Whole Brain Social Incentive Delay fMRI Prereg

Jessica Flannery

10/21/2020

Working Title

Three-Time-Point Longitudinal Changes in Adolescent Social Reward Processing: Differences in Neural Social Anticipation and Receipt

Introduction

Adolescence is a pivotal period of social-affective learning, often denoted by hormonal changes and structural and functional maturation changes within and between regions involved in learning, reward-processing, social cognition, and regulation. Reward processing, both anticipation and receipt are critical processes for learning and associated with individual differences in risk-taking behaviors.

Reward processing overview.

The social incentive delay (SID) task is designed to assess reward processing and allows for the assessment of both reward anticipation and receipt, specifically as it relates socially-relevant stimuli. Compared to the monetary incentive delay (MID), the social incentive delay (SID) task uses socially relevant information as the target of reward. SID tasks are shown to recruit similar regions, but thought to also evoke regions associated with social-cognition, such at the TPJ, dmPFC, precuneus, STG (Greimel, et al., 2018; Martins et al., 2020). An adult meta analysis of reward anticipation demonstrated increased activation in the amygdala, striatum, left IFG and decreased activation in regions such as the bilateral IFG, angular gyri, and right parahippocampal gyrus. Punishment anticipation showed similar increased activation in the amygdala, striatum, thalamus, and frontal gyri, but decreased activation in the paracingulate cortex, bilateral frontal poles, left temporal pole, bilateral SMG. Overlap between anticipated reward and punishment include caudate nucleus, bilateral putamen, thalamus, frontal medial cortex, right insula, and bilateral amygdala (Martins et al., 2020).

In the same adult meta analysis social reward receipt included increased activation in regions such as anterior cingulate, bilateral amygdala and hippocampus, bilateral medial frontal cortex and decreased activation in regions such as bilateral superior frontal gyrus, right role, insula, and bilateral precuneus. Receipt of social punishment, however, showed increased activation in the anterior insula, IFG and relatively decreased activation in precuneus, angular gyrus, left caudate and right putamen. Anticipation of reward and receipt of reward overlapped in regions of frontal medial cortex, cingulate cortex, thalamus, left amygdala, and STG and right superior frontal gyrus (Martins et al., 2020).

Connectivity differences between anticipation and receipt.

Although similar regions have been implicated in both anticipation and receipt of social feedback, a large cross-sectional adolescent study demonstrated that connectivity in key reward regions differed between anticipation and receipt. During reward anticipation, both thalamus and VS showed similar positive connections to anterior supplementary motor area (pre-SMA), dorsal anterior cingulate cortex (dACC), left anterior insula, and medial occipital lobe (Cao et al, 2019). During reward receipt, the authors suggest the projections to OFC and insula may signal reward evaluation (Cao et al., 2019; Silverman et al., 2015). Cao et al., 2019 only found reliable engagement for the vmPFC (.8 probability), which is described to reflect reward magnitude. During reward receipt, the vMPFC was negative connected to insula and dorsal striatum. Authors suggest that VmPFC-Dorsal striatum may represent learning about the reward and vmPFC-anterior insula reflecting

valuation for reward. For negative receipt, there was a main effect of decreased ventral straitum. "VS positive connections with medial frontal gyrus, dACC, pre-SMA, precentral gyrus, putamen, insula, thalamus, precuneus, cuneus, lingual gyrus, and calcarine sulcus." (Cao et al, 2019).

Development of reward processing.

Common approaches to studying developmental changes in reward processing have focused on cross-sectional comparisons between children, adolescents, and adults. In particular, these studies have focused on changes in monetary versus social reward processing. While some studies have found no differences between monetary and social reward processing during adolescence, suggesting a general reward process, other studies suggest that inconsistent patterns of activation between reward anticipation and receipt are due to variability in the reward value and may additionally engage regions involved in social cognition. In cross-sectional comparison studies across development, adults show a relatively stronger response to monetary incentives, whereas adolescents showed either no differentiation, or show greater response to social incentives. In another cross-sectional age comparison study, Hoogendam et al., 2013, found a linear increase in reward-associated regions during reward receipt and linear decrease in these regions during reward receipt. In direct comparison, they suggest results indicate a developmental shift toward increased reward anticipation versus reward receipt across age (Hoogendam et al., 2013). The dual-process model of behavioral and neurobiological development posits this is due to mismatch in neural maturation, such that limbic regions involved in reward processing, such as the ventral striatum and amygdala peak during adolescence, and regions involved in regulatory, such as the lateral prefrontal cortex and dmPFC are still developing.

However, patterns of reward processing between anticipation and receipt are not consistent within adolescent studies or cross-sectional age comparison studies. In particular, ventral striatum involvement in both anticipation and receipt of social reward results are mixed, with studies showing both hyper or hypo-activation of the ventral striatum during either anticipation and receipt (Bjork et al., 2010; Ernst et al., 2005, Galvan et al., 2006; Hoogendam et al., 2013). Mixed findings regarding VS engagement during reward receipt may in part be due to the type of outcome (i.e., how rewarding it is perceived; Cao et al, 2019; Flores et al., 2015), if outcome is more or less rewarding than anticipated (prediction error; Cao et al., 2019), or age (Hoogendam et al., 2013).

Furthermore, while in some cross-sectional studies, adolescents show either similar or increased reward activation to social versus monetary rewards in regions involved reward valuation, processing and learning such as the striatum, nucleus accumbens, caudate, insula, a recent large-scale cross-sectional study of adolescent reward processing, suggests reward receipt only reliably engaged the ventral medial prefrontal cortex (Cao et al, 2019). Cao et al., 2019 suggest that vmPFC engagement encodes reward receipt, whereas VS encodes prediction error of reward receipt.

Pubertal development and sex difference may be another reason for inconsistent results. Forbes, et al., 2010 found that later pubertal development was associated with decreased ventral striatum and increased medial PFC to rewards. exception of whole brain developmental affective processes. In a recent longitudinal assessment of affective processing during adoelscence, while both age and puberty showed a positive quadratic effect for activity in the hippocampus and amygdala, testosterone levels showed a negative quadratic in the hippocampus and bilateral amygdala (Vijayakumar, et al., 2019). At the whole brain, males and females showed different developmental patterns in the perigenual ACC. For age, males showed a postive qadtratic effect, but a negative quaratic effect for puberty, measured by the petersen Development Scale (PDS). Relatedly, negative quadratic effects were shown for age in regions across the prefrontal cortex and positive quadratic effects were shown for age in the bilateral precentral, right postcentral cortices, and the right lingual gyrus (Vijayakumar, et al., 2019). Togehther, these findings suggest that while ROI analyses may be more consistent with expected limbic-related peak activation during adolescence, that whole brain analyses may show more complicated patterns of both negative and positve quadratic effects across neural networks. Furthermore, these age developmental patterns appear to differ across sex, pubertal stage, and testosterone.

Gaps.

To date, majority of studies have either used cross-sectional age comparisons or used region of interest approaches to assess developmental changes in reward processing during adolescence. With the exception

of one study that assessed this question cross-sectionally, we are unaware of other studies that assess the age-expected differences between reward anticipation and receipt. Furthermore, there is considerably less information on how anticipated or receipt of punishment or lack of social reward change across development. Given the mixed findings, particularly in regards in VS activity and the engagement of regions involved in social cognition, expanding this question beyond longitudinal assessment for *a priori* regions of interest (i.e., longitudinal whole brain changes) may provide unique insight into these developmental processes. Lastly, although recent studies suggest connectivity analyses can help tease apart conflicting findings regarding reward processing during anticipation and receipt, to our knowledge, this has not been assessed to see if and how longitudinal changes in these regions influences patterns of reward processing.

Current Study

The current study uses a three-year longitudinal design to assess how neural response to anticipated and receipt of social reward and punishment change across adolescence using a social incentive delay task. Specifically, the task was designed to use socially-relevant stimuli: age-match peers. We plan to assess maturation changes in whole brain reward anticipation and receipt by age and puberty. Given the mixed VS findings, we also plan to assess longitudinal changes in VS connectivity during reward anticipation and receipt.

Aims & Hypotheses

1. How do whole brain neural responses to anticipated peer reward and punishment change across development? Do reward and punishment show unique trajectories?
2. How do whole brain neural responses change for neural receipt of social reward and threat? Does social receipt versus social reward show different developmental patterns? Does social receipt show a different developmental pattern than social anticipation?

For both main hypotheses, we will assess the between subject linear and quadratic effects of age and puberty. Due to our study design of an accelerated cohort longitudinal design, participants spanned across grades, and some participants changed school environments (i.e., transitioned to high school) across the three time points. Therefore, sensitivity analyses will be run with grade and the quadratic grade term.

By design, this is an exploratory analysis. The aim of these questions is to expand our understanding of developmental changes in neural development beyond a prior ROI approaches or cross-sectional age comparisons. By assessing both age and pubertal developmental effects, we aim to differentiate how functional patterns of social anticipation and receipt may change as factor of age, but also as a factor of pubertal development. Initial hypotheses will focus on pubertal changes, but sensitivity analyses will be run to see if this differs as factor of assessing pubertal timing. Details explained in the analysis section.

That said, we have *a priori* hypotheses about the neural regions linear or non-linear effects based on prior studies that have either assessed cross-sectional age differences in the SID or assessed longitudinal changes using regions of interest analyses.

Overall, if our findings were aligned with common neurobiological models of adolescent development, we would expect to see positive nonlinear (inverted U-shape) change in limbic regions, such as the ventral striatum and amygdala, and a negative linear effect in pre-frontal and regulatory associated regions, such as the dorsal medial pre-frontal cortex, cingulate cortex, and inferior frontal gyris.

Prior studies would suggest sex differences may be stronger for anticipation than receipt, specifically increase left putamen, middle temporal gyrus, and right precuneus activity for males (Cao et al., 2019). This also exists in adulthood. We therefore predict this difference may increase across age.

We'd further expect these patterns to differ based on a) anticipation or receipt and b) response to reward and punishment. Across age, if our findings aligned with Hoogendam et al., 2013, we would predict increased engagement of the VS during reward anticipation and decreased engagement of VS during social receipt. If we saw increased engagement in VS during social punishment receipt and increased engagement in vmPFC during social reward receipt, that would be aligned evidence that VS is associated with reward prediction error and vmPFC is more associated with reward receipt (Cao et al., 2019).

In addition to regions commonly described in this task, by taking a whole brain approach, we anticipate nonlinear engagement in regions involved in social cognition and the default mode network will also change across time. In particular, we anticipate nonlinear positive (inverted U shape) engagement in regions such as the angular gyrus, temporal parietal junction, SMG, temporal poles, and IPL. Based on Vijayakumar, et al., 2019, we also expect these whole-brain longitudinal changes to show similar patterns of negative and positive quadratic effects of affective processing that may differ based on pubertal stage, age, and sex. Given the difference in our paradigms, we don't anticipate identical effects, but expect similarity in social reward and punishment receipt.

**Essential	Checklist**
Hypotheses include direction of expected results	*X*
Interactions describe expected shape	*X*
Manipulated variables include manipulation checks or explain why not	*NA*

**Recommended	Checklist**
Figure or table to describe expected results	___
Rationals or frameworks included for why certain hypotheses are being tested	*X*
Which outcome would be predicted by which theory	*X*

Adapted from van't Veer & Giner-Sorolla, 2016

Exisiting Data

Preregistration is designed to make clear the distinction between confirmatory tests, specified prior to seeing the data, and exploratory analyses conducted after observing the data. Therefore, creating a research plan in which existing data will be used presents unique challenges. Please select the description that best describes your situation.

- Registration prior to analysis of the data
- Registration following analysis of the data

Details of which analysis have and have not been run are detailed below.

Explanation of Exisiting Data

If you indicate that you will be using some data that already exist in this study, please describe the steps you have taken to assure that you are unaware of any patterns or summary statistics in the data. This may include an explanation of how access to the data has been limited, who has observed the data, or how you have avoided observing any analysis of the specific data you will use in your study. The purpose of this question is to assure that the line between confirmatory and exploratory analysis is clear. e.g., links to prior papers, osf project page, prior posters or talks, or descriptions.

One time point of these data has been analyzed and published here for social anticipation of feedback: Telzer et al., 2020, https://srcd.onlinelibrary.wiley.com/doi/pdf/10.1111/cdev.13466?casa_token=-XqDB1adchEAAAAA:9DWPBFCCFhuU3g9VzOcBHVhWjra2YOOYH9tlXZRE2asKBzAUx_dO9XWMeSHCWSco6Jdq7US40-hbSg

These findings provide the foundation of expected effects for the task, but do not assess longitudinal changes in these patterns. Patterns of activation from the first time point will be used as a basis of elicited activation from this task, but it does not provide insight into expected developmental changes. Furthermore, this first paper does not focus on the neural effects of social receipt.

In an effort to learn if and how we could model three points of data at the whole brain, afni 3dLME models were developed and run with linear quadratic terms for age and grade.

Preliminary model code included

```
#!/bin/bash
#---------------------------
#

3dLME -prefix 3dLME_v1 -jobs 4 \
    -model 'Ant*age' \
    -qVars "age" \
    -qVarCenters "0" \
    -resid grade_resids \
    -ranEff '~1+age' \
    -SS_type 3 \
    -num_glt 10 \
    -gltLabel 1 'Reward' -gltCode 1 'Ant : 1*rew' \
    -gltLabel 2 'Punishment' -gltCode 2 'Ant : 1*pun' \
    -gltLabel 3 'Rew_Neu' -gltCode 3 'Ant : 1*rew -1*neu' \
    -gltLabel 4 'Pun_Neu' -gltCode 4 'Ant : 1*pun -1*neu' \
    -gltLabel 5 'Rew_Pun' -gltCode 5 'Ant : 1*rew -1*pun' \
    -gltLabel 6 'Val_neu' -gltCode 6 'Ant : 1*rew & 1*pun -1*neu' \
    -gltLabel 7 'age.pun-neu' -gltCode  7 'Ant : 1*pun -1*neu age : ' \
    -gltLabel 8 'age.rew-neu' -gltCode  8 'Ant : 1*rew -1*neu age : ' \
    -gltLabel 9 'age.val' -gltCode  9 'Ant : 1*rew & 1*pun age : ' \
    -gltLabel 10 'age.val-neu' -gltCode  10 'Ant : 1*rew & 1*pun -1*neu age : ' \
    -dataTable \
```

and

```
#!/bin/bash
#---------------------------
#

3dLME -prefix 3dLME_age2 -jobs 4 \
    -model 'Ant*age_c+Ant*age2_c' \
    -qVars "age_c,age2_c" \
    -qVarCenters "0,0" \
    -resid age2_resids \
    -ranEff '~1+age_c' \
    -SS_type 3 \
    -num_glf 4 \
    -glfLabel 1 'val_age' -glfCode 1 'Ant : 1*rew & 1*pun age_c :' \
    -glfLabel 2 'val_age2' -glfCode 2 'Ant : 1*rew & 1*pun age2_c :' \
    -glfLabel 3 'rew_age' -glfCode 3 'Ant : 1*rew age_c :' \
    -glfLabel 4 'pun_age2' -glfCode 4 'Ant : 1*pun age2_c :' \
    -num_glt 17 \
    -gltLabel 1 'Reward' -gltCode 1 'Ant : 1*rew' \
    -gltLabel 2 'Punishment' -gltCode 2 'Ant : 1*pun' \
    -gltLabel 3 'Rew_Neu' -gltCode 3 'Ant : 1*rew -1*neu' \
    -gltLabel 4 'Pun_Neu' -gltCode 4 'Ant : 1*pun -1*neu' \
    -gltLabel 5 'Rew_Pun' -gltCode 5 'Ant : 1*rew -1*pun' \
```

```
-gltLabel 6 'Val_neu' -gltCode 6 'Ant : 1*rew & 1*pun -1*neu' \
-gltLabel 7 'age.pun-neu' -gltCode  7 'Ant : 1*pun -1*neu age_c : ' \
-gltLabel 8 'age.rew-neu' -gltCode  8 'Ant : 1*rew -1*neu age_c : ' \
-gltLabel 9 'age.val' -gltCode  9 'Ant : 1*rew & 1*pun age_c : ' \
-gltLabel 10 'age.val-neu' -gltCode  10 'Ant : 1*rew & 1*pun -1*neu age_c : ' \
-gltLabel 11 'age.rew-pun' -gltCode  11 'Ant : 1*rew -1*pun  age_c : ' \
-gltLabel 12 'age2.pun-neu' -gltCode  12 'Ant : 1*pun -1*neu age2_c : ' \
-gltLabel 13 'age2.rew-neu' -gltCode  13 'Ant : 1*rew -1*neu age2_c : ' \
-gltLabel 14 'age2.val' -gltCode  14 'Ant : 1*rew & 1*pun age2_c : ' \
-gltLabel 15 'age2.val-neu' -gltCode  15 'Ant : 1*rew & 1*pun -1*neu age2_c : ' \
-gltLabel 16 'age2.val-neu' -gltCode  16 'Ant : 1*rew & 1*pun -1*neu age2_c : ' \
-gltLabel 17 'age2.rew-pun' -gltCode  17 'Ant : 1*rew -1*pun  age2_c : ' \
-dataTable \
Subj Ant grade_c grade2_c age_c age2_c InputFile \
```

Notably, these test models did not contain the final version of t and f tests, or the full set up subjects. These models were not FWE corrected and only assessed to test model function and approripriate modeling strategy. These preliminary models and understanding of model strategy interpretation were discussed with afni message board. These models will be updated for the manuscript analysis, as shown below. In addition to model updates, these models were either a) not centered (model 1) or centered at mean age, instead of our planned centering at min age. In addition, these models do not include sex, which we plan to include in our final model.

While we examined these models to see if they ran properly within the basic output, no effects were extracted to examine the pattern of developmental patterns. Furthermore, any preliminary patterns, aligned or not aligned with prior hypotheses, will not influence analysis plans or hypotheses. No analysis have been done with puberty data, ROIs, or behavior for this paper. We have not run our whole brain 3dLME model for puberty or any follow up ROI analyses and have not graphed any relationships.

Details of Larger Study

Is your preregistration part of a larger project?

- Yes

If yes, provide a brief description of the larger study. Note, this does not need to include a list of all measures included in the larger study, but it is meant to provide context for the larger scope of the project.

This project is part of two larger projects, NeuroTeen and Teen Transition Project. Teen Transition Project is a project that recruited and obtained data collection in the schools. NeuroTeen is the subset of participants from Teen Transition Project that came into the lab yearly. At the lab visit, they completed several meausures, including several questionnaires on mood, risk behaviors, and family and friend relationships, several biological measures, and an fMRI scan.

Explanation of how information from larger study or related studies within the larger project have/have not influenced your hypotheses/ measurement decisions.

To date, longitudinal data from Neuroteen and Teen Transition Project have not been published. The specific task that is the focus of this study, a social incentive delay, has been assessed by lab members in relation to other research questions, but the only analysis details that I am aware of at the time of this preregistration is the published paper at time one of this study (Telzer, et al., 2020).

Data Collection Procedures

Please describe the process by which you will collect your data. If your data are already collected, write out collection procedures as you would for your manuscript.

After your paragraph, use the table as a checklist to ensure you included all information suggested for reproducible results.

```
Data was collected prior to preregistration.
Prior to publication, these details will be confirmed with the current sample.
```

Based on data from first paper at time point one, participants were recruited from three rural public middle schools in the southeast United States. Between 66.7% and 72.1% of students in these schools was classified as economically disadvantaged based on school reports (North Carolina School Report Cards, 2017), and 69.5% of students in the district was eligible for free or reduced-price lunch based on district reports. Participants were recruited from a larger study of 873 students in sixth and seventh grade. Participants from the larger study provided interest in being contacted for a future fMRI study. Interested participants were then called and screened on the phone for eligibility (i.e., MRI contraindications) and recruited for the fMRI study within the same academic year as the larger study. No other exclusion criteria were used. We screened 284 families, of whom 91 were ineligible due to learning disabilities, braces, head trauma, or other MRI contraindications, and 45 were eligible but did not participate due to scheduling difficulties or no longer interested in participating, resulting in a final sample of 148 adolescents. Thus, of those contacted and eligible, 77.5% participated. However, at time two, an additional 30 adolescents were recruited to the sample. Therefore, **X%** completed up to two timepoints. For this study, adolescents and their primary caregiver attended the fMRI session, during which consent, and assent were obtained. Participants completed an fMRI scan that lasted approximately1.5hr, during which they completed the SID task (described in the below section), as well as four other tasks that are not the focus of this manuscript. Following the scan, participants completed several self-report measures using computer-assisted soft-ware in a private room and provided other hormonal measures, not a focus of this manuscript. Adolescents were compensated with a monetary remuneration of 90 dollars, small prizes for completing the full scan and staying still (e.g., headphones, candy; worth 20 dollars), snacks during the visit, and a meal. Parents were compensated with a monetary remuneration of $50, as well as a meal, compensation for gas, and parking. Participants completed the lab sessions and fMRI session once a year for three years, following the sample procedures detailed for wave 1. Adolescents and parents gave written assent/consent in accordance with the university's Institutional Review Board at each time point.

**Collection	Checklist**
Population	*X*
Recruitment efforts	*X*
Inclusion/Exclusion criteria	*X*
Clinical criteria (if applicable)	*NA*
Matching strategy (if applicable)	*NA*
Payment for participation	*X*
IRB, consent/assent obtained	*X*
Number of subjects participated and analyzed	*TBD*
Age	*X*
Sex	*X*
Handedness	*X*
Variables equated across groups?*	NA
Study timeline**	*X*

*For group comparisons, what variables (if any) were equated across groups?

**Study timeline (e.g., number of visits, length of visits, what was measured/collected at each visit)

Recommendations of fMRI details come from Nichols et al., 2016; Poldrack et al., 2008. For extended checklist guidelines for this section following data analysis, see PHBM COBIDAS report Nichols et al., 2016.

Sample Size & Stopping Rule

Target sample size: To obtain our target sample size, we plan to recruit:

```
Because the data were already collected, we did not apply an additional target,
stopping rules, or contingencies in stopping rules.
By using a longitudinal model that allows for missing data, we tried to optimize
the amount of data and subjects retained in our model.
```

Justification of sample size: Provide a justification for target or current sample size and rationale for why you would be powered to test hypotheses.

Because we have three time points of data, we can model linear and quadratic effects of age between subjects, but to avoid over-fitting the model, we will restrict our random effects to linear effects.

By using afni's 3dLME, we are able to retain subjects with 1-3 data points and thereby increasing our n. This allows us to use all subjects and not artificially restrict our sample size to only subjects with complete timepoints.

Specify the type of outcome used as the basis of power computations, e.g. signal in a prespecified ROI, or whole image voxelwise (or clusterwise, peakwise, etc.).

We did not use an *a priori* power test for the study. However, we used Geuter et al., 2018 which indicates a sample size of at least 80 in needed to regions with a medium effect size, as a benchmark to assess if our study was reasonably powered to assess the main aims of our study. We are unaware of power analysis to assess longitudinal changes at the whole brain. Given that our study seeks to assess whole brain changes over time, we anticipate a larger sample is needed and our sample size well exceeds the minimum sample size requirement.

**Effect Size Estimate	Checklist**
Effect size used	NA*
Source of predicted effect size (prior lit, pilot etc.)	NA*
Significant level	NA*
Target power	NA*
Outcome used to calculate	NA*

- *NA = notably, there are many types of fMRI study types that as the time of writing this, we do not have clear standards on how to calculate power. If that is the case, describe reason believe study would be powered or what would benchmark you are using to determine minimum number of subjects/timepoints needed for study design.*

Stopping rule:

If you have a stopping rule for recruitment, add it in here. Contingencies for if your target sample size is not met: (e.g., will hypotheses be adopted to better powered question if you cannot meet you target sample size?)

```
Recruitment happened prior to preregistration and therefore we do
not have a stopping rule and already aware that we have the sample size
needed to conduct our study.
```

Measured Behavioral Variables

Describe each variable that you will measure. You do not need to include any variables that you plan on collecting if they are not going to be included in the confirmatory analyses of this study.

Outcome measures (specific measure, scale/range of measure, which subscale/component of measure you will use):

Outcome measure will largely focus on whole brain longitudinal effects. Analyses will be corrected for multiple comparison using cluster-extent thresholding.

Predictor measures (specific measure, scale/range of measure, which subscale/component of measure you will use):

This paper will focus on developmental effects, which includes age and pubertal development scale. Age and puberty will be centered at min value, centered, and also squared for a quadratic effect. The quadratic effect will also be centered. Sensitivity analyses will also assess these patterns by grade. Grade will also be centered at min grade and squared and centered. Puberty will be measured with the Petersen Development Scale (PDS).

To assist in developmental trend interpretation, select ROIs will be correlated with select behavior. To follow up on analyses from time 1, we will assess how these maturation changes associate with Risk-Taking Behaviors at time 3. Participants completed a modified version of the Adolescent Risk-taking Scale (Alexander et al.,1990). Adolescents reported on their frequency of engaging in 14 risky behaviors on a 4-point scale(0=never,1=once or twice, 2=several times, 3=many times). The scale included questions about rule breaking (e.g.,"I have snuck out of my house without my parents knowing"), sexual activity (e.g.,"I have had sex with someone I just met"), sub-stance use (e.g.,"I have gotten drunk or high at a party"), and dangerous behavior (e.g.,"I did some-thing risky or dangerous on a dare"). A total mean score for all items was calculated (a=.769). Prior to extraction and assessing behavior measures, the preregistration will be updated with the details of analysis plan, measures, and ROIs.

Covariate measures (specific measure, scale/range of measure, which subscale/component of measure you will use):

Based on sex differences in puberty and pubertal timing, reward activation patterns for the MID and SID (Dhingra et al., 2020; Wang et al., 2020), age and pubertal models will be run with and without sex in the model as covariate and interaction term . Furthermore, pubertal timing shown to differ for White, Black and non-White Hispanic youth. Therefore all age and puberty plots will be shown broken down by sex and race.

How was behavioral task performance measured (if task fMRI; e.g., response time, accuracy)?

Although behavioral responses are not the focus of this study, we will assess how performance and response time on the social incentive delay change across age. If there is a significant change in reaction time or accuracy across the three time-points, we will run a sensitivity analysis at the whole brain to assess the main effect of reaction time and for our main results, controlling for reaction time.

Contingency plans for behavioral analysis (e.g., plans if x% of behavioral data is missing; poor variability in behavioral measure). E.g., If the X questionnaire is missing for more than 10% of participants we will not use it or if X does not show variability in response (either ceiling or floor effects) in which we cannot look at behavioral pattern of interest, we will not use that questionnaire and use Y questionnaire instead.

We will run follow up analysis to see if our developmental patterns of change to anticipated and receipt of social feedback associate with behavioral outcomes at time 3. Prior to the decision of which measurement would be used, this preregistration will be updated.

For extended checklist guidelines for this section following data analysis, see PHBM COBIDAS report Nichols et al., 2016.

Additional Operational Definitions

Region Specificity (e.g., defined based on anatomical definition, Prior study cluster, Neurosynth definition (make sure to be specific here!), Parcellation definition)

Details will be added prior to ROI analysis. VS and amygdala will be define anatomically, using the same ROI previously defined in study one. Additional regions will be assessed by clustered-extent thresholds at the whole brain. ROI extraction will be defined by whole brain behavior. ROI extraction will serve two main purposes: 1. Graph the effects found at the whole brain and 2. Correlate with select behavior (outside of

fMRI task) in order to assist in interpretation of results.Prior to ROI extraction for additional analyses, this preregistration will be updated.

Any other definitions used across study: (e.g., how is "risk" defined; how was "depressed")

Maturation/Development will be defined be age (age at scan – DOB) and puberty (pubertal stage via PDS and pubertal timing via age-adjusted pubertal stage score). Puberty will be assessed with caution, taking into consideration outlined in recent papers, including Cheng et al., 2020. Puberty for this study will be assessing "perceived pubertal stage", as it is a self-report of pubertal stage. While it is well documented that puberty can influence brain development, it is also known that sex, race, BMI, stress, or parental education, may impact the onset of puberty. We will include these correlations and interpret results within the context of these correlations. For example, it is possible that earlier maturation for some groups may mean we are limited in interpreting the effects of pubertal timing for everyone. Pubertal models will be compared to model with sex included as an interaction term and as a control variable.

Transformations

If you plan on transforming, centering, recoding the data, or will require a coding scheme for categorical variables, please describe that process.

Contingency plans for transformation: (e.g., transformations that will occur if data are skewed or for model convergence/multicollinearity)

Code, if applicable: for scoring behavioral data.

```
#age age puberty will be centered by minimum age and minimum puberty stage
# age_c <- age - min_age
# puberty_c <- PDS - min_puberty_stage
# #age and puberty stage will be squared, then centered
# age2 <- age*age
# age2_c <- age2 - min_age2
# puberty2 <- PDS*PDS
# puberty2_c <- puberty2 - min_puberty2
#
# #puberty timing will be assess by taking age adjusted residual of puberty
# Pubertal_timing_fit <- lm(PDS ~ age, data = df)
# df$Pubertal_timing <- residuals(Pubertal_timing)
```

Age, puberty, and grade are all highly correlated and therefore will not be included in the same model. Sensitivity analyses for puberty will include the age-adjusted residual effect of puberty. Sensitivity analyses will also assess how the model varies by assessing linear and quadratic effect of grade.

Analysis Data Exclusion

How will you determine which data points or samples (if any) to exclude from your analyses? How will outliers be handled?

If any subjects were/will be scanned but then rejected/could be rejected from analysis after data collection, state reasons for rejection/possible rejections. (e.g., If a participant has X percentage of volumes with motion, participant will be excluded) Contingency plans: (e.g., plans for missing field map, plans for dropout, missing mprage etc.)

Subject exclusion based on scan quality occurred prior to analysis plan for this paper and will use the criteria previously established and discussed in Telzer et al., 2020.

How will you deal with incomplete or missing data (e.g., missing timepoints or missing/incomplete data within or between runs; what percent missing will be included)?

Incomplete data will be included in the analyses. As long as a participant has at least one time point of data, they are able to be included in the model.

Experimental Design

For all tables, you can fill in the table or write paragraph below as you would for paper and use table as checklist of topics covered.

```
Task description is based on Telzer, et al., 2020 at time one.
Details will be added and confirmed across waves.
These decisions were made prior to the preregistration and will not
be edited to ensure results do not bias our analysis plans.
```

SID Task. Participants completed the Social Incentive Dealy (SID) task while under-going fMRI to measure neural responses when anticipating receiving social rewards and avoiding social punishments. The SID is modified from the widely used Monetary Incentive Delay Task (Knutson, Westdorp, Kaiser, & Hommer, 2000), and reliably engages the VS (Cremers, Veer, Spinhoven, Rom-bouts, & Roelofs, 2015; Kohls et al., 2013). For instance, anticipation of both social and monetary rewards recruits the VS (Spreckelmeyer et al., 2009),and the anticipation of avoidable social punishments recruits the VS similarly to VS activation during the anticipation of social reward gain (Kohls et al., 2013).Each trial of the SID began with a cue that signaled whether the potential feedback would be a reward, punishment, or neutral (500 ms). The cue was a different shape for each condition. The cue was followed by a jittered crosshair (between 0.48 and 3.9 s,M=2.0 s), which was followed by the target (a white square; 300 ms), at which point participants were instructed to press a button as quickly as possible. The display of social feedback (1,450 ms) was dependent on the trial type and participants'reaction time. In the reward condition, a hit (i.e., fast enough response) earned the feedback of a happy face (i.e., social reward feedback), and a miss (i.e., too slow response) earned a blurred face (i.e., neutral feedback). During the punishment condition, a hit earned a blurred face (i.e., neutral feedback) and a miss earned an angry face (i.e., social punishment feed-back). Both hits and misses were followed by a blurred face in the neutral condition. After the feed-back, a jittered crosshair (between 0.51 and 4.2 s,M=2.3 s) was presented before the next trial began. Trials were presented in an event-related design, with reward, punishment, and neutral trials randomly ordered. Participants completed two rounds of the task, totaling 116 trials (48 reward, 48 punishment, 20 neutral).To prevent a ceiling or floor performance effect and ensure participants performed roughly at 50% accuracy so that they received relatively equal amount of positive and negative feedback, the time required for a successful hit was adaptive, starting at 0.30s for the first trial and adding or subtracting 0.02 s after a miss or hit, respectively, with an upper bound of 0.50 s and a lower bound of 0.16s. In order to make the task motivationally salient, age-matched adolescent faces posing emotional facial expressions were utilized as rewards and punishments. The faces were photographs of ethnically diverse male and female adolescents (24 faces, 12 female) taken from the National Institute of Mental Health Child Emotional Faces Picture Set(NIMH-ChEFS). Participants were trained on the meaning of each cue and completed 12 practice trials prior to entering the scanner.

Design Specifications

**Design	Checklist**
Design type (task, rest; event-related, block)	____
Conditions & Stimuli (detailed as possible, pictures encouraged)	____
Number of blocks, trials or experimental units per session and/or subject	____
Timing and Duration (length of each trial and interval between trials)	*X*
Length of experiment (length of full scan and each run)	*X*
Was the design optimized for efficiency, and if so, how?	____
Presentation software (name, version, operating system; code if possible)	____

Task Specification

**Task	Checklist**
Instructions to subjects (what were they asked to do?)	*X*
Stimuli (what were they; how many were there; did it repeat across trials?)	*X*
Stimuli presentation & response collection	*X*
Randomization/pseudo-randomized (why/why not done & how)	*X*
Run order (of tasks within scanner)	_____

Recommendations of fMRI details come from Nichols et al., 2016; Poldrack et al., 2008. For extended checklist guidelines for this section following data analysis, see PHBM COBIDAS report Nichols et al., 2016.

Data acquisition

Use table as checklist of topics covered in your paragraph on data acquisition.

```
fMRI data acqusition occured prior to the pregregistration.
A brief description is provided below, but will be detailed prior
to publication to increase transparency and reproducibility.
```

fMRI Data Acquisition. Imaging data were collected using a three Tesla Siemens Prisma MRI scanner. The SID was presented on a computer screen and projected through a mirror. A high-resolution structural T2* -weighted echo-planar imaging (EPI) volume (TR=2,000 ms; TE=25 ms; matrix=92992; field of view (FOV)=230 mm; 37 slices; slice thickness=3 mm ;voxel size 2.592.593 mm3) was acquired coplanar with a T2* -weighted structural matched-band-width (MBW), high-resolution, anatomical scan(TR=5,700 ms; TE=65 ms; matrix=1929192;FOV=230 mm; 38 slices; slice thickness=3 mm).In addition, a T1* magnetization-prepared rapid-acquisition gradient echo(TR=2,400 ms; TE=2.22 ms; matrix=2569256; FOV=256 mm; sagittal plane; slice thickness=0.8 mm; 208 slices) was acquired. The orientation for the EPI and MBW scans was oblique axial to maximize brain coverage and to reduce noise.

Set up & Acquisitions

**Set up & Acquisitions	Checklist**
Participant Preparation	
Mock scanning (Report type of mock scanner and protocol; i.e. duration, types of sim scans, exper)	_____ or NA
Specific accommodations (e.g., pediatric, parent present? Asleep?)	_____ or NA
Experimental personnel (number of planned personnel to interact with subjects)	_____ or NA
MRI System	
Manufacturer, field strength (in Tesla), model name	*X*
MRI acquisition	
Pulse sequence (gradient/spin echo etc.)	_____
Image type (EPI, spiral, 3D etc.)	_____
Essential sequence & imaging parameters	
For All	
Echo time (TE)	*X*
Repetition time (TR)	*X*
For multishot acquisitions, additionally the time per volume	_____ or NA
Flip angle (FA)	_____
Acquisition time (duration of acquisition)	_____
Functional MRI	
Number of volumes	_____
Inversion Recovery Sequences	
Sparse sampling delay (delay in TR) if used	NA
B0 Field Maps	
Inversion time (TI) for inversion recovery sequence	_____ or NA

**Set up & Acquisitions	Checklist**
Echo time difference (dTE). Diffusion MRI	_____ or NA
Number of directions	_____ or NA
Direction optimization, if used and type	_____ or NA
b-values	_____ or NA
Number of b=0 images	_____ or NA
Number of averages (if any)	_____ or NA
Single shell, multishell (specify equal or unequal spacing)	_____ or NA
Single or dualspinecho, gradient mode (serial or parallel)	_____ or NA
If cardiac gating used	_____ or NA
Imaging Parameters	
Field of view	*X*
In plane matrix size, slice thickness and interslice gap, for 2D acquisitions	_____ or NA
Slice orientation (Axial, sagittal, coronal or oblique)	_____ or NA
Angulation: If acquistion not aligned with scanner axes, specified	_____ or NA
angulation to ACPC line (see Slice position procedure)	_____ or NA
3D matrix size, for 3D acquisitions	_____ or NA
Additional sequence & imaging parameters	
Phase encoding	_____ or NA
Parallel imaging method & parameters	_____ or NA
Multiband parameters	_____ or NA
Readout parameters	_____ or NA
Fat suppression (for anatomical, state if used)	_____ or NA
Shimming	_____ or NA
Slice order & timing	_____ or NA
Brain coverage (e.g., whole brain, was cerebellum, brain stem included)	_____ or NA
Scanner-side preprocessing*	_____ or NA
Scan duration (in seconds)	_____ or NA
Other non-standard procedures	_____ or NA
T1 stabilization (discarded "dummy" scans acquired discarded by scanner)	_____ or NA
Diffusion MRI gradient table (Also referred to as the bmatrix, but not to be confused with the 3×3 matrix that describes diffusion weighting for a single diffusion weighted measurement) scanner-side preprocessing: (e.g., Including: Reconstruction matrix size differing from acquisition matrix size; Prospective-motion correction (including details of any optical tracking, and how motion parameters are used); Signal inhomogeneity correction; Distortion-correction.)	_____ or NA

Recommendations of fMRI details come from Nichols et al., 2016; Poldrack et al., 2008. For extended checklist guidelines for this section following data analysis, see PHBM COBIDAS report Nichols et al., 2016.

Preprocessing

Can fill in the table or write paragraph below as you would for paper and use table as checklist of topics covered.

```
Preprocessing occured prior to the preregistration.
A brief description of this preprocesssing plan is detailed below.
Prior to publication, details of this preprocessing pipeline will be
further detailed for replication and transparency.
Notably, no changes will be made to the preprocessing pipeline for
this manuscript to ensure preprocessing decisions were not altered
based on desire outcomes.
```

Preprocessing was conducted using FSL (FMRIB's Software Library, version 6.0; www.fmrib.ox.ac.uk/fsl) and included the following steps: Skull stripping using BET(Smith, 2002); motion correction with MCFLIRT(Jenkinson, Bannister, Brady, & Smith, 2002); spatial smoothing with Gaussian kernel of full width at half maximum 6 mm; high-pass temporal filtering with a filter width of 128s (Gaussian-weighted least-squares straight line fitting, with sigma=64.0 s); grand-mean intensity normalization of the entire 4D data set by a single multiplicative factor; and individual level ICA denoising for motion and physiological noise using MELODIC(version 3.15; Beckmann & Smith, 2004), combined with an automated signal classifier (Tohka et al.,2008; Neyman-Pearson threshold=0.3). For the spatial normalization, the EPI data were registered to the T1 image with a linear transformation, followed by a white-matter boundary-based transfomation (Greve & Fischl, 2009) using FLIRT, linear and nonlinear transformations to standard Montreal Neurological Institute (MNI) 2-mm brain were per-formed using Advanced Neuroimaging Tools(Avants et al., 2011), and then spatial normalization of the EPI images to the MNI.

**Preliminary quality control	Checklist**
Motion monitoring (For functional or diffusion acquisitions, any visual or quantitative checks for severe motion; likewise, for structural images, checks on motion or general image quality.)	_____
Incidental findings (Protocol for review of any incidental findings, and how they are handled in particular with respect to possible exclusion of a subject's data.)	_____

Data preprocessing

_____ For each piece of software used, give the version number (or, if no version number is available, date of last application of updates)

_____ If any subjects required different processing operations or settings in the analysis, those differences should be specified explicitly

**Pre-processing: general	Checklist**
Specify order of preprocessing operations	_____
Describe any data quality control measures	_____
Unwarping of B0 distortions	_____
Slice timing correction	_____
Reference slice and type of interpolation used (e.g., "Slice timing correction to the first slice as performed, using SPM5's	Fourier phase shift interpolation")
Motion correction	*X*
Reference scan, image similarity metric, type of interpolation used, degrees-of-freedom (if not rigid body) and, ideally, optimization method, e.g., "Head motion corrected with FSL's MCFLIRT by maximizing the correlation ratio between each timepoint and the middle volume, using linear interpolation."	_____
Motion susceptibility correction used	_____
Smoothing	*X*
Size and type of smoothing kernel (provide justification for size; e.g., for a group study, "12 mm FHWM Gaussian smoothing applied to ameliorate differences in intersubject localization"; for single subject fMRI "6 mm FWHM Gaussian smoothing used	to reduce noise")

**Intersubject registration	Checklist**
Intersubject registration method used	_____

**Intersubject registration	Checklist**
Illustration of the voxels present in all subjects ("mask image") can be helpful, particularly for restricted fields of view (to illustrate overlap of slices across all subjects). Better still would be an indication of average BOLD sensitivity within each voxel in the mask	_____
Transformation model and optimization	_____
Transformation model (linear/affine, nonlinear), type of any non-linear transformations (polynomial, discrete cosine basis), number of parameters (e.g., 12 parameter affine, $3 \times 2 \times 3$ DCT basis), regularization, image-similarity metric, and interpolation method	_____
Object image information (image used to determine transformation to atlas)	_____
Anatomical MRI? Image properties (see above)	_____
Co-planar with functional acquisition?	_____
Functional acquisition co-registered to anatomical? if so, how?	_____
Segmented gray image?	_____
Functional image (single or mean)	_____
Atlas/target information	_____
Brain image template space, name, modality and resolution (e.g., "FSL's MNI Avg152, T1 $2 \times 2 \times 2$ mm"; "SPM2's MNI gray matter template $2 \times 2 \times 2$ mm")	_____
Coordinate space	_____
(Typically MNI, Talairach, or MNI converted to Talairach	*X*
If MNI converted to Talairach, what method? e.g., Brett's mni2tal?	_____
How were anatomical locations (e.g., gyral anatomy, Brodmann areas) determined? (e.g., paper atlas, Talairach Daemon, manual inspection of individuals' anatomy, etc.)	_____
Smoothing	_____
Size and type of smoothing kernel (provide justification for size; e.g., for a group study, "12 mm FHWM Gaussian smoothing applied to ameliorate differences in intersubject localization"; for single subject fMRI "6 mm FWHM Gaussian smoothing used	to reduce noise")

Recommendations of fMRI details come from Nichols et al., 2016; Poldrack et al., 2008. For extended checklist guidelines for this section following data analysis, see PHBM COBIDAS report Nichols et al., 2016.

Statistical modeling

For all prompts and tables, can fill in the table or write paragraph below as you would for paper and use table as checklist of topics covered.

Planned comparison *If the experiment has multiple conditions, what are the specific planned comparisons, or is an omnibus ANOVA used?*

General issues *For novel methods that are not described in detail in a separate paper, provide explicit description and validation of method. Remember to include package and package version used for each test.*

First level (fx) modeling

```
First level modeling occurred prior to preregistration.
A brief description of the first level models are detailed below.
Prior to publication, the first level models will be further detailed
for replication and transparency. Notably, no changes will be made to
the first level models for this manuscript to ensure fx modeling
decisions were not altered based on desire outcomes.
```

Individual level, fixed-effects analyses were estimated using the general linear model convolved with a canonical hemodynamic response function in SPM12. The task was modeled as event-related with height conditions, including three anticipation conditions (reward, punishment, neutral), two outcome conditions for both reward (hit, miss) and punishment (hit, miss), and one outcome condition for neutral. Anticipation conditions were modeled as the onset of the cue and the duration of the cue and jitter prior to the target, and outcome conditions were modeled at the onset of and the full duration of the feedback. Six motion parameters were modeled as regressors of no interest. Using the parameter estimates from the general linear model (GLM), linear contrast images comparing each of the conditions of interest were calculated for each individual. The primary contrasts of interest for this study were reward anticipation versus neutral anticipation and punishment anticipation versus neutral anticipation, and reward receipt versus neutral and punishment receipt versus neutral.

**First level (fx) modeling	Checklist**
Event related design predictors	
Modeled duration, if other than zero	*X*
Parametric modulation	*NA*
Block Design predictors(Note whether baseline was explicitly modeled)	_____
HRF basis, typically one of:	
Canonical only	_____
Canonical plus temporal derivative	_____
Canonical plus temporal and dispersion derivative. Smooth basis (e.g. SPM "informed" or Fourier basis; FSL's FLOBS)	_____
Finite Impulse Response model	_____
Drift regressors (e.g. DCT basis in SPM, with specified cutoff)	_____
Movement regressors; specify if squares and/or temporal derivative used	_____
Any other nuisance regressors, and whether they were entered as interactions (e.g. with a task effect in 1st level fMRI, or with group effect)	_____
Any orthogonalization of regressors, and set of other regressors used to orthogonalize against	_____
Contrast construction (Exactly what terms are subtracted from what? Define these in terms of task or stimulus conditions (e.g., using abstract names such as AUDSTIM, VISSTIM) instead of underlying psychological concepts	_____
Autocorrelation model type (e.g., AR(1), AR(1) + WN, or arbitrary autocorrelation function), and whether global or local (e.g., for SPM2/SPM5, 'Approximate AR(1) autocorrelation model estimated at omnibus F-significant voxels ($P < 0.001$), used globally over the whole brain'; for FSL, 'Autocorrelation function estimated locally at each voxel, tapered and regularized in space.')	_____

Second level (group) modeling

Group level models will be run using AFNI's 3dLME. This program allows for voxel-level whole brain analysis of linear mixed effects (maximum-likelihood, multi-level model), while allowing for repeated measures and within subject missing data (Chen et al., 2013). This modeling structure also allows us to minimize the number of models we run and therefore constrains our need for multiple comparison correction across the number of models.

Template model structure included here:

```
3dLME -prefix 3dLME_age2 -jobs 4 \
    -model 'Ant*age_c+Ant*age2_c' \
    -qVars "age_c,age2_c" \
    -qVarCenters "0,0" \
    -resid age2_resids \
    -ranEff '~1+age_c' \
```

```
    -SS_type 3 \
    -num_glt 17 \
    -gltLabel 1 'Reward' -gltCode 1 'Ant : 1*rew' \ #confirmatory
    -gltLabel 2 'Punishment' -gltCode 2 'Ant : 1*pun' \ #confirmatory
    -gltLabel 3 'Neutral' -gltCode 3 'Ant : 1*neu' \ #confirmatory
    -gltLabel 4 'Rew_Neu' -gltCode 4 'Ant : 1*rew -1*neu' \ #confirmatory
    -gltLabel 5 'Pun_Neu' -gltCode 5 'Ant : 1*pun -1*neu' \ #confirmatory
    -gltLabel 6 'Rew_Pun' -gltCode 6 'Ant : 1*rew -1*pun' \ #confirmatory
    -gltLabel 7 'Val_neu' -gltCode 7 'Ant : 1*rew & 1*pun -1*neu' \ #confirmatory
    -glfLabel 8 'rew_age' -glfCode 8 'Ant : 1*rew age_c :' \
    -glfLabel 9 'pun_age' -glfCode 9 'Ant : 1*pun age_c :' \
    -glfLabel 10 'neu_age' -glfCode 10 'Ant : 1*neu age_c :' \
    -gltLabel 11 'age.pun-neu' -gltCode  11 'Ant : 1*pun -1*neu age_c : ' \
    -gltLabel 12 'age.rew-neu' -gltCode  12 'Ant : 1*rew -1*neu age_c : ' \
    -gltLabel 13 'age.rew-pun' -gltCode  13 'Ant : 1*rew -1*pun age_c : ' \
    -gltLabel 14 'age.val' -gltCode  14 'Ant : 1*rew & 1*pun age_c : ' \
    -gltLabel 15 'age.val-neu' -gltCode  15 'Ant : 1*rew & 1*pun -1*neu age_c : ' \
    -glfLabel 16 'rew_age2' -glfCode 16 'Ant : 1*rew age2_c :' \
    -glfLabel 17 'pun_age2' -glfCode 17 'Ant : 1*pun age2_c :' \
    -glfLabel 18 'neu_age2' -glfCode 18 'Ant : 1*neu age2_c :' \
    -gltLabel 19 'age2.pun-neu' -gltCode  19 'Ant : 1*pun -1*neu age2_c : ' \
    -gltLabel 20 'age2.rew-neu' -gltCode  20 'Ant : 1*rew -1*neu age2_c : ' \
    -gltLabel 21 'age.rew-pun' -gltCode  21 'Ant : 1*rew -1*pun  age_c : ' \
    -gltLabel 22 'age2.val' -gltCode  22 'Ant : 1*rew & 1*pun age2_c : ' \
    -gltLabel 23 'age2.val-neu' -gltCode  23 'Ant : 1*rew & 1*pun -1*neu age2_c : ' \
    -dataTable \
```

To probe interactions, we will include additional glt tests, such as the following.

```
-gltLabel 1 'ageearly.rew-pun' -gltCode 9 'Ant : 1*rew -1*pun age : 0' \
-gltLabel 2 'agemid.rew-pun' -gltCode 9 'Ant : 1*rew -1*pun age : 1' \
-gltLabel 3 'agemid.rew-pun' -gltCode 9 'Ant : 1*rew -1*pun age : 2' \
```

**Second level (group) modeling	Checklist**
Statistical model and estimation method, inference type (mixed/random effects or fixed), e.g., "Mixed effects inference with one sample t-test on summary statistic" (SPM2/SPM5), e.g., "Mixed effects inference with Bayesian 2-level model with fast approximation to posterior probability of activation." (FSL)	
If fixed effects inference used, justify	_____
If more than 2-levels, describe the levels and assumptions of the model (e.g., are variances assumed equal between groups)	_____
Repeated measures?	
If multiple measurements per subject, list method to account for within subject correlation, exact assumptions made about correlation/variance e.g., SPM: "Within-subject correlation estimated at F-significant voxels (P <0.001), then used globally over whole brain"; or, if variances for each measure are allowed to vary, "Within-subject correlation and relative variance estimated. . . "	_____
For group model with repeated measures, specify:	
How condition effects are modeled (e.g. as factors, or as linear trends)	_____
Whether subject effects are modeled (i.e. as regressors, as opposed to with a covariance structure)	_____
For group effects: clearly state whether or not covariates are split by group (i.e. fit as a groupbycovariate interaction)	

**Second level (group) modeling	Checklist**
Model type (Some suggested terms include:	
"Mass Univariate"	_____
"Multivariate" (e.g. ICA on whole brain data)	_____
"Mass Multivariate" (e.g. MANOVA on diffusion or morphometry tensor data)	_____
"Local Multivariate" (e.g. "searchlight")	_____
"Multivariate, intrasubject predictive" (e.g. classify individual trials in eventrelated fMRI)	_____
"Multivariate intersubject predictive" (e.g. classify subjects as patient vs. control)	_____
"Representational Similarity Analysis")	_____
Model settings (The essential details of the model For mass univariate, first level fMRI, these include:	
Drift model, if not already specified as a dependent variable (e.g. locally linear detrending of data & regressors, as in FSL)	_____
Autocorrelation model (e.g. global approximate AR(1) in SPM; locally regularized autocorrelation function in FSL)	_____
For mass univariate second level fMRI these include:	
Fixed effects (all subjects' data in one model)	_____
Random or mixedeffects model, implemented with:	
Ordinary least squares (OLS, aka unweighted summary statistics approach; SPM default, FSL FEAT's "Simple OLS")	_____
weighted least squares (i.e. FSL FEAT's "FLAME 1"), using voxelwise estimate of between subject variance	_____
Global weighted least squares (i.e. SPM's MFX)	_____
With any group (multisubject) model, indicate any specific variance structure, e.g.	
Unequal variance between groups (and if globally pooled, as in SPM)	_____
If repeated measures, the specific covariance structure assumed (e.g. compound symmetric, or arbitrary; if globally pooled)	_____

ROI analysis

`Details of ROI analysis we be added prior to analysis.

**ROI analysis	Checklist**
How were ROIs defined (e.g., functional, anatomical, parcel localizer)?	_____
How was signal extracted within ROI? (e.g., average parameter estimates, FIR deconvolution?)	_____
If percent signal change reported, how was scaling factor determined (e.g., height of block regressor or height of isolated event regressor)?	_____
Is change relative to voxel-mean, or whole-brain mean?	_____
Justify definition of ROI and analysis conducted with it: (e.g., if your ROI is defined based on the cluster; how will you ensure your ROI analyses are not circular?)	_____

If not previously specified above, what statistical model will you use to test each hypothesis? Please include the type of model (e.g. ANOVA, multiple regression, SEM, etc) and the specification of the model (this includes each variable that will be included as predictors, outcomes, or covariates). Please specify any interactions that will be tested and remember that any test not included here must be noted as an exploratory test in your final article.

Recommendations of fMRI details come from Nichols et al., 2016; Poldrack et al., 2008. For extended checklist guidelines for this section following data analysis, see PHBM COBIDAS report Nichols et al., 2016.

Statistical inference

For all prompts and tables, can fill in the table or write paragraph below as you would for paper and use table as checklist of topics covered.

Statistical inference will use whole-brain cluster extent thresholding, which controls for family wise errors. Because this approach can penalize smaller regions, results will also be displayed with subthreshold results and group-level statistical maps will be uploaded to neurovault.
Unless otherwise specified, reported results will exceed the minimum cluster extent threshold needed for a 0.05 family-wise error (FWE) rate given a voxel-wise threshold of p = 0.001 for each contrast, as determined by AFNI 3dClustSim, version AFNI_16.1.06 (Mar 6 2016). This version of AFNI was updated to resolve identified software bugs (Eklund et al., 2016). Smoothness estimates entered into 3dClustSim were an average of subject level spatial autocorrelation function (acf) parameters based on individual participants' residuals from their fx level model, as calculated by 3dFWHMx using the -acf flag. Results will also be run and reported for correction of FWE set to .025, to account for two primary models (age and puberty).

ROI analysis statistical inference will be detailed prior to analysis.

**Statistical inference	Checklist**
Search region (Type of search region for analysis, and the volume in voxels or CC)	*X*
If not whole brain, state how region was determined; method for constructing region should be independent of present statistic image	*TBD*
Whole brain or "small volume"; carefully describe any small volume correction used for each contrast	*X*
If a small-volume correction mask is defined anatomically, provide named anatomical regions from a publicly available ROI atlas	*TBD*
If small-volume correction mask is functionally defined, clearly describe the functional task and identify any risk of circularity	*TBD*
All small-volume corrections should be fully described in the methods section, not just mentioned in passing in the results	*TBD*
Statistical type (Typically one of:	
Voxelwise (aka peakwise in SPM)	_____ or NA
Cluster-wise	_____ or NA
Cluster size	_____ or NA
Cluster mass	_____ or NA
Thresholdfree Cluster Enhancement (TFCE)	_____ or NA
For cluster size or mass, report:	
Cluster-forming threshold	_____ or NA
For all clusterwise methods, report:	
Neighborhood size used to form clusters (e.g. 6, 18 or 26)	_____ or NA
For TFCE, report:	
Use of nondefault TFCE parameters.)	_____ or NA
P value computation (Report if anything but standard parametric inference used to obtain (uncorrected) Pvalues. If nonparametric method was used, report method (e.g. permutation or bootstrap) and number of permutations/samples used.)	
Multiple test correction (For massunivariate, specify the type of correction and how it is obtained, especially if not the typical usage.)	_____ or NA
Usually one of:	
Familywise Error	
Random Field Theory (typical)	*X*

**Statistical inference	Checklist**
Permutation	NA
Monte Carlo	NA
Bonferroni	NA
False Discovery Rate	
Benjamini & Hochberg FDR (typical)	NA
Positive FDR	NA
Local FDR	NA
Clusterlevel FDR	NA
None/Uncorrected	NA
If permutation or Monte Carlo, report the number of permutations/samples. If Monte Carlo, note the brain mask and smoothness used, and how smoothness was estimated	NA
If correction is limited to a small volume, the method for selecting the region should be stated explicitly	NA
If no formal multiple comparisons method is used, the inference must be explicitly labeled "uncorrected."	*X*
If FWE found by random field theory list the smoothness in mm FWHM and the RESEL count	*TBD*
If FWE found by simulation (e.g., AFNI AlphaSim), provide details of parameters for simulation	NA
If not a standard method, specify the method for finding significance (e.g., "Custom in-lab software was used to construct statistic maps and thresholded at FDR< 0.05 (Benjamini and Hochberg, 1995)"	NA
If cluster-wise significance, state cluster-defining threshold (e.g., P = 0.001)	*X*
False negative discussion	
Any discussion of failure to reject the null hypothesis (e.g., lack of activation in a particular region) should be accompanied by SNR or effect size of the actually observed effect (allows reader to infer power to estimate an effect)	*X*

Whole brain cluster-extent correction is subject to false negatives, due to loss of power. We will therefore provide are statistical maps to neurovault and provide supplemental materials that show the thresholded, as well as sub-threshold effects. The lack of an effect will not be interpreted as no change, but rather detectable change at the whole brain.

Functional Connectivity

```
Functional connectivity will be assessed *a priori* with anatomically defined VS
as the seed at each time point. Details will be added prior to analysis.
Secondary functional connectivity analysis will be considered if there are
linear or quadratic effects in a lateral or medial prefrontal cortex region.
```

**Functional Connectivity	Checklist**
Confound adjustment & filtering Report:	
Method for detecting movement artifacts, movement-related variation, and remediation (e.g. 'scrubbing', 'despiking', etc)	_____ or NA
Use of global signal regression, exact type of global signal used and how it was computed	_____ or NA
Whether a high or lowpass temporal filtering is applied to data, and at which point in the analysis pipeline. Note, any temporal regression model using filtered data should have its regressors likewise filtered	_____ or NA
Multivariate method: Independent Component Analysis Report:	

**Functional Connectivity	Checklist**
Algorithm to estimate components	_____ or NA
Number of components (if fixed), or algorithm for estimating number of components	_____ or NA
If used, method to synthesize multiple runs	_____ or NA
Sorting method of IC's, if any	_____ or NA
Detailed description of how components were chosen for further analysis	_____ or NA
Dependent variable definition	
For seed¬-based analyses report:	
Definition of the seed region(s)	_____ or NA
Rationale for choosing these regions	_____ or NA
For region-¬based analyses report:	
Number of ROIs	_____ or NA
How the ROI's are defined (e.g. citable anatomical atlas; auxiliary fMRI experiments); note if ROIs overlap	_____ or NA
Assignment of signals to regions (i.e. how a time series is obtained from each region, e.g. averaging or first singular vector)	_____ or NA
Note if considering only bilateral (L+R) merged regions	_____ or NA
Note if considering only interhemispheric homotopic connectivity	_____ or NA
Functional connectivity measure/model Report:	
Measure of dependence used, e.g. Pearson's (full) correlation, partial correlation, mutual information, etc; also specify:	
Use of Fisher's Z-transform (Yes/No) and, if standardised, effective N is used to compute standard error (to account for any filtering operations on the data)	_____ or NA
Estimator used for partial correlation	_____ or NA
Estimator used for mutual information	_____ or NA
Regression model used to remove confounding effects (Pearson or partial correlation)	_____ or NA
Effectivity connectivity Report:	
Model	_____ or NA
Algorithm used to fit model	_____ or NA
If per subject model, method used to generalize inferences to population.	_____ or NA
Itemize models considered, and method used for model comparison	
Graph analysis	
Report the 'dependent variable' and 'functional connectivity measure' used (see above). Specify either:	
Weighted graph analysis or	_____ or NA
Binarized graph analysis is used, clarifying the method used for thresholding (e.g. a 10% density threshold, or a statistically ¬defined threshold); consider the sensitivity of your findings to the particular choice of threshold used	_____ or NA
Itemise the graph summaries used (e.g. clustering coefficient, efficiency, etc), whether these are global or per¬node/per¬edge summaries. In particular with fMRI or EEG, clarify if measures applied to individual subject networks or group networks	_____ or NA

Recommendations of fMRI details come from Nichols et al., 2016; Poldrack et al., 2008. For extended checklist guidelines for this section following data analysis, see PHBM COBIDAS report Nichols et al., 2016.

Follow-up Analyses

If not specified previously, will you be conducting any confirmatory analyses to follow up on effects in your statistical model, such as subgroup analyses, pairwise or complex contrasts, or follow-up tests from interactions? Remember that any analyses not specified in this research plan must be noted as exploratory.

Follow-up analysis to better understand our results will include a) association with behavior, to be detailed prior to analysis b) connectivity, to be detailed prior to analysis, c) ROI extraction to visually graph and perform model comparison outside of the whole-brain model, to be detailed prior to analysis.

Exploratory Analyses

If you plan to explore your data set to look for unexpected differences or relationships, you may describe those tests here. An exploratory test is any test where a prediction is not made up front, or there are multiple possible tests that you are going to use. A statistically significant finding in an exploratory test is a great way to form a new confirmatory hypothesis, which could be registered at a later time.

By nature of a whole-brain analysis, our results are exploratory. They will be assessed controlling for multiple comparison, though interpretation will try to emphasize the discussion of results in terms of effects sizes, when possible.