Saad, Z.S., Bandettini, P.A., 2015. Tracking ongoing cognition in individuals using brief, whole-brain functional connectivity patterns. PNAS 112 (28), 8762–8767. doi:10.1073/pnas.1501242112.
Gratton, C., Laumann, T.O., Gordon, E.M., Adeyemo, B., Petersen, S.E., 2016. Evidence for two independent factors that modify brain networks to meet task goals. Cell Rep. 17 (5), 1276–1288. doi:10.1016/j.celrep.2016.10.002.
Gratton, C., Laumann, T.O., Nielsen, A.N., Greene, D.J., Gordon, E.M., Gilmore, A.W., Nelson, S.M., Coalson, R.S., Snyder, A.Z., Schlaggar, B.L., Dosenbach, N.U.F., 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. doi:10.1016/j.neuron.2018.03.035, .e5.
Grayson, D.S., Fair, D.A., 2017. Development of large-scale functional networks from birth to adulthood: a guide to the neuroimaging literature. Neuroimage 160, 15–31. doi:10.1109/MTAS.2004.1371634.
Greene, D.J., Koller, J.M., Hampton, J.M., Wesevich, V., Van, A.N., Nguyen, A.L., Hoyt, C.R., McIntyre, L., Earl, E.A., Klein, R.L., Shimony, J.S., Petersen, S.E., Schlaggar, B.L., Fair, D.A., Dosenbach, N.U.F., 2018. Behavioral interventions for reducing head motion during MRI scans in children. Neuroimage 171 (September 2017), 234–245. doi:10.1016/j.neuroimage.2018.01.023.
Hasson, U., Malach, R., Heeger, D.J., 2010. Reliability of cortical activity during natural stimulation. Trends Cogn. Sci. 14 (1), 40–48. doi:10.1016/j.tics.2009.10.011.
Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject synchronization of cortical activity during natural vision. Science 303 (MARCH), 1634–1640. doi:10.1126/science.1089506 .
Hutchison, R.M., Morton, J.B., 2015. Tracking the brain's functional coupling dynamics over development. J. Neurosci. 35 (17), 6849–6859. doi:10.1523/JNEUROSCI.4638-14.2015.
James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning. In: Casella, G., Fienberg, S., Olkin, I. (Eds.), Springer Texts in Statistics. Springer doi:10.1016/j.peva.2007.06.006.
Ji, J.L., Spronk, M., Kulkarni, K., Repovš, G., Anticevic, A., Cole, M.W., 2019. Mapping the human brain's cortical-subcortical functional network organization. Neuroimage 185 (October 2018), 35–57. doi:10.1016/j.neuroimage.2018.10.006.
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. Nat. Neurosci. 20 (4), 513–515. doi:10.1038/nn.4511.
Khundrakpam, B.S., Reid, A., Brauer, J., Carbonell, F., Lewis, J., Ameis, S., Karama, S., Lee, J., Chen, Z., Das, S., Evans, A.C., 2013. Developmental changes in organization of structural brain networks. Cereb. Cortex 23 (9), 2072–2085. doi:10.1093/cercor/bhs187.
Konishi, Y., Taga, G., Yamada, H., Hirasawa, K., 2002. Functional brain imaging using fMRI and optical topography in infancy. Sleep Med. 3 (SUPPL. 2), 41–43. doi:10.1016/S1389-9457(02)00163-6.
Korgaonkar, M.S., Fornito, A., Williams, L.M., Grieve, S.M., 2014. Abnormal structural networks characterize major depressive disorder: A connectome analysis. Biol. Psychiatry 76 (7), 567–574. doi:10.1016/j.biopsych.2014.02.018.
Lai, C.H., Wu, Y.Te, Hou, Y.M, 2017. Functional network-based statistics in depression: Theory of mind subnetwork and importance of parietal region. J. Affect. Disord. 217 (539), 132–137. doi:10.1016/j.jad.2017.03.073.
Long, X., Benischek, A., Dewey, D., Lebel, C., 2017. Age-related functional brain changes in young children. Neuroimage 155 (April), 322–330. doi:10.1016/j.neuroimage.2017.04.059.
Long, Z., Duan, X., Wang, Y., Liu, F., Zeng, L., Zhao, J., Chen, H., 2015. Disrupted structural connectivity network in treatment-naive depression. Prog. Neuropsychopharmacol. Biol. Psychiatry 56, 18–26. doi:10.1016/j.pnpbp.2014.07.007.
Lopes, R., Delmaire, C., Defebvre, L., Moonen, A.J., Duits, A.A., Hofman, P., Leentjens, A.F.G., Dujardin, K, 2017. Cognitive phenotypes in parkinson's disease differ in terms of brain-network organization and connectivity. Hum. Brain Mapp. 38 (3), 1604–1621. doi:10.1002/hbm.23474.
Lv, Y., Margulies, D.S., Villringer, A., Zang, Y.F., 2013. Effects of finger tapping frequency on regional homogeneity of sensorimotor cortex. PLoS One 8 (5), 6–11. doi:10.1371/journal.pone.0064115.
Lynch, L.K., Lu, K.H., Wen, H., Zhang, Y., Saykin, A.J., Liu, Z., 2018. Task-evoked functional connectivity does not explain functional connectivity differences between rest and task conditions. Hum. Brain Mapp. 39 (12), 4939–4948. doi:10.1002/hbm.24335.
Madhyastha, T., Peverill, M., Koh, N., McCabe, C., Flournoy, J., Mills, K., King, K., Pfeifer, J., McLaughlin, K.A., 2018. Current methods and limitations for longitudinal fMRI analysis across development. Develop. Cogn. Neurosci. 33 (November 2017), 118–128. doi:10.1016/j.dcn.2017.11.006.
Marusak, H.A., Calhoun, V.D., Brown, S., Crespo, L.M., Sala-Hamrick, K., Gotlib, I.H., Thomason, M.E., 2017. Dynamic functional connectivity of neurocognitive networks in children. Hum. Brain Mapp. 38 (1), 97–108. doi:10.1002/hbm.23346.
Miranda-Dominguez, O., Mills, B.D., Carpenter, S.D., Grant, K.A., Kroenke, C.D., Nigg, J.T., Fair, D.A., 2014. Connectotyping: Model based fingerprinting of the functional connectome. PLoS One (11) 9. doi:10.1371/journal.pone.0111048.
Noble, S., Spann, M.N., Tokoglu, F., Shen, X., Constable, R.T., Scheinost, D., 2017. Influences on the Test-Retest Reliability of Functional Connecfile:///Users/ss2569/Desktop/Research/Manuscripts/InPrep/FLUX.DCN.conceptual.manuscript/readings/1.state-trait-variation/rest/stability/Gratton_etal2018_StableNetworks.pdftivity MRI and its Rel. Cereb. Cortex 27 (11), 5415–5429. doi:10.1093/cercor/bhx230.
Pajula, J., Kauppi, J.P., Tohka, J., 2012. Inter-subject correlation in fMRI: method validation against stimulus-model based analysis. PLoS One 7 (8). doi:10.1371/journal.pone.0041196.
Petroni, A., Cohen, S.S., Ai, L., Langer, N., Henin, S., Vanderwal, T., Milham, M.P., Parra, L.C., 2018. The variability of neural responses to naturalistic videos change with age and sex. ENeuro 5 (1), 1–13. doi:10.1523/ENEURO.0244-17.2017.
Power, J.D., Barnes, K.A., Snyder, A.X., Schlaggar, B., Petersen, S.E., 2012. Spurious but systematic correlation in functional connectivity MRI networks arise form subject motion. Neuroimage 23 (1), 1–7. doi:10.1038/jid.2014.371.
Power, J.D., Mitra, A., Laumann, T.O., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2014. Methods to detect, characterize, and remove motion artifact in resting-state fMRI. Neuroimage 84, 320–341.
Pua, E.P.K., Malpas, C.B., Bowden, S.C., Seal, M.L, 2018. Different brain networks underlying intelligence in autism spectrum disorders. Hum. Brain Mapp. 39 (8), 3253–3262. doi:10.1002/hbm.24074.
Rohr, C.S., Arora, A., Cho, I.Y.K., Katlariwala, P., Dimond, D., Dewey, D., Bray, S, 2018. Functional network integration and attention skills in young children. Develop. Cognitive Neurosci. 30 (November 2017), 200–211. doi:10.1016/j.dcn.2018.03.007.
Rohr, C.S., Vinette, S.A., Parsons, K.A.L., Cho, I.Y.K., Dimond, D., Benischek, A., Lebel, C., Dewey, D., Bray, S, 2017. Functional connectivity of the dorsal attention network predicts selective attention in 4-7 year-old girls. Cereb. Cortex 27 (9), 4350–4360. doi:10.1093/cercor/bhw236.
Rosenberg, M.D., Scheinost, D., Greene, A.S., Avery, E.W., Kwon, Y.H., Finn, E.S., Ramani, R., Qiu, M., Constable, T.R., Chun, M.M., 2020. Functional connectivity predicts changes in attention over minutes, days, and months. Proc. Natl Acad. Sci. 34, 1–32. doi:10.1073/pnas.1912226117.
Simony, E., Honey, C.J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., Hasson, U., 2016. Dynamic reconfiguration of the default mode network during narrative comprehension. Nat. Commun. 7 (May 2015). doi:10.1038/ncomms12141.
Smyser, C.D., Inder, T.E., Shimony, J.S., Hill, J.E., Degnan, A.J., Snyder, A.Z., Neil, J.J., 2010. Longitudinal analysis of neural network development in preterm infants. Cereb. Cortex 20 (12), 2852–2862. doi:10.1093/cercor/bhq035.
Snyder, A.Z., 2016. Intrinsic brain activity and resting state networks. In: Pfaff, D.W., Volkow, N.D. (Eds.), Neuroscience in the 21st Century: From Basic to Clinical. Springer Science, pp. 1–4155. doi:10.1007/978-1-4939-3474-4.
Somandepalli, K., Kelly, C., Reiss, P.T., Zuo, X.N., Craddock, R.C., Yan, C.G., Petkova, E., Castellanos, F.X., Milham, M.P., Di Martino, A., 2015. Short-term test-retest reliability of resting state fMRI metrics in children with and without attention-deficit/hyperactivity disorder. Develop. Cognitive Neurosci. 15, 83–93. doi:10.1016/j.dcn.2015.08.003.
Somerville, L., Bookheimer, S., Buckner, R., Burgess, G., Van Essen, D.C., Barch, D.M., 2018. The lifespan human connectome project in development: a large-scale study of brain connectivity development in 5-21 year olds. Neuroimage 456–468. doi:10.1016/j.neuroimage.2018.08.050.
Sonkusare, S., Breakspear, M., Guo, C., 2019. Naturalistic stimuli in neuroscience: critically acclaimed. Trends Cogn. Sci. 23 (8), 699–714. doi:10.1016/j.tics.2019.05.004.
Supekar, K., Musen, M., Menon, V., 2009. Development of large-scale functional brain networks in children. PLoS Biol. 7 (7). doi:10.1371/journal.pbio.1000157.
Thomason, M.E., Dennis, E.L., Joshi, A.A., Joshi, S.H., Dinov, I.D., Chang, C., Henry, M.L., Johnson, R.F., Thompson, P.M., Toga, A.W., Glover, G.H., Van Horn, J.D., Gotlib, I.H., 2011. Resting-state fMRI can reliably map neural networks in children. Neuroimage 55 (1), 165–175. doi:10.1016/j.neuroimage.2010.11.080.
Uematsu, A., Matsui, M., Tanaka, C., Takahashi, T., Noguchi, K., Suzuki, M., Nishijo, H., 2012. Developmental trajectories of amygdala and hippocampus from infancy to early adulthood in healthy individuals. PLoS One 7 (10). doi:10.1371/journal.pone.0046970.
van den Heuvel, M.P., Kersbergen, K.J., De Reus, M.A., Keunen, K., Kahn, R.S., Groenendaal, F., De Vries, L.S., Benders, M.J.N.L, 2015. The neonatal connectome during preterm brain development. Cereb. Cortex 25 (9), 3000–3013. doi:10.1093/cercor/bhu095.
van der Meer, J.N., Breakspear, M., Chang, L.J., Sonkusare, S., Cocchi, L., 2020. Movie viewing elicits rich and reliable brain state dynamics. Nat. Commun. 11 (1), 1–14. doi:10.1038/s41467-020-18717-w.
Vanderwal, T., Eilbott, J., Castellanos, F.X., 2019. Movies in the magnet: naturalistic paradigms in developmental functional neuroimaging. Develop. Cogn. Neuroscience 36 (October), 1–15. doi:10.1016/j.dcn.2018.10.004.
Vanderwal, T., Eilbott, J., Finn, E.S., Craddock, R.C., Turnbull, A., Castellanos, F.X., 2017. Individual differences in functional connectivity during naturalistic viewing conditions. Neuroimage 157 (June), 521–530. doi:10.1016/j.neuroimage.2017.06.027.
Vanderwal, T., Eilbott, J., Kelly, C., Woodward, T.S., Milham, M.P., Castellanos, F.X., 2019. Stability and similarity of the pediatric connectome as developmental outcomes. BioRxiv https://doi.org/http://dx.doi.org/10.1101/828137.
Vanderwal, T., Kelly, C., Eilbott, J., Mayes, L.C., Castellanos, F.X., 2015. Inscapes: A movie paradigm to improve compliance in functional magnetic resonance imaging. Neuroimage 122, 222–232. doi:10.1016/j.neuroimage.2015.07.069.
Zajac, L., Koo, B.B., Bauer, C.M., Killiany, R., 2017. Seed location impacts whole-brain structural network comparisons between healthy elderly and individuals with alzheimer's disease. Brain Sci. 7 (4). doi:10.3390/brainsci7040037.
Zalesky, A., Fornito, A., Bullmore, E.T., 2010. Network-based statistic: Identifying differences in brain networks. Neuroimage 53 (4), 1197–1207. doi:10.1016/j.neuroimage.2010.06.041.