nature neuroscience

Article https://doi.org/10.1038/s41593-023-01358-9

Large-scale encoding of emotion concepts becomes increasingly similar between individuals from childhood to adolescence

In the format provided by the authors and unedited

Supplemental Materials

Appendix A: Full characterization of video features using the EmoCodes system

A1: Associations among video features

The full features that we characterize for each video are listed in Table SA1. Video features were characterized using the EmoCodes system[1] following standard procedures. At least two independent raters rated each video. For non-emotional codes (body, face, number of characters, spoken words, written words), consensus was required before using in further analyses. For emotion codes, the ratings were averaged across the two raters before further computation or analysis. Overlap between ratings (before consensus or averaging) ranged from 0.65-1.00, indicating high consistency between raters. Traces for each code included in the general linear models estimating activation are shown in Figure SA1.

General codes were computed by multiplying the character's Boolean valence ratings by the character's overall intensity rating, then summing across characters. For specific emotion ratings, Boolean emotion ratings were summed across modality (face, body, voice), multiplied by the character's intensity rating, and summed across characters.

Table SA1: Details of each feature obtained from each video through either frame-by-frame video coding (manual) or by using video processing python functions (automatic). The right two columns indicate the percentage of the video that had that label "present" (non-zero). DM=*Despicable Me*, 10 minutes; TP=*The Present*, 3 minutes and 20 seconds.

Label	Type	Description	DM % non-zero	TP % non-zero
positive	manual	General positive affect expressed across all characters.	41.2%	75.2%
negative	manual	General negative affect expressed across all characters.	61.3%	26.0%
anger	manual	A strong feeling of annoyance, hostility, or displeasure. Anger can be provoked or unprovoked and is typically high intensity and high negative valence.	33.6%	22.4%
happy	manual	A feeling of pleasure or contentment. To distinguish happiness from excitement, happiness in this coding scheme is a moderate to low arousal positive emotion.	35.6%	71.7%
fear	manual	A negative emotion brought about by a threat to one's physical or psychological safety. Fear can come about due to actual or perceived threats and is typically characterized by high arousal as well as a flight, fight, or freezing response.	24.0%	2.3%
excited	manual	Great enthusiasm or eagerness, typically in anticipation of a desired event. Excitement is high arousal and positive.	9.9%	16.9%
sad	manual	A negative emotion typically brought about by unfavorable events or thoughts. Sadness can be high or low arousal.	24.7%	3.9%
body	manual	Presence of body parts/biological motion.	90.5%	89.0%
faces	manual	Presence of faces. Both full and partial faces are included.	88.8%	79.9%
number of characters	manual	Total number of characters on the screen.	89.7%	93.3%
spoken words	manual	Denotes when identifiable words are being uttered (excludes other speech such as grunts or fillers).	54.4%	21.3%
written words	manual	Denotes when legible words are visible on the screen.	16.8%	3.5%
brightness	automatic	Mean frame-by-frame luminance.	N/A	N/A
saliency	automatic	Fraction of the frame classified as highly salient using the Itti & Koch algorithm[2,3]	N/A	N/A
sharpness	automatic	Degree of blur or sharpness in each frame.	N/A	N/A
vibrance	automatic	Variance among color channels in each frame.	N/A	N/A
optical flow	automatic	Frame-by-frame motion.	N/A	N/A
tempo	automatic	Rolling average of music tempo using a sliding 30-second window.	N/A	N/A
loudness	automatic	Root mean squared of the audio amplitude.	N/A	N/A

Figure SA1: Traces for each feature coded using the EmoCodes system for each Despicable Me (10 minutes, panel **A**) and The Present (3 minutes and 20 seconds, panel **B**). The below traces were convolved with the hemodynamic response function before estimating activation.

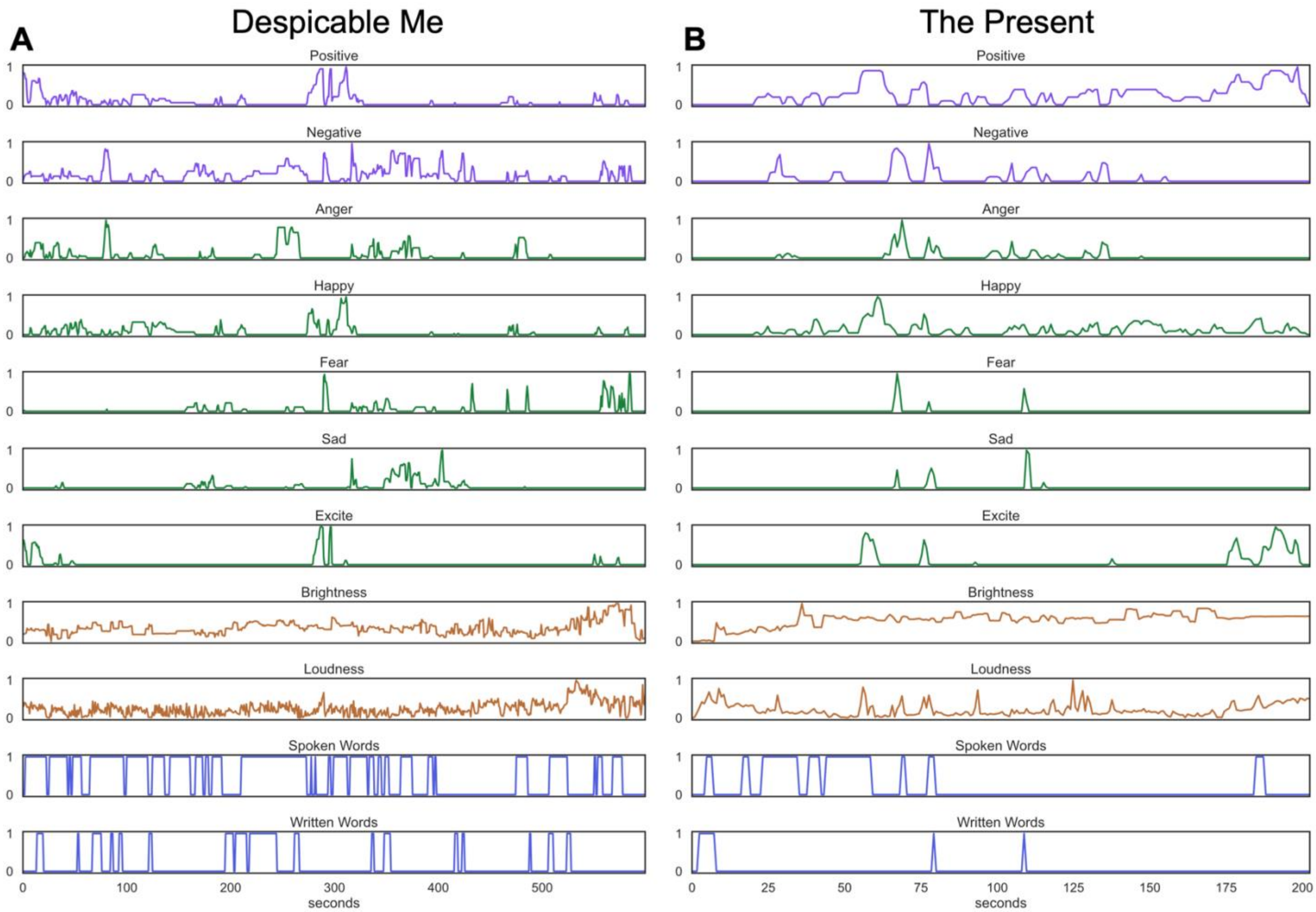
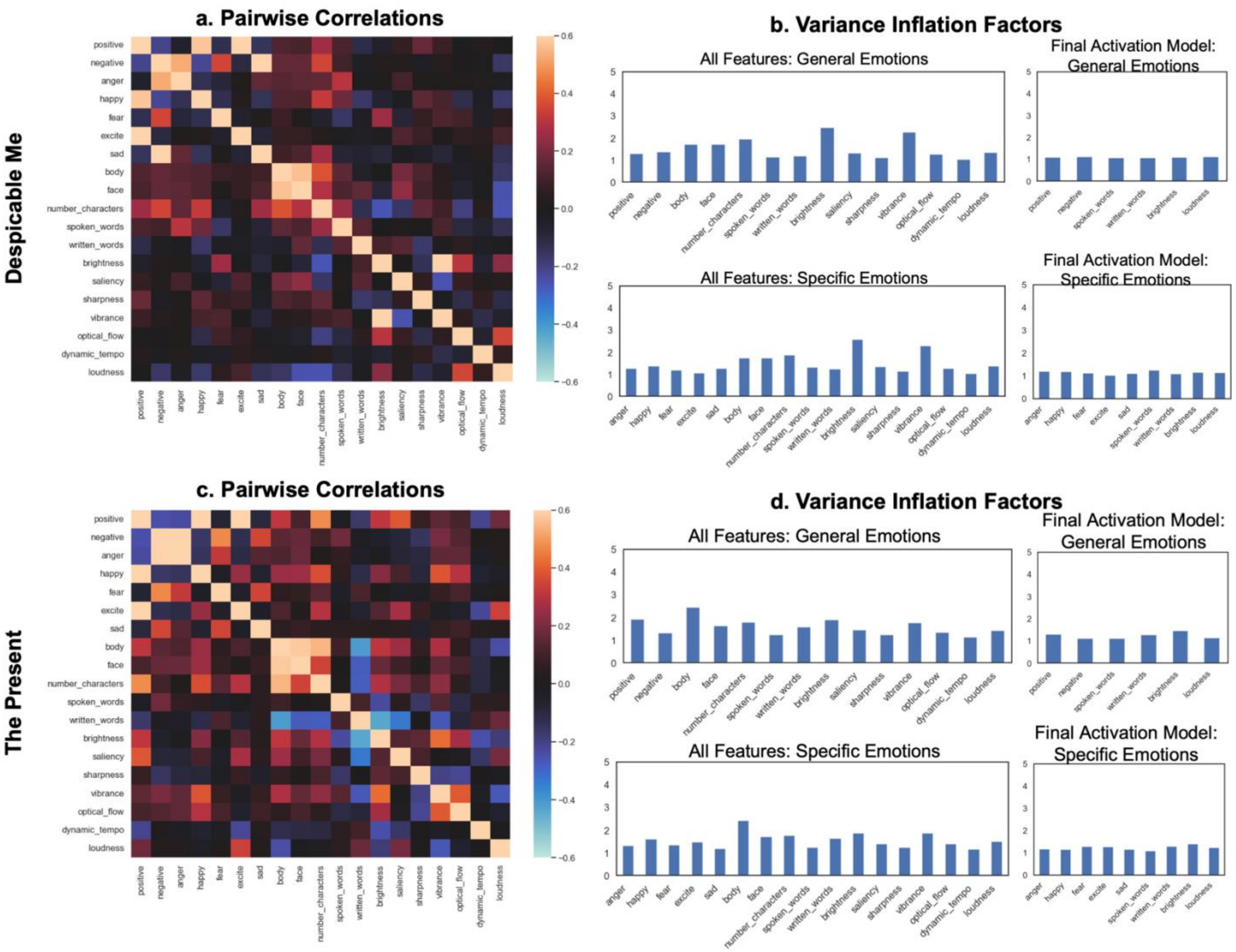


Figure SA2: Full feature analysis for the *Despicable Me* clip (top) and *The Present* (bottom). **Panel a**. Heatmap of the pair-wise Pearson correlations among all video features identified in the video. **Panel b**. Variance inflation factors (VIF) among all features (left) and the features included in the activation model (right). Because it is expected—and found—that the specific emotions would be correlated to the general emotions, we computed VIF scores separately for each (general is top, specific is bottom). A VIF of less than 2 is considered not notably collinear, between 2 and 5 is considered somewhat collinear though still independent enough for the purposes of linear regression, and a score above 5 is considered collinear and problematic to include in the same regression model.



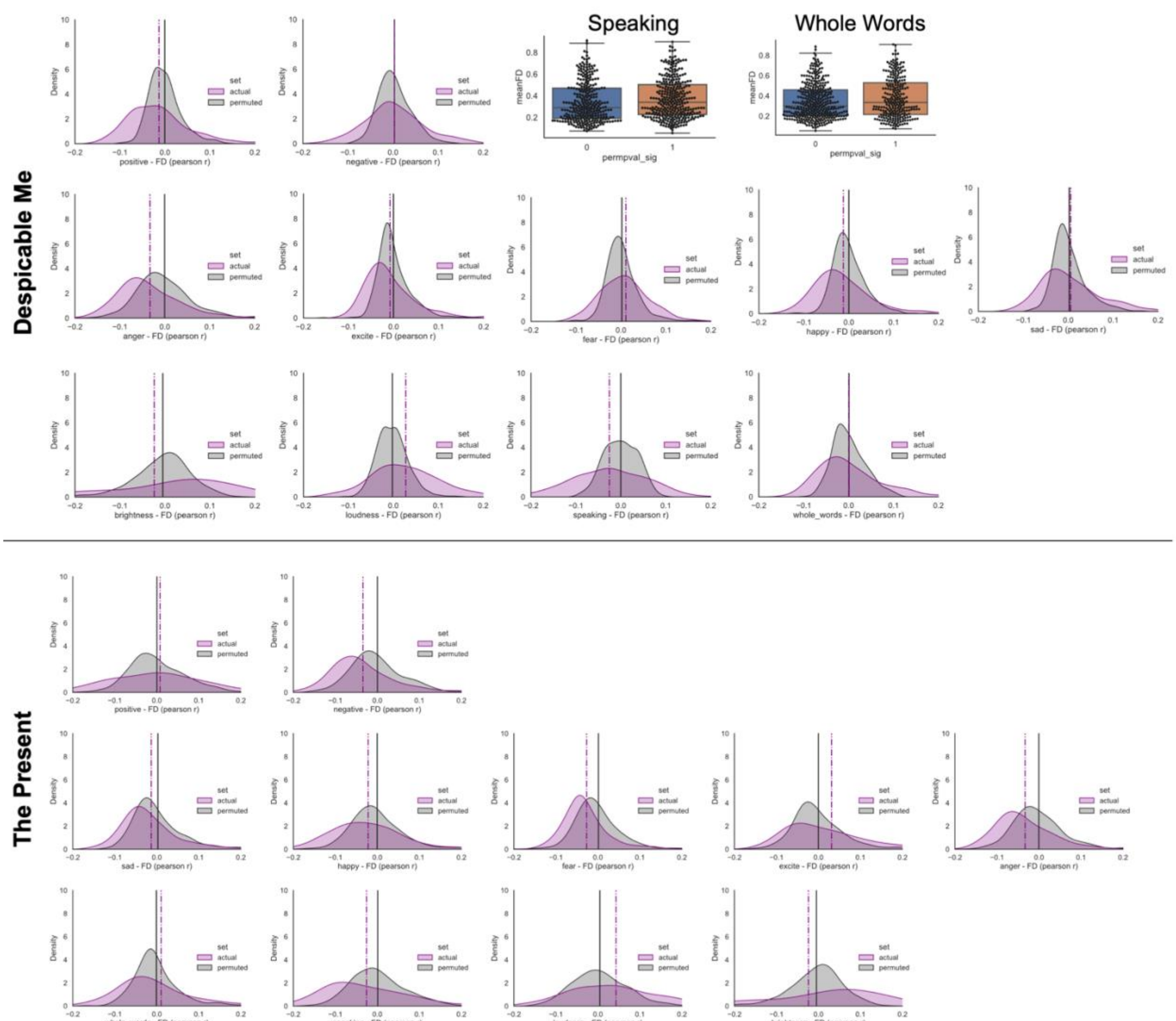
A2. Motion metrics

We next examined if there were any associations between participant motion and video features. For each participant, we conducted Pearson correlations between framewise displacement across the video and each video feature that we extracted activation to: positive, negative, anger, fear, excitement, happiness, sadness, brightness, loudness, speaking, and written words. We computed these correlations again using permuted data to generate a null distribution for statistical comparison. These plots are shown in figures SA4 and SA5. For each video feature and FD correlation, we computed a permutation-based p-value (2-sided).

If there were at least 20 children with a significant correlation between FD and a given video feature, we then use a t-test to examined differences in age or average motion between the children with a significant association between motion and the video feature. Only for the Speaking and Whole Words features for Despicable Me was there a significant difference in average motion between children with a significant association between the code and FD and those without (Speaking-FD meanFD two-sided t=-2.13, p=0.034; Whole Words-FD meanFD: two-sided t=-2.37, p=0.018). In both cases the children with a significant association between FD and that feature had higher average motion.

Together, these analyses suggest no association between participant motion and the emotion regressors examined in the main text analysis.

Figure SA3: Analysis of participant motion and *Despicable Me* (top) and *The Present* (bottom) video features. Distribution plots show the Pearson r values across the sample and for a permuted dataset. Notably, the actual and permuted distributions are highly overlapping for each analysis, indicating very few participants with a significant association between motion and video features. For each analysis, if there were at least 20 participants with a significant association between the video feature and motion, a t-test was conducted to test if those children were systematically older/younger or moved more on average. Only 2 of those tests were significant (group differences are plotted in the upper right).



Appendix B: Additional support vector machine analysis results

B1: Classifying activation maps within and across videos

We sought to maximize generalizability by using both movie clips for the main analyses. However, due to both the differences in the narratives of each video and recent work suggesting that video clips can be optimized to detect specific effects[2,3], we have repeated our analyses separately and across clips.

Table SB1: Classifying emotions from activation within and across movies. For all models, training and test data were determined based on collection site, thus there was no overlap in samples. We were able to successfully train a model to classify activation maps for general and specific emotions greater than chance when training and testing data were limited to maps from one video. The only exception was for the specific emotions classifier in which we trained using data from one video and tested in data from the other video. DM=Despicable Me, TP= The Present.

Training Data	Testing Data	Labels Predicting	Chance Accuracy	Train Accuracy [95% CI]	Test Accuracy [95% CI]	p-value
General Emotions						
both	both	positive, negative	50%	89% [81%, 94%]	88% [86%, 89%]	<0.001
DM only	DM only	positive, negative	50%	95% [85%, 100%]	87% [84%, 90%]	<0.001
TP only	TP only	positive, negative	50%	93% [83%, 98%]	93% [91%, 94%]	<0.001
DM only	TP only	positive, negative	50%	95% [85%, 100%]	59% [56%, 62%]	0.062
TP only	DM only	positive, negative	50%	93% [83%, 98%]	61% [57%, 65%]	0.027
Specific Emotions						
both	both	anger, fear, sad, happy, excite	20%	77% [68%, 81%]	74% [72%, 75%]	<0.001
DM only	DM only	anger, fear, sad, happy, excite	20%	86% [77%, 92%]	77% [75%, 79%]	<0.001
TP only	TP only	anger, fear, sad, happy, excite	20%	89% [81%, 95%]	91% [89%, 92%]	<0.001
DM only	TP only	anger, fear, sad, happy, excite	20%	86% [77%, 92%]	15% [14%, 18%]	0.928
TP only	DM only	anger, fear, sad, happy, excite	20%	89% [81%, 95%]	22% [20%, 25%]	0.201

B2: Using a curvilinear kernel to predict maturity metrics from activation maps

Table SB2: Only the models trained on each anger and loudness to predict age performed significantly better using a curvilinear kernel than the linear kernel model on at least one performance metric. All statistical tests were one-sided. MSE=mean square error.

Activation Data	Spearman r [95% CI]	Spearman p-value	Pearson r [95% CI]	Pearson p-value	MSE	MSE p-value
Chronological Age						
Main Analyses						
negative	0.22 [0.15, 0.28]	<0.001	0.22 [0.14, 0.27]	<0.001	8.18	0.875
positive	0.20 [0.13, 0.26]	<0.001	0.20 [0.12, 0.26]	<0.001	8.23	0.886
anger	0.31 [0.23, 0.36]	<0.001	0.30 [0.23, 0.36]	<0.001	7.82	0.786
excite	0.19 [0.12, 0.26]	<0.001	0.20 [0.13, 0.27]	<0.001	8.21	0.880
fear	0.21 [0.14, 0.28]	<0.001	0.21 [0.14, 0.28]	<0.001	8.23	0.889
happy	0.22 [0.15, 0.28]	<0.001	0.22 [0.15, 0.28]	<0.001	8.20	0.878
sad	0.26 [0.19, 0.32]	<0.001	0.26 [0.19, 0.31]	<0.001	8.11	0.862
Specificity Analyses						

brightness	0.25 [0.2, 0.29]	<0.001	0.24 [0.19, 0.28]	<0.001	8.35	0.969
loudness	0.20 [0.15, 0.24]	<0.001	0.20 [0.15, 0.24]	<0.001	8.48	0.980
speaking	0.22 [0.15, 0.28]	<0.001	0.22 [0.15, 0.28]	<0.001	8.16	0.867
words	0.14 [0.06, 0.21]	<0.001	0.14 [0.06, 0.21]	<0.001	8.48	0.926
Puberty						
Main Analyses						
negative	0.16 [0.08, 0.24]	<0.001	0.17 [0.08, 0.24]	<0.001	20.35	0.641
positive	0.13 [0.05, 0.21]	0.007	0.12 [0.04, 0.2]	0.015	19.84	0.593
anger	0.17 [0.08, 0.25]	<0.001	0.17 [0.08, 0.25]	<0.001	20.12	0.613
excite	0.17 [0.08, 0.25]	<0.001	0.17 [0.08, 0.24]	<0.001	19.87	0.594
fear	0.13 [0.04, 0.22]	0.008	0.15 [0.07, 0.23]	0.002	20.42	0.647
happy	0.20 [0.11, 0.28]	<0.001	0.18 [0.09, 0.26]	<0.001	20.37	0.647
sad	0.23 [0.14, 0.3]	<0.001	0.24 [0.14, 0.3]	<0.001	20.40	0.648
Specificity Analyses						
brightness	0.14 [0.07, 0.2]	<0.001	0.14 [0.08, 0.2]	<0.001	20.34	0.700
loudness	0.13 [0.06, 0.18]	<0.001	0.11 [0.04, 0.16]	0.001	20.63	0.741
speaking	0.13 [0.04, 0.2]	0.009	0.15 [0.07, 0.22]	0.003	19.51	0.555
words	0.02 [-0.07, 0.11]	0.653	0.03 [-0.06, 0.12]	0.575	21.03	0.703

B3: Predicting maturity from activation separately for each movie

Table SB3: Maturity support vector regression results when limiting the activation data to *Despicable Me*. Results indicate model performance in the unseen test data, using a liner kernel. Based on overlap (or lack of overlap) of 95% confidence intervals, Positive, Anger, Spoken Words, and Written Words activation map models performed better at predicting chronological age when using *Despicable Me* data alone as compared to the full models. Models for other activation maps did not perform notably better than when using activation maps from both videos (see Table 1 in the main text for comparison). Puberty was not better predicted by *Despicable Me* activation maps. All statistical tests were one-sided.

	Activation Data	Spearman *r* [95% CI]	Spearman p-value	Pearson *r* [95% CI]	Pearson p-value	MSE	MSE p-value
Despicable Me	**Chronological Age**						
	Main Analyses						
	negative	0.14 [0.03, 0.23]	0.020	0.16 [0.05, 0.24]	0.008	9.88	0.632
	positive	0.39 [0.26, 0.45]	<0.001	0.36 [0.24, 0.43]	<0.001	9.15	0.402
	anger	0.36 [0.25, 0.43]	<0.001	0.36 [0.26, 0.43]	<0.001	8.30	0.419
	excite	0.41 [0.29, 0.48]	<0.001	0.39 [0.28, 0.47]	<0.001	8.28	0.340
	fear	0.18 [0.07, 0.27]	0.002	0.18 [0.07, 0.27]	0.003	9.83	0.829
	happy	0.24 [0.13, 0.33]	<0.001	0.23 [0.13, 0.32]	<0.001	9.45	0.654
	sad	0.13 [-0.0, 0.23]	0.031	0.11 [-0.02, 0.22]	0.072	10.80	0.649
	Specificity Analyses						

	brightness	0.26 [0.17, 0.32]	<0.001	0.28 [0.19, 0.33]	<0.001	11.80	0.959
	loudness	0.11 [0.02, 0.18]	0.010	0.11 [0.03, 0.18]	0.009	15.10	0.999
	spoken words	0.34 [0.22, 0.4]	<0.001	0.32 [0.22, 0.4]	<0.001	9.91	0.570
	written words	0.26 [0.14, 0.35]	<0.001	0.27 [0.15, 0.35]	<0.001	9.17	0.600
Despicable Me	**Puberty**						
	Main Analyses						
	negative	0.21 [0.07, 0.32]	0.004	0.21 [0.09, 0.31]	0.003	21.70	0.520
	positive	0.19 [0.06, 0.3]	0.008	0.18 [0.06, 0.29]	0.011	22.20	0.525
	anger	0.35 [0.21, 0.44]	<0.001	0.31 [0.18, 0.41]	<0.001	19.03	0.385
	excite	0.2 [0.08, 0.31]	0.006	0.19 [0.07, 0.3]	0.009	21.40	0.511
	fear	0.18 [0.05, 0.29]	0.011	0.2 [0.07, 0.3]	0.007	20.14	0.493
	happy	0.18 [0.06, 0.29]	0.012	0.19 [0.06, 0.29]	0.009	20.80	0.502
	sad	0.11 [-0.02, 0.22]	0.132	0.11 [-0.02, 0.22]	0.141	22.39	0.495
	Specificity Analyses						
	brightness	0.25 [0.13, 0.33]	<0.001	0.23 [0.11, 0.31]	<0.001	24.39	0.728
	loudness	0.14 [0.05, 0.23]	0.006	0.14 [0.04, 0.22]	0.008	30.35	0.947
	spoken words	0.32 [0.19, 0.41]	<0.001	0.32 [0.2, 0.41]	<0.001	19.52	0.348
	written words	0.16 [0.04, 0.28]	0.026	0.19 [0.06, 0.29]	0.009	21.41	0.553

Table SB4: Maturity support vector regression results when limiting the activation data to *The Present*. Results indicate model performance in the unseen test data, using a liner kernel. These models did not perform significantly better than when using activation maps from both videos (see Table 1 in the main text for comparison). All statistical tests were one-sided.

	Activation Data	Spearman *r* [95% CI]	Spearman p-value	Pearson *r* [95% CI]	Pearson p-value	MSE	MSE p-value
	Chronological Age						
	Main Analyses						
	negative	0.23 [0.12, 0.31]	<0.001	0.24 [0.13, 0.32]	<0.001	8.56	0.140
	positive	0.09 [-0.01, 0.18]	0.095	0.12 [0.02, 0.2]	0.034	11.15	0.447
	anger	0.2 [0.08, 0.27]	<0.001	0.21 [0.09, 0.29]	<0.001	10.22	0.089
The Present	excite	0.16 [0.05, 0.24]	0.004	0.15 [0.04, 0.23]	0.006	9.94	0.250
	fear	0.15 [0.04, 0.23]	0.006	0.16 [0.05, 0.23]	0.004	10.15	0.122
	happy	0.14 [0.04, 0.23]	0.011	0.13 [0.02, 0.21]	0.022	9.75	0.361
	sad	0.21 [0.11, 0.29]	<0.001	0.21 [0.11, 0.28]	<0.001	8.98	0.259
	Specificity Analyses						
	brightness	0.24 [0.15, 0.31]	<0.001	0.25 [0.15, 0.3]	<0.001	12.90	0.616

	loudness	0.2 [0.11, 0.27]	<0.001	0.21 [0.11, 0.26]	<0.001	9.66	0.276
	spoken words	0.13 [0.03, 0.22]	0.022	0.14 [0.04, 0.23]	0.009	11.75	0.424
	written words	0.13 [0.02, 0.22]	0.020	0.14 [0.03, 0.22]	0.014	10.14	0.403
The Present	**Puberty**						
	Main Analyses						
	negative	0.23 [0.1, 0.33]	<0.001	0.25 [0.13, 0.33]	<0.001	18.86	0.213
	positive	0.1 [-0.02, 0.2]	0.148	0.1 [-0.02, 0.2]	0.121	23.95	0.488
	anger	0.15 [0.03, 0.27]	0.019	0.2 [0.07, 0.3]	0.003	22.97	0.199
	excite	0.11 [-0.02, 0.22]	0.106	0.13 [0.01, 0.23]	0.055	24.48	0.526
	fear	0.07 [-0.05, 0.19]	0.275	0.13 [0.01, 0.22]	0.055	22.22	0.221
	happy	0.23 [0.1, 0.32]	<0.001	0.23 [0.1, 0.33]	<0.001	19.23	0.261
	sad	0.19 [0.06, 0.3]	0.005	0.21 [0.09, 0.31]	0.001	19.57	0.301
	Specificity Analyses						
	brightness	0.11 [-0.03, 0.22]	0.015	0.11 [-0.01, 0.22]	0.021	32.70	0.884
	loudness	0.12 [0.02, 0.19]	0.011	0.14 [0.05, 0.21]	0.003	21.50	0.328
	spoken words	-0.02 [-0.15, 0.11]	0.771	0.05 [-0.1, 0.17]	0.48	23.37	0.380
	written words	0.13 [0.0, 0.24]	0.047	0.14 [0.02, 0.25]	0.033	21.29	0.379

B4: Predicting site of scan and motion metrics from activation maps

Table SB5: Support vector classification results predicting site of scan from data. No model performed better than chance.

Data Set	Train Accuracy	Test Accuracy	Test 95% CI	Permuted p-value
all	54%	53%	[52%, 54%]	0.066
negative	56%	57%	[53%, 61%]	0.164
positive	60%	58%	[54%, 62%]	0.111

Table SB6: Support vector regression results predicting mean framewise displacement (FD) from activation maps. Models were not accurate as evidenced by high mean square error (MSE). A modest association between predicted and actual labels were found for models predicting mean motion from negative and fear activation maps, suggesting a weak difference in activation with increasing mean motion. All statistical tests were one-sided.

Activation Data	Spearman *r* [95% CI]	Spearman p-value	Pearson *r* [95% CI]	Pearson p-value	MSE	MSE p-value
negative	0.15 [0.06, 0.21]	<0.001	0.15 [0.06, 0.21]	<0.001	0.07	0.966
positive	-0.02 [-0.09, 0.05]	0.627	-0.01 [-0.08, 0.06]	0.751	0.08	0.983
anger	0.08 [-0.01, 0.15]	0.059	0.10 [0.01, 0.16]	0.017	0.08	0.885
excite	0.04 [-0.05, 0.11]	0.388	0.04 [-0.04, 0.12]	0.296	0.08	0.985
fear	0.11 [0.02, 0.17]	0.007	0.11 [0.02, 0.17]	0.006	0.08	0.970
happy	0.09 [0.01, 0.15]	0.028	0.08 [-0.01, 0.14]	0.060	0.07	0.990
sad	0.04 [-0.04, 0.11]	0.327	0.05 [-0.04, 0.12]	0.248	0.08	0.998

Appendix C: Supplemental Figures and Tables

Table SC1: Demographic characteristics of the final sample. Puberty was measured using the Peterson Puberty Scale[16] and motion was measured as framewise displacement[17]. Statistical tests were two-sided.

	Low-Motion Data (N=823)		
Characteristic	**Discovery/ Training (N=424)**	**Replication/ Testing (N=399)**	**Statistic p-value**
Age *Mean years (SD)*	10.3 (2.7)	10.5 (2.8)	t=1.38 p=0.167
Puberty *Mean (SD)*	9.5 (4.1)	9.7 (4.2)	t=0.72 p=0.471
Male *%*	63%	58%	X^2= 1.99 p=0.159
Right-handed *%*	75%	74%	X^2= 0.35 p=0.743
Motion *Mean mm (SD)*	0.27 (0.15)	0.45 (0.17)	t=15.18 p<0.001

Table SC2: Emotion activation classification and maturity prediction results. For each analysis, activation maps across the two videos were pooled. Models were trained on data from the RUBIC site (Discovery) using ten-fold cross validation with participant ID entered as a grouping variable and tested on unseen data from the CBIC site (Replication). Support vector regression was not able to accurately predict maturity using any of the activation maps. Modest associations between maturity indices and activation were found, suggesting modest changes in activation across maturity. All statistical tests were one-sided.

Activation Data Used	Spearman *r* (p-value)	Spearman *r* 95% CI	Pearson *r* (p-value)	Pearson *r* 95% CI	MSE	MSE p-value
Chronological Age						
Main Analyses						
negative	0.16 (<0.001)	[0.07, 0.23]	0.17 (<0.001)	[0.08, 0.23]	9.50	0.364
positive	0.15 (<0.001)	[0.06, 0.20]	0.12 (0.002)	[0.04, 0.18]	11.25	0.685
anger	0.17 (<0.001)	[0.09, 0.24]	0.18 (<0.001)	[0.10, 0.23]	10.44	0.313
excite	0.13 (<0.001)	[0.05, 0.19]	0.13 (0.002)	[0.04, 0.18]	10.35	0.520
fear	0.15 (<0.001)	[0.07, 0.21]	0.14 (<0.001)	[0.06, 0.19]	10.39	0.394
happy	0.09 (0.022)	[0.01, 0.16]	0.1 (0.014)	[0.02, 0.16]	10.19	0.626
sad	0.16 (<0.001)	[0.08, 0.21]	0.15 (<0.001)	[0.07, 0.21]	10.24	0.535
Specificity Analyses						
brightness	0.18 (<0.001)	[0.11, 0.22]	0.17 (<0.001)	[0.10, 0.20]	12.31	0.882
loudness	0.10 (<0.001)	[0.03, 0.15]	0.08 (0.007)	[0.01, 0.13]	12.47	0.985
speaking	0.11 (0.009)	[0.03, 0.17]	0.11 (0.005)	[0.04, 0.17]	12.26	0.788
words	0.12 (0.003)	[0.05, 0.18]	0.14 (<0.001)	[0.06, 0.19]	10.42	0.672
Puberty Scores						
Main Analyses						
negative	0.14 (0.003)	[0.05, 0.21]	0.14 (0.004)	[0.05, 0.21]	22.07	0.414
positive	0.07 (0.172)	[-0.03, 0.16]	0.08 (0.096)	[-0.01, 0.16]	24.10	0.530
anger	0.10 (0.045)	[0.01, 0.18]	0.12 (0.016)	[0.02, 0.20]	24.50	0.413
excite	0.09 (0.056)	[0.00, 0.17]	0.07 (0.137)	[-0.01, 0.15]	22.98	0.454
fear	0.08 (0.123)	[-0.02, 0.16]	0.10 (0.038)	[0.02, 0.17]	23.78	0.428
happy	0.16 (0.001)	[0.07, 0.23]	0.15 (0.001)	[0.07, 0.23]	22.23	0.487
sad	0.14 (0.003)	[0.05, 0.22]	0.13 (0.007)	[0.03, 0.21]	21.51	0.347
Specificity Analyses						
brightness	0.17 (<0.001)	[0.08, 0.22]	0.15 (<0.001)	[0.07, 0.20]	28.68	0.806
loudness	0.10 (0.005)	[0.02, 0.15]	0.09 (0.009)	[0.02, 0.15]	25.74	0.737
speaking	0.13 (0.007)	[0.04, 0.21]	0.14 (0.004)	[0.04, 0.22]	24.17	0.478
words	0.10 (0.043)	[-0.00, 0.18]	0.10 (0.048)	[0.01, 0.18]	22.92	0.533

Figure SC1: A schematic illustration of our analytical approach**.**

1) Support vector classification was conducted to test if activation to contextualized emotions were dissociable and where in the brain emotion-specific information was represented. **2)** Support vector regression was used to examine linear and curvilinear associations between activation to emotions and maturity. **3)** Inter-subject representational similarity analysis (IS-RSA) was used to test which nonlinear model of development fit the data best by comparing 3 different maturity similarity metrics and neural activation similarity. **4)** dynamic similarity analysis was used to identify scenes within each video that evoked high synchrony across the sample and in the age group identified as more similar in the IS-RSA. Identified scenes were then examined quantitatively and qualitatively.

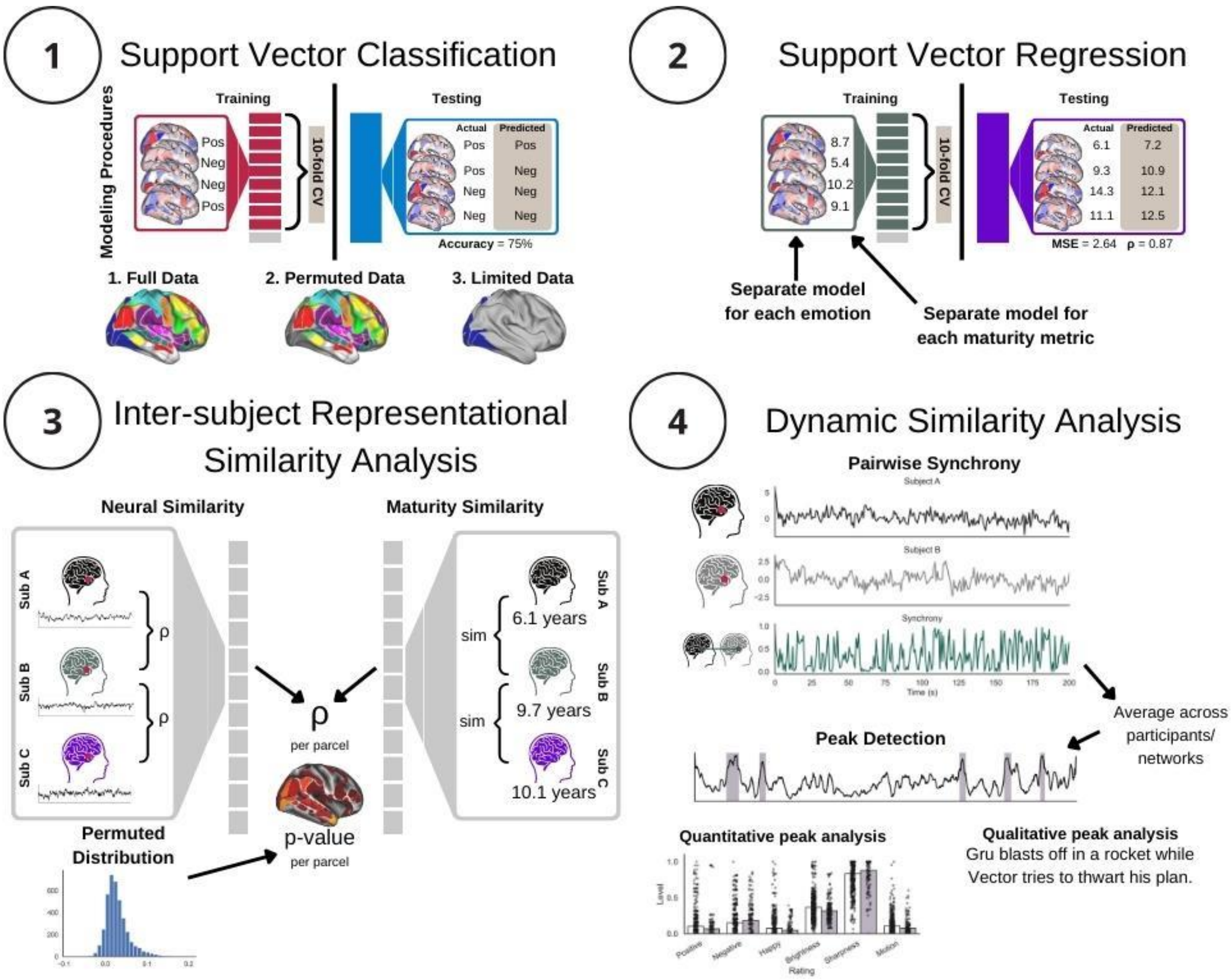
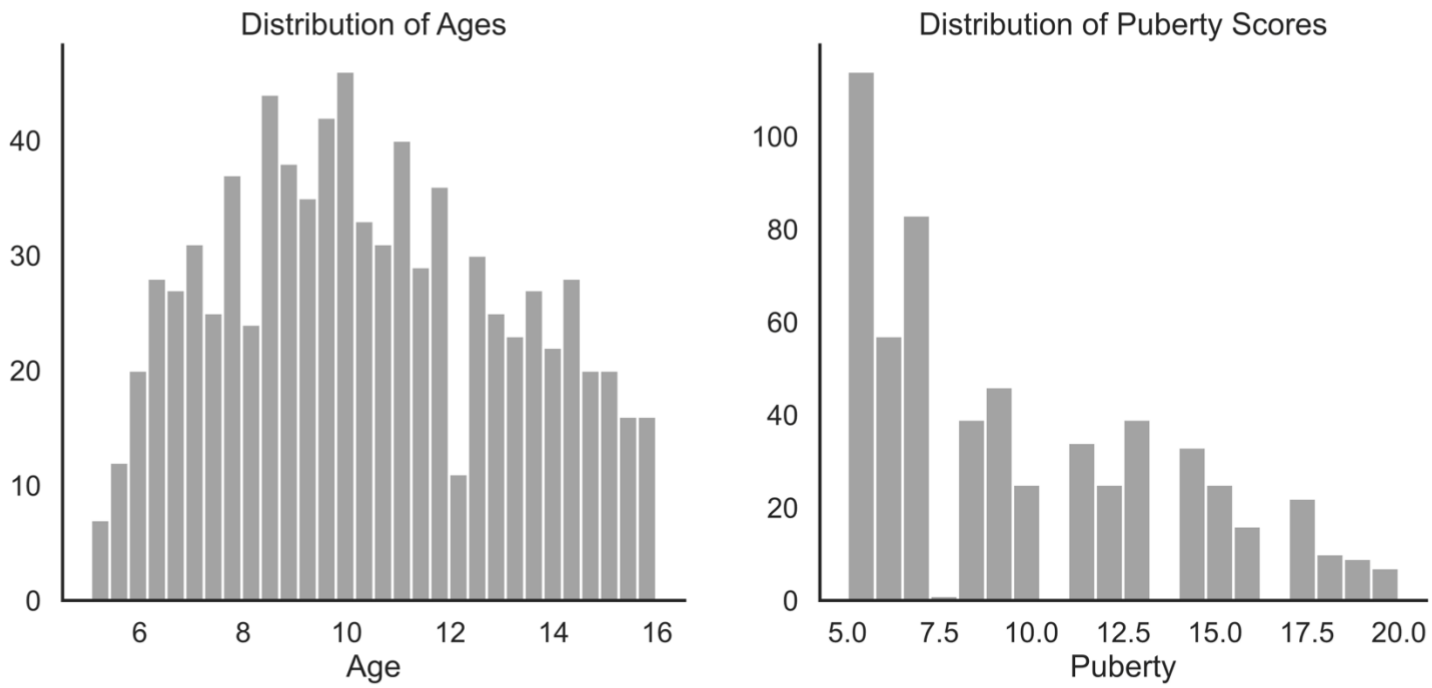


Figure SC2: Distribution plots of age and puberty scores for the final sample of participants (N=823).



References

1. Camacho MC, Williams EM, Balser D, et al. EmoCodes: a Standardized Coding System for Socio-emotional Content in Complex Video Stimuli. *Affect Sci*. Published online January 20, 2022. doi:10.1007/s42761-021-00100-7

2. Eickhoff SB, Milham M, Vanderwal T. Towards clinical applications of movie fMRI. *NeuroImage*. 2020;217:116860. doi:10.1016/j.neuroimage.2020.116860

3. Finn ES, Bandettini PA. Movie-watching outperforms rest for functional connectivity-based prediction of behavior. *NeuroImage*. 2021;235:117963. doi:10.1016/j.neuroimage.2021.117963

4. Tracy JL, Randles D. Four Models of Basic Emotions: A Review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt. *Emot Rev*. 2011;3(4):397-405. doi:10.1177/1754073911410747

5. Panksepp J, Watt D. What is Basic about Basic Emotions? Lasting Lessons from Affective Neuroscience. *Emot Rev*. 2011;3(4):387-396. doi:10.1177/1754073911410741

6. Lettieri G, Handjaras G, Ricciardi E, et al. Emotionotopy in the human right temporo-parietal cortex. *Nat Commun*. 2019;10(1):5568. doi:10.1038/s41467-019-13599-z

7. Lettieri G, Handjaras G, Setti F, et al. Default and control network connectivity dynamics track the stream of affect at multiple timescales. *Soc Cogn Affect Neurosci*. 2022;17(5):461-469. doi:10.1093/scan/nsab112

8. Satpute AB, Lindquist KA. The Default Mode Network's Role in Discrete Emotion. *Trends Cogn Sci*. Published online 2019. doi:10.1016/j.tics.2019.07.003

9. Kragel PA, LaBar KS. Multivariate neural biomarkers of emotional states are categorically distinct. *Soc Cogn Affect Neurosci*. 2015;10(11):1437-1448. doi:10.1093/scan/nsv032

10. Kragel PA, Knodt AR, Hariri AR, Labar KS. Decoding Spontaneous Emotional States in the Human Brain. *PLOS Biol*. Published online 2000. doi:10.1371/journal.pbio.2000106

11. Kragel PA, Labar KS. Decoding the Nature of Emotion in the Brain. *Trends Cogn Sci*. 2016;20(6):444-455. doi:10.1016/j.tics.2016.03.011

12. Chang LJ, Jolly E, Cheong JH, et al. Endogenous variation in ventromedial prefrontal cortex state dynamics during naturalistic viewing reflects affective experience. doi:10.1101/487892

13. Ruba AL, Pollak SD. The Development of Emotion Reasoning in Infancy and Early Childhood. *Annu Rev Dev Psychol*. 2020;2(1):503-531. doi:10.1146/annurev-devpsych-060320-102556

14. Widen SC. Children's Interpretation of Facial Expressions: The Long Path from Valence-Based to Specific Discrete Categories. *Emot Rev*. 2013;5(1):72-77. doi:10.1177/1754073912451492

15. Nook EC, Sasse SF, Lambert HK, McLaughlin KA, Somerville LH. The Nonlinear Development of Emotion Differentiation: Granular Emotional Experience Is Low in Adolescence. *Psychol Sci*. 2018;29(8):1346-1357. doi:10.1177/0956797618773357

16. Petersen AC, Crockett L, Richards M, Boxer A. A self-report measure of pubertal status: Reliability, validity, and initial norms. *J Youth Adolesc*. 1988;17(2):117-133. doi:10.1007/BF01537962

17. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*. 2014;84:320-341. doi:10.1016/j.neuroimage.2013.08.048