Archival Report

Biological Psychiatry

Sex-Specific Vulnerability to Externalizing Problems: Sensitivity to Early Stress and Nucleus Accumbens Activation Over Adolescence

Lauren R. Borchers, Justin P. Yuan, Josiah K. Leong, Booil Jo, Rajpreet Chahal, Joshua Ryu, Andrew Nam, Saché M. Coury, and Ian H. Gotlib

ABSTRACT

BACKGROUND: Exposure and sensitivity to early-life stress (ELS) are related to increased risk for psychopathology in adolescence. While cross-sectional studies have reported blunted nucleus accumbens (NAcc) activation in the context of these associations, researchers have not yet assessed the effects of ELS on developmental trajectories of activation. We examined whether trajectories are affected by stress and the moderating role of biological sex in predicting vulnerability to symptoms of psychopathology.

METHODS: Adolescents (n = 173) completed 3 assessments at 2-year intervals across puberty (ages 9–18 years). At baseline, we assessed objective ELS and stress sensitivity using the Traumatic Events Screening Inventory for Children. At all time points, we assessed NAcc activation using the Monetary Incentive Delay task and externalizing, internalizing, and total problems using the Youth Self-Report. We examined correlations between NAcc trajectories (extracted using linear mixed-effects models) with ELS and stress sensitivity and conducted multivariate regression analysis to examine the interaction of NAcc trajectories and biological sex in predicting symptoms of psychopathology.

RESULTS: Symptoms increased over adolescence. Stress sensitivity, but not objective ELS, was associated with decreasing trajectories of NAcc activation. Biological sex interacted with NAcc trajectories to predict psychopathology; boys, but not girls, with decreasing NAcc activation had more severe externalizing problems in adolescence. These findings were replicated in the putamen and caudate but not in the medial prefrontal cortex or control brain regions.

CONCLUSIONS: NAcc activation may be a sex-specific marker of externalizing problems in adolescence. Efforts to reduce stress sensitivity may help to decrease symptoms of psychopathology in adolescent boys.

https://doi.org/10.1016/j.biopsych.2024.01.011

Early-life stress (ELS) is alarmingly prevalent (1,2). Nearly 40% of children in the United States have a Child Protective Services investigation prior to age 18 (3), and over 13% of children have confirmed maltreatment (4). These cases reflect more severe forms of ELS such as abuse or neglect; far more children experience moderate levels of ELS, such as being bullied, experiencing financial strain, or moving. Importantly, however, not all children respond similarly to stressors in their environment.

Objectively assessed ELS affects a range of domains including physical health (2) and cognitive performance (5) and is associated with various forms of psychopathology (6,7). Stress sensitivity, the tendency to have an outsized subjective response to stressors, has also been linked to psychopathology (8,9). Importantly, girls tend to be more reactive to stressors in their environment than boys (10). Furthermore, whereas girls are more likely to develop internalizing problems such as depression and anxiety, boys are more likely to experience externalizing problems such as disruptive behavioral disorders (11).

ELS has also been found to alter reward processing during adolescence. The brain region that is associated most strongly with reward processing is the nucleus accumbens (NAcc) (12). Perhaps not surprisingly, therefore, researchers have documented associations between ELS and alterations in NAcc activation during adolescence. Numerous cross-sectional studies have reported anomalous NAcc activation during reward tasks in participants who have experienced ELS (13). For example, Mehta *et al.* (14) found that previously institutionalized adolescents had blunted NAcc activation during the anticipation of reward versus the anticipation of nonreward. Furthermore, activation in the NAcc was blunted in adolescents, but not in children, who had experienced more severe ELS (15). Adults who experienced more severe ELS have also been found to have decreased activation in the NAcc (16,17). Notably, additional factors such as prenatal exposure to nicotine (18) are also associated with blunted NAcc activation, and the summarized literature is not experimental. Therefore, while these studies provide a contextual framework for how ELS may affect the brain and symptoms of psychopathology, they do not imply causality. Although Ho *et al.* (9) found that greater stress sensitivity was related to poorer white matter coherence in the right frontal uncinate fasciculus in

ISSN: 0006-3223

adolescents, no study has yet examined the relationship between stress sensitivity and brain function.

Several investigators have linked alterations in reward circuitry to psychopathology in adolescence. For example, Goff *et al.* (15) reported that decreased NAcc activation was associated with more severe symptoms of depression. Similarly, Hanson *et al.* (19) reported that the relationship between emotional neglect and depressive symptoms was partially mediated by decreased activation in the ventral striatum in adolescents. There have also been equivocal findings in this area. Whereas Hawes *et al.* (20) found that children with disruptive behavioral disorders had blunted NAcc activation, Bjork *et al.* (21) found that adolescents with externalizing disorders had greater NAcc activation than healthy control participants. Yau *et al.* (22) reported that in children with a parent with alcoholism, but not in control participants, NAcc activation was positively related to externalizing problems. Finally, using data from the Adolescent Brain Cognitive Development (ABCD) Study, Schettini *et al.* (23) found that boys with smaller right NAcc volumes had more externalizing problems.

It is important to note that almost all these studies are cross-sectional. Few researchers have conducted repeated assessments of neural activation, and even fewer have described trajectories of NAcc activation in adolescents (see Table S1 for a list of the design and contrasts for each study). Thus, it is possible that alterations in reward circuit activation precede the onset of psychopathology. For example, high-risk youths have been found to have decreased activation in the reward response compared with control participants (24,25), although researchers have not yet elucidated the timing and developmental course of these effects. The current study was designed, in part, to address this issue.

Adolescence is characterized by protracted brain development in the striatum and by an increase in symptoms of psychopathology, both of which have been linked to exposure to ELS. We hypothesized that participants who experienced more severe objective ELS and exhibited greater stress sensitivity would have attenuated NAcc activation across adolescence, which in turn would increase their risk for symptoms of psychopathology. We also examined whether biological sex and NAcc activation interact to predict subsequent internalizing and externalizing problems. Importantly, other brain regions have been implicated in ELS and in psychopathology. Therefore, we also analyzed activation in the putamen, caudate, and medial prefrontal cortex (mPFC) in addition to control regions.

METHODS AND MATERIALS

Participants and Procedure

A total of 225 adolescents from the San Francisco Bay Area participated in a longitudinal study assessing the effects of ELS on psychobiological development. At time 1, participants were between ages 9 and 13 years, and boys and girls were matched on pubertal status (therefore, boys were older than girls on average at all assessments) (Table 1). Participants returned for a second (time 2) and third (time 3) follow-up assessment at approximately 2-year intervals (Figure 1). At all 3 time points, parents or legal guardians provided written consent and children provided verbal assent to participate in the study. The Stanford Institutional Review Board approved all study procedures.

The current study included participants who completed the ELS interview, underwent functional magnetic resonance imaging (fMRI), and self-reported symptoms of psychopathology. For more information regarding exclusion criterion for the study and fMRI data, see Recruitment and Exclusion Criterion in the Supplement. Our core analysis strategy—longitudinal mixed-effects modeling using maximum likelihood estimation—allowed us to fully utilize all available data (173 participants who completed at least 1 time point).

Table 1. Characteristics of the Sample by Biological Sex

Variable	Boys, n = 70	Girls, n = 103	Statistic, p Value
Race			
Asian/Asian American	7 (10.0%)	14 (13.6%)	$\chi^2_5 = 3.40$, $p = .638$
Biracial	17 (24.3%)	18 (17.5%)	
Black/African American	6 (8.6%)	5 (4.9%)	
Hispanic/Latin-X	4 (5.7%)	10 (9.7%)	
Other Race	5 (7.1%)	6 (5.8%)	
White	31 (44.3%)	50 (48.5%)	
Pubertal Status Time 1	1.94 (0.70)	2.08 (0.80)	$t_{171} = -1.14$, $p = .256$
Income-to-Needs Ratio Time 1	1.30 (0.53)	1.26 (0.57)	$t_{149} = 0.52$, $p = .602$
Objective Early-Life Stress	6.98 (4.65)	6.26 (5.08)	$t_{171} = 0.94$, $p = .348$
Stress Sensitivity	−0.17 (0.62)	0.04 (0.51)	$t_{171} = -2.41$, $p = .017$
Age at Time 1, Years	11.88 (0.93)	11.18 (1.06)	$t_{139} = 3.82$, $p < .001$
Age at Time 2, Years	13.76 (1.04)	13.10 (1.19)	$t_{115} = 2.91$, $p = .004$
Age at Time 3, Years	15.81 (0.99)	15.32 (1.17)	$t_{100} = 2.59$, $p = .011$
Total Problems at Time 1	38.51 (21.30)	37.29 (22.89)	$t_{171} = 0.35$, $p = .723$
Internalizing at Time 3	10.79 (7.32)	17.32 (9.51)	$t_{132} = -4.32$, $p < .001$
Externalizing at Time 3	11.14 (6.95)	11.57 (8.07)	$t_{132} = -0.32$, $p = .747$

Values are presented as n (%) or mean (SD).

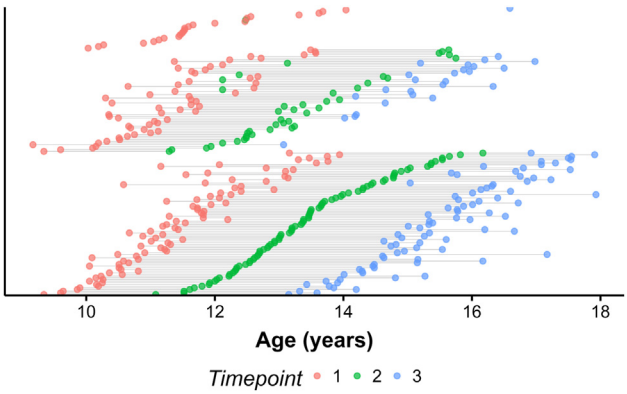


Figure 1. Distribution of age across time points. Each row is a different participant. Red dots are time 1, green dots are time 2, blue dots are time 3. The top section shows participants with usable scan data at 1 time point followed by those who had 2 and 3 usable scans, respectively.

For descriptions of the assessment of income-to-needs ratio and pubertal staging, see Income-to-Needs Ratio and Pubertal Assessment in the Supplement.

ELS Assessment

We administered the Traumatic Events Screening Inventory for Children (26) at baseline, obtaining information about 30 different stressors in childhood in addition to an open-ended question. If a stressor was endorsed, participants reported their subjective severity rating indicating how scared they felt at the time of the stressor. Coders who were blinded to the interview determined the objective severity of each event using a modified version of the UCLA Life Stress Interview coding system (27). Agreement among coders was high (intraclass correlation coefficient [ICC] = 0.99) (28). The severity of ELS was calculated based on the sum of the maximum score of each type of stressor reported so as to not give too much weight to the scores of participants who reported many events. The prevalence of objective ELS in our sample is presented in Table S2.

Stress sensitivity was calculated by residualizing a child's cumulative subjective stress severity score from the panel's cumulative objective stress severity score (i.e., variance in subjective ratings after regressing out objective ratings). Negative values indicate low stress sensitivity or a reduced stress response; positive values indicate high stress sensitivity or a heightened stress response (9).

Youth Self-Report

Participants completed the Youth Self-Report (YSR) (29). This 112-item questionnaire indexes the severity of problems experienced over the last 6 months measured on a 3-point scale (0, not true; 2, very true or often true). These items yield total raw scores for internalizing, externalizing, and total problems (see Figure S1 for the distribution of scores across time points). Internalizing problems are calculated by summing the anxious/depressed, somatic complaints, and withdrawn subscales. Externalizing problems are calculated by summing the aggressive behavior and rule-breaking behavior subscales.

MRI Scan Acquisition

Scanning sequences were acquired using a 3T MRI scanner (Discovery MR750 scanner; GE Healthcare) equipped with a 32-channel head coil (Nova Medical). See MRI Scan Acquisition in the Supplement for information about the acquisition parameters and see Scanner Upgrade/COVID-19 Status in the Supplement for information about controlling for COVID-19 data collection status and the scanner upgrade, hereby referred to as collection status.

Monetary Incentive Delay Task

Children completed a child-friendly version of the Monetary Incentive Delay (KIDMID) task at each time point. The rewarding stimulus (typically monetary) was replaced with points redeemed for prizes (Figure S2) (30). Monetary tasks have reliably been found to recruit the NAcc during reward in children, adolescents, and adults (30,31). The KIDMID task consists of 72 pseudorandomized 6-second trials (locked to 3-volume acquisitions). Participants viewed an incentive cue, anticipated the incentive, responded to the target, and viewed the trial outcome (±5; ±0). For information regarding the procedure, missed trials, reaction time, and performance, see KIDMID Procedure and Performance on the KIDMID task in the Supplement (Table S3).

We examined the bilateral NAcc to probe reward circuitry. We used an 8-mm sphere; the center coordinates are based on Wu *et al.* (32). We converted the center coordinates from Talairach space to Montreal Neurological Institute space using the icm2mni function (https://www.brainmap.org/icbm2tal/). The NAcc coordinates (x = ±12, y = 12, z = −7) are from a meta-analysis of 27 studies of reward anticipation (31). We assessed activation in 3 additional reward regions spanning the striatum and prefrontal cortex (putamen, caudate, and mPFC) and 6 control regions spanning parietal, default mode, and visual areas (angular gyrus, cingulate gyrus [posterior division], occipital fusiform gyrus, occipital pole, and inferior frontal gyrus opercularis and triangularis) to examine the specificity of our findings to NAcc activation (see Reward-Related ROIs and Control Regions in the Supplement) (Figures S6 and S7).

Preprocessing

We preprocessed the functional images using fMRIprep version 20.2.1 (33). fMRIPrep aligns anatomical and functional images, resamples functional data into the desired template space, applies slice-timing correction, identifies motion outliers, and derives regressors such as white matter and cerebrospinal fluid. Details regarding the preprocessing are discussed in Preprocessing in the Supplement. We conducted additional standard preprocessing steps using AFNI version 18.2.04 (34), presented in AFNI Preprocessing in the Supplement. Then, we extracted percent signal change for the NAcc by calculating the mean signal across the whole task and then subtracting the mean from the activation at each volume acquisition (or repetition time) and dividing by the mean signal to derive a continuous measure of percent signal change.

Statistical Analyses

Sample Characteristics. All analyses were conducted using R version 4.0.2. First, we characterized the sample in terms of demographic and clinical variables, examining biological sex differences in our measures of interest.

Time Course and Whole-Brain Analysis. We analyzed the time course data across hit and missed trials (Figure S3) and conducted whole-brain analysis at each time point (see Whole-Brain Analysis in the Supplement and Figure S4) to ensure that the KIDMID task recruited the NAcc.

Stability of NAcc Activation. We tested the stability (reliability) of neural activation in the NAcc over the 4-year period in both conditions and contrasts using ICC (calculated based on 2-way mixed-model; 3,*k*) (35).

We conducted a series of linear mixed-effects models and extracted each individual's estimated intercept and slope (referred to as the trajectory) of NAcc activation and of internalizing and externalizing problems across adolescence. We winsorized estimated neural trajectories to minimize the impact of extreme cases (see Sensitivity Analysis: Non-Winsorized Values in the Supplement). All analyses controlled for collection status, and we applied the Bonferroni correction to account for multiple comparisons (adjusted p value = .017). The models used for the primary analyses are presented in Equations in the Supplement.

Primary Analysis. First, we examined whether greater objective ELS and stress sensitivity predicted decreasing NAcc trajectories. Next, we assessed the main effect and interaction between NAcc trajectories and biological sex in predicting symptoms of psychopathology. Finally, we examined whether greater objective ELS and stress sensitivity predicted more severe internalizing and externalizing problems in adolescence. In the presence of a significant main effect or interaction, we conducted post hoc analyses to determine whether findings were specific to internalizing or externalizing problems. In the presence of a significant interaction involving biological sex, we conducted a follow-up simple slope analysis to examine whether the findings were significant in boys and/or girls. A schematic with our primary statistical framing and findings is presented in the Supplement (Figure S5).

Supplemental Analysis. To ensure that our findings were insensitive to modeling decisions, we conducted sensitivity analyses 1) examining raw psychopathology and T scores at time 3 (see Sensitivity Analysis: Time 3 YSR Data in the Supplement), 2) using nonwinsorized NAcc values (see Sensitivity Analysis: Non-Winsorized Values in the Supplement), and 3) examining only the 63 participants with usable data at all 3 time points (see Sensitivity Analysis: Missing at Random in the Supplement). In exploratory analyses, we examined whether objective ELS and stress sensitivity were related to the NAcc intercept and whether NAcc trajectories interacted with biological sex to predict trajectories of internalizing or externalizing problems.

Finally, to examine the specificity of activation to the NAcc, we examined activations in 3 additional reward-related regions (putamen, caudate, mPFC) and 6 control regions (angular gyrus, cingulate gyrus [posterior division], occipital fusiform gyrus, the occipital pole, and inferior frontal gyrus opercularis and triangularis).

RESULTS

Sample Characteristics

Demographic characteristics and sample means are presented in Table 1. The sample is representative of the broader San Francisco Bay Area population (i.e., racially and ethnically diverse, high income). Information about handedness, distributions of the YSR (Figure S1), and relations of objective ELS and stress sensitivity with demographic variables is presented in Demographics and Stress Variables in the Supplement.

Time Course and Whole-Brain Analysis

Across all trials, time course analyses revealed that there was greater activation in the NAcc during the anticipation of reward than during the anticipation of nonreward (time 1: $t_{145} = 5.64$, $p < .001$; time 2: $t_{122} = 10.01$, $p < .001$; time 3: $t_{106} = 8.32$, $p < .001$) (Figure S3). Across hit trials, there was greater activation in the NAcc during the receipt of reward than the receipt of nonreward (time 1: $t_{145} = 3.73$, $p < .001$; time 2: $t_{122} = 3.12$, $p = .002$; time 3: $t_{106} = 2.73$, $p = .007$) (Figure S3). The whole-brain analysis displaying activation in the NAcc during the KIDMID task at each time point is presented in Figure S4.

Stability of NAcc Activation

Across all trials, the anticipation of reward versus anticipation of nonreward contrast had an ICC value of 0.32 in the NAcc. This value is weak to modest but suggests shared variance in signal over time (see Discussion for a more nuanced review of relevant issues). However, there was very poor reliability of the receipt of reward versus receipt of nonreward contrast (ICC = 0.08) across all trials (and 0.04 across hit/correct trials). These values for anticipation are consistent with ICC values that have been reported in other longitudinal fMRI studies (36,37) and are higher than those in the ABCD dataset, which has an explicit objective of modeling longitudinal brain activation (38). Given our interest in the longitudinal development of reward circuitry (focus on shared variance across development), we examined the anticipation of reward versus anticipation of nonreward contrast in all subsequent analyses.

Development of Psychopathology and NAcc Activation

When controlling for collection status and biological sex, internalizing ($\beta = 0.11$, $p = .025$) and externalizing ($\beta = 0.12$, $p = .014$) problems increased across adolescence. Furthermore, there was a main effect of biological sex on internalizing problems such that girls had more severe problems than boys ($\beta = 0.30$, $p = .014$). Internalizing and externalizing problems in adolescence were significantly correlated in both boys ($r = 0.62$, $p < .001$) and girls ($r = 0.68$, $p < .001$). NAcc activation during the anticipation of reward versus the anticipation of nonreward did not increase significantly across adolescence ($\beta = 0.11$, $p = .054$).

Stress, NAcc Trajectories, Biological Sex, and Psychopathology

Greater stress sensitivity (β = −0.15, p = .043), but not objective ELS (β = 0.09, p = .258), predicted decreasing NAcc trajectories. There was a significant interaction of biological sex and NAcc trajectories in predicting psychopathology (β = 0.10, $p < .001$) even after controlling for baseline total problems. The interaction was significant for externalizing (β = 0.43, p = .002), but not for internalizing (β = −0.08, p = .528), problems. Follow-up simple slope analysis indicated that boys (β = −0.28, $p < .001$), but not girls (β = 0.12, p = .210), with decreasing NAcc trajectories had greater externalizing problems in adolescence (Figure 2). Finally, objective ELS (β = 0.06, p = .008) and stress sensitivity (β = 0.05, p = .019) predicted more severe symptoms of psychopathology. Specifically, objective ELS predicted internalizing (β = 0.23, p = .002) and externalizing (β = 0.18, p = .016) problems, whereas stress sensitivity only predicted internalizing problems (β = 0.18, p = .014; externalizing: β = 0.04, p = .634). Consistent with the MacArthur framework, we did not test for mediation given that there was no significant association between stress sensitivity and externalizing problems in adolescence (39).

Supplemental Analysis

We obtained a similar pattern of findings when using raw and T scores from time 3 (see Sensitivity Analysis: Time 3 YSR Data in the Supplement) and nonwinsorized values (see Sensitivity Analysis: Non-Winsorized Values in the Supplement). We also reran our primary analyses including only participants with usable data across all 3 time points (who did not differ on demographic, clinical, or neural measures at baseline) (see Attrition in the Supplement) and obtained a similar pattern of results (see Sensitivity Analysis: Missing at Random in the Supplement), suggesting that our results were insensitive to the missing data assumption: missing at random (40).

Exploratory analyses examining the NAcc intercept and trajectories of psychopathology are presented in Exploratory Analyses in the Supplement.

We obtained a similar pattern of findings in the putamen and caudate as we observed in the NAcc. Specifically, trajectories of activation in these regions were related to stress sensitivity, and decreasing activation predicted more externalizing problems in boys (see Reward-Related ROIs in the Supplement). Trajectories of activation across the NAcc, putamen, caudate, and mPFC during the anticipation of reward versus anticipation of nonreward were correlated (rs = 0.17–0.87) except for the mPFC and caudate (r = 0.12, p = .121). Importantly, we did not observe a similar pattern of findings in the mPFC or control brain regions (see Reward-Related ROIs and Control Regions in the Supplement), underscoring the specificity of our findings concerning stress sensitivity, biological sex, and externalizing problems to activation in the striatum.

DISCUSSION

ELS has been found to be related to alterations in reward circuitry and an increased risk for developing psychopathology. Here, we built on previous research by conducting and analyzing data from repeated assessments over adolescence and by examining both objectively assessed ELS and participants' sensitivity to stress. In a series of longitudinal analyses, we found that internalizing and externalizing problems increased across adolescence. Furthermore, we found that stress sensitivity, but not objectively assessed ELS, was associated with a higher NAcc intercept and decreasing trajectories of NAcc activation over adolescence. Finally, adolescent boys, but not girls, with decreasing NAcc activation exhibited more severe externalizing problems in adolescence. We found the same pattern of results in the putamen and caudate, highlighting the role of the striatum in stress sensitivity and risk for developing externalizing problems in adolescent boys.

Researchers have posited that ELS is characterized by blunted NAcc activation in adolescence (14,15); however, researchers have not examined longitudinal patterns of NAcc activation and the effects of stress sensitivity. Although our findings require replication, they provide important insights

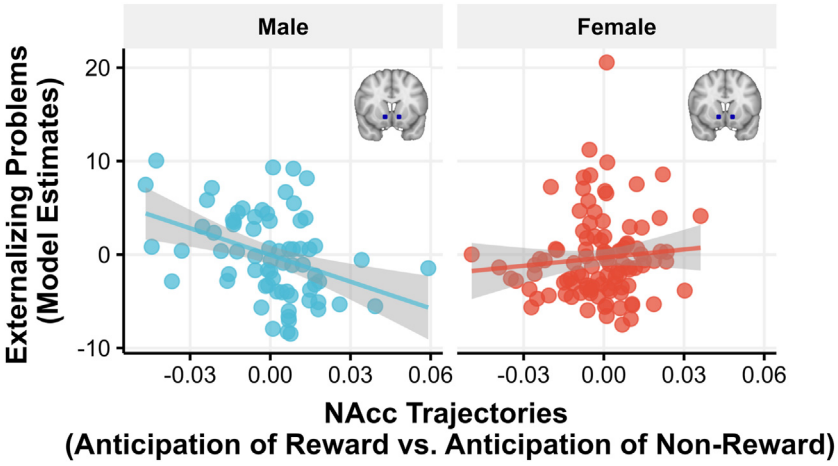


Figure 2. Decreasing nucleus accumbens (NAcc) trajectories predict externalizing problems in boys.

concerning NAcc activation and the development of psychopathology across development.

Investigators have posited that rates of maturation in adolescents are different in the striatum than they are in regions like the prefrontal cortex (41), possibly contributing to higher sensation seeking and impulsivity, and to diminished top-down control, in adolescence (42). In fact, whereas the prefrontal cortex is characterized by protracted brain development into adulthood (43), during adolescence there are increases in dopamine (41) and high levels of NAcc activation (44). Although speculative, the observed positive association between stress sensitivity and NAcc intercepts at age 9 may reflect stress acceleration, which makes it more difficult for children to control externalizing-related problems.

Our findings complement research showing that adolescents and adults with attention-deficit/hyperactivity disorder have blunted task-related NAcc activation (45–47). Our findings also extend findings reported by Hawes *et al.* (20) that adolescents with a sole diagnosis of disruptive behavioral disorders have more blunted NAcc activation than adolescents who also have callous-unemotional traits. We found that decreasing NAcc activation over adolescence, which may reflect insensitivity to reward, predicted the development of externalizing problems in boys. Relatedly, adolescents who were exposed to higher stress were less sensitive to reward in a decision-making task (48). If boys experience decreasing sensitivity to reward over adolescence, it may contribute to the increased impulsivity and sensation seeking that have been documented in adolescents with more severe externalizing problems (49) (see Hyperactivation in the NAcc in the Supplement).

Notably, our findings complement results reported by Hanson *et al.* (19) that blunted NAcc activation in a sample of children and adolescents mediated the relationship between emotional neglect and depressive symptoms. Because stress sensitivity was not related to externalizing problems in the current study, we did not conduct a mediation analysis; nevertheless, both studies highlight the importance of conducting longitudinal fMRI assessments, and both studies implicate blunted activation in reward regions as a risk factor for the development of subsequent psychopathology.

We also assessed the temporal stability of percent signal change in the NAcc during the KIDMID task. We found low stability of the anticipation of reward versus anticipation of nonreward contrast. Certainly, this is an important consideration in interpreting our findings, but we should note that our reliability coefficients are comparable to that reported by Braams *et al.* (44), who assessed NAcc activation in 8- to 27-year-old participants twice over 2 years. In fact, the ABCD dataset has lower stability than we do here (38). It is possible that in longitudinal fMRI studies, repeated administrations of the same task desensitizes participants to the stimuli (50), leading them to habituate to the stimuli over time. Notably, participants in our sample felt less excited toward our reward cue across time (see Post-Scan Questionnaires in the Supplement). Altering perceptual components may decrease habituation but introduce methodological challenges.

Low ICC values may reflect neurodevelopmental changes over time in our sample. The approximately 4-year interval spanned by our assessments encompasses a large portion of adolescence. Reward processing and, specifically, the striatum undergo profound changes during this period [e.g., (44)], which may contribute to the discrepancies in blood oxygen level–dependent signal across time points. Distinguishing the contribution of methodological difficulties from true developmental changes is a crucial task for longitudinal studies [see Herting *et al.* (51) for a review of test-retest reliability in longitudinal task-based fMRI]. For a more detailed discussion of ICC, see Neural Stability Considerations and Suggestions for Improving Stability in Longitudinal Neuroimaging Studies in the Supplement.

We should note some limitations of this study. First, we had a community sample that did not experience the most severe levels of ELS. Experiences of more severe abuse and neglect likely have stronger effects on psychopathology and neural activation; thus, assessing trajectories of reward circuitry in more severely maltreated youths is an important direction for future research. Second, we conducted a number of statistical tests; the relationship between stress sensitivity and the trajectory of NAcc activation was small and did not survive correction for multiple comparisons, thereby underscoring the need to examine these constructs in larger samples. In addition, despite our efforts to ensure a robust and reliable signal, the ICC value for the NAcc in our study was low, although this may, in part, reflect developmental changes. In addition, while we think that our study yields important insights concerning the impact of stress sensitivity on adolescent reward processing and symptomatology, our findings should nevertheless be interpreted with caution. Finally, we focused on the NAcc in this study. Although we conducted supplemental analyses examining the putamen, caudate, mPFC, and several control regions, future researchers with larger samples should examine the longitudinal effects of ELS and stress sensitivity on different brain networks.

Despite these limitations, our study has several significant strengths, including the longitudinal examination of trajectories of neural activation from childhood through adolescence, which allowed us to assess intraindividual variability and development more dynamically than is done by cross-sectional assessments and to analyze dimensionally both objectively assessed ELS and stress sensitivity. By indexing symptoms of psychopathology at younger ages, we hope to improve the identification of samples at risk for developing clinically significant symptoms at older ages. Early prevention efforts may alleviate the experience of stress across development.

Conclusions

Conclusions drawn from cross-sectional examinations of NAcc activation in adults speak to the adverse effects of ELS decades after exposure to the stressors. Conducting longitudinal fMRI assessments in adolescents promises to advance our understanding of neurodevelopment because we are capturing changes in activation while teens are still in a stressful environment or relatively shortly after exposure to the stressor. Most studies conducted to date have assessed associations of ELS with reward brain regions cross-sectionally; in contrast, we collected and analyzed neuroimaging and behavioral data from 3 time points, which allowed us to assess longitudinal relationships among ELS and stress sensitivity,

trajectories of reward-related brain activation, and symptoms of psychopathology. Our findings indicate that elucidating trajectories of reward circuitry over adolescence is important for understanding the development of externalizing problems, particularly in adolescent boys.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the National Institute of Mental Health (Grant Nos. R37MH101495 [to IHG] and F32MH120975 [to RC]).

We thank the families who participated in this study and the research assistants at the Stanford Neurodevelopment, Affect, and Psychopathology Lab who collected the data.

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Department of Psychology, Stanford University, Stanford, California (LRB, JPY, RC, JR, AN, SMC, IHG); Department of Psychological Science, University of Arkansas, Fayetteville, Arkansas (JKL); and Department of Psychiatry and Behavioral Sciences, Stanford School of Medicine, Stanford, California (BJ).

Address correspondence to Lauren R. Borchers, Ph.D., at lrborchers@stanford.edu.

Received Nov 29, 2022; revised Jan 5, 2024; accepted Jan 11, 2024.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.biopsych.2024.01.011.

REFERENCES

1. Felitti VJ, Anda RF, Nordenberg D, Williamson DF, Spitz AM, Edwards V, *et al.* (1998): Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The adverse childhood experiences (ACE) study. Am J Prev Med 14:245–258.
2. Green JG, McLaughlin KA, Berglund PA, Gruber MJ, Sampson NA, Zaslavsky AM, Kessler RC (2010): Childhood adversities and adult psychiatric disorders in the national comorbidity survey replication I: Associations with first onset of DSM-IV disorders. Arch Gen Psychiatry 67:113–123.
3. Kim H, Wildeman C, Jonson-Reid M, Drake B (2017): Lifetime prevalence of investigating child maltreatment among US children. Am J Public Health 107:274–280.
4. Wildeman C, Emanuel N, Leventhal JM, Putnam-Hornstein E, Waldfogel J, Lee H (2014): The prevalence of confirmed maltreatment among US children, 2004 to 2011. JAMA Pediatr 168:706–713.
5. Pechtel P, Pizzagalli DA (2011): Effects of early life stress on cognitive and affective function: An integrated review of human literature. Psychopharmacol (Berl) 214:55–70.
6. LeMoult J, Humphreys KL, Tracy A, Hoffmeister JA, Ip E, Gotlib IH (2020): Meta-analysis: Exposure to early life stress and risk for depression in childhood and adolescence. J Am Acad Child Adolesc Psychiatry 59:842–855.
7. McLaughlin KA, Green JG, Gruber MJ, Sampson NA, Zaslavsky AM, Kessler RC (2010): Childhood adversities and adult psychiatric disorders in the national comorbidity survey replication II: Associations with persistence of DSM-IV disorders. Arch Gen Psychiatry 67:124–132.
8. Bogdan R, Nikolova YS, Pizzagalli DA (2013): Neurogenetics of depression: A focus on reward processing and stress sensitivity. Neurobiol Dis 52:12–23.
9. Ho TC, King LS, Leong JK, Colich NL, Humphreys KL, Ordaz SJ, Gotlib IH (2017): Effects of sensitivity to life stress on uncinate fasciculus segments in early adolescence. Soc Cogn Affect Neurosci 12:1460–1469.
10. Ho TC, Pham HT, Miller JG, Kircanski K, Gotlib IH (2020): Sympathetic nervous system dominance during stress recovery mediates associations between stress sensitivity and social anxiety symptoms in female adolescents. Dev Psychopathol 32:1914–1925.
11. Lewinsohn PM, Hops H, Roberts RE, Seeley JR, Andrews JA (1993): Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III—R disorders in high school students. J Abnorm Psychol 102:133–144.
12. Knutson B, Adams CM, Fong GW, Hommer D (2001): Anticipation of increasing monetary reward selectively recruits nucleus accumbens. J Neurosci 21:RC159.
13. Novick AM, Levandowski ML, Laumann LE, Philip NS, Price LH, Tyrka AR (2018): The effects of early life stress on reward processing. J Psychiatr Res 101:80–103.
14. Mehta MA, Gore-Langton E, Golembo N, Colvert E, Williams SCR, Sonuga-Barke E (2010): Hyporesponsive Reward Anticipation in the Basal Ganglia following Severe Institutional Deprivation Early in Life. J Cogn Neurosci 22:2316–2325.
15. Goff B, Gee DG, Telzer EH, Humphreys KL, Gabard-Durnam L, Flannery J, Tottenham N (2013): Reduced nucleus accumbens reactivity and adolescent depression following early-life stress. Neuroscience 249:129–138.
16. Hanson JL, Albert D, Iselin AM, Carré JM, Dodge KA, Hariri AR (2016): Cumulative stress in childhood is associated with blunted reward-related brain activity in adulthood. Soc Cogn Affect Neurosci 11:405–412.
17. Boecker R, Holz NE, Buchmann AF, Blomeyer D, Plichta MM, Wolf I, *et al.* (2014): Impact of early life adversity on reward processing in young adults: EEG-fMRI results from a prospective study over 25 years. PLoS One 9:e104185.
18. Müller KU, Mennigen E, Ripke S, Banaschewski T, Barker GJ, Büchel C, *et al.* (2013): Altered reward processing in adolescents with prenatal exposure to maternal cigarette smoking. JAMA Psychiatry 70:847–856.
19. Hanson JL, Hariri AR, Williamson DE (2015): Blunted ventral striatum development in adolescence reflects emotional neglect and predicts depressive symptoms. Biol Psychiatry 78:598–605.
20. Hawes SW, Waller R, Byrd AL, Bjork JM, Dick AS, Sutherland MT, *et al.* (2021): Reward processing in children with disruptive behavior disorders and callous-unemotional traits in the ABCD study. Am J Psychiatry 178:333–342.
21. Bjork JM, Chen G, Smith AR, Hommer DW (2010): Incentive-elicited mesolimbic activation and externalizing symptomatology in adolescents. J Child Psychol Psychiatry 51:827–837.
22. Yau WY, Zubieta JK, Weiland BJ, Samudra PG, Zucker RA, Heitzeg MM (2012): Nucleus accumbens response to incentive stimuli anticipation in children of alcoholics: Relationships with precursive behavioral risk and lifetime alcohol use. J Neurosci 32:2544–2551.
23. Schettini E, Wilson S, Beauchaine TP (2021): Internalizing–externalizing comorbidity and regional brain volumes in the ABCD study. Dev Psychopathol 33:1620–1633.
24. Gotlib IH, Hamilton JP, Cooney RE, Singh MK, Henry ML, Joormann J (2010): Neural processing of reward and loss in girls at risk for major depression. Arch Gen Psychiatry 67:380–387.
25. Luking KR, Pagliaccio D, Luby JL, Barch DM (2016): Reward processing and risk for depression across development. Trends Cogn Sci 20:456–468.
26. Ribbe D (1996): Psychometric review of traumatic event screening instrument for children (TESI-C). Meas Stress Trauma Adapt 386–387.
27. Rudolph KD, Hammen C, Burge D, Lindberg N, Herzberg D, Daley SE (2000): Toward an interpersonal life-stress model of depression: The developmental context of stress generation. Dev Psychopathol 12:215–234.
28. King LS, Humphreys KL, Camacho MC, Gotlib IH (2019): A person-centered approach to the assessment of early life stress: Associations with the volume of stress-sensitive brain regions in early adolescence. Dev Psychopathol 31:643–655.
29. Achenbach TM (1991): Manual for the Youth Self-Report and 1991 Profile. Burlington, VT: Department of Psychiatry, University of Vermont.
30. Leong JK, Ho TC, Colich NL, Sisk L, Knutson B, Gotlib IH (2021): White-matter tract connecting anterior insula to nucleus accumbens

predicts greater future motivation in adolescents. Dev Cogn Neurosci 47:100881.
31. Knutson B, Greer SM (2008): Anticipatory affect: Neural correlates and consequences for choice. Philos Trans R Soc B Biol Sci 363:3771–3786.
32. Wu Y, Na X, Zang Y, Cui Y, Xin W, Pang R, *et al.* (2014): Upregulation of tumor necrosis factor-alpha in nucleus accumbens attenuates morphine-induced rewarding in a neuropathic pain model. Biochem Biophys Res Commun 449:502–507.
33. Esteban O, Markiewicz CJ, Blair RW, Moodie CA, Isik AI, Erramuzpe A, *et al.* (2019): fMRIPrep: A robust preprocessing pipeline for functional MRI [no. 1]. Nat Methods 16:111–116.
34. Cox RW (1996): AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:162–173.
35. Shrout PE, Fleiss JL (1979): Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 86:420–428.
36. Baranger DAA, Lindenmuth M, Nance M, Guyer AE, Keenan K, Hipwell AE, *et al.* (2021): The longitudinal stability of fMRI activation during reward processing in adolescents and young adults. Neuroimage 232:117872.
37. Ordaz SJ, Foran W, Velanova K, Luna B (2013): Longitudinal growth curves of brain function underlying inhibitory control through adolescence. J Neurosci 33:18109–18124.
38. Kennedy JT, Harms MP, Korucuoglu O, Astafiev SV, Barch DM, Thompson WK, *et al.* (2022): Reliability and stability challenges in ABCD task fMRI data. NeuroImage 252:119046.
39. Kraemer HC, Wilson GT, Fairburn CG, Agras WS (2002): Mediators and moderators of treatment effects in randomized clinical trials. Arch Gen Psychiatry 59:877–883.
40. Little RJA, Rubin DB (1989): The analysis of social science data with missing values. Sociol Methods Res 18:292–326.
41. Galvan A, Hare TA, Parra CE, Penn J, Voss H, Glover G, Casey BJ (2006): Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. J Neurosci 26:6885–6892.
42. Steinberg L (2008): A social neuroscience perspective on adolescent risk-taking. Dev Rev 28:78–106.
43. Eiland L, Romeo RD (2013): Stress and the developing adolescent brain. Neuroscience 249:162–171.
44. Braams BR, van Duijvenvoorde ACK, Peper JS, Crone EA (2015): Longitudinal changes in adolescent risk-taking: A comprehensive study of neural responses to rewards, pubertal development, and risk-taking behavior. J Neurosci 35:7226–7238.
45. Plichta MM, Vasic N, Wolf RC, Lesch KP, Brummer D, Jacob C, *et al.* (2009): Neural hyporesponsiveness and hyperresponsiveness during immediate and delayed reward processing in adult attention-deficit/hyperactivity disorder. Biol Psychiatry 65:7–14.
46. Scheres A, Milham MP, Knutson B, Castellanos FX (2007): Ventral striatal hyporesponsiveness during reward anticipation in attention-deficit/hyperactivity disorder. Biol Psychiatry 61:720–724.
47. Ströhle A, Stoy M, Wrase J, Schwarzer S, Schlagenhauf F, Huss M, *et al.* (2008): Reward anticipation and outcomes in adult males with attention-deficit/hyperactivity disorder. NeuroImage 39:966–972.
48. Guyer AE, Kaufman J, Hodgdon HB, Masten CL, Jazbec S, Pine DS, Ernst M (2006): Behavioral alterations in reward system function: The role of childhood maltreatment and psychopathology. J Am Acad Child Adolesc Psychiatry 45:1059–1067.
49. Carlson SR, Pritchard AA, Dominelli RM (2013): Externalizing behavior, the UPPS-P Impulsive Behavior scale and Reward and Punishment Sensitivity. Personal Individ Dif 54:202–207.
50. Telzer EH, McCormick EM, Peters S, Cosme D, Pfeifer JH, van Duijvenvoorde ACK (2018): Methodological considerations for developmental longitudinal fMRI research. Dev Cogn Neurosci 33:149–160.
51. Herting MM, Gautam P, Chen Z, Mezher A, Vetter NC (2018): Test-retest reliability of longitudinal task-based fMRI: Implications for developmental studies. Dev Cogn Neurosci 33:17–26.

SUPPLEMENTARY INFORMATION

Sex-Specific Vulnerability to Externalizing Problems: Sensitivity to Early Stress and Nucleus Accumbens Activation Over Adolescence

Borchers *et al.*

Supplementary Material

Table 1S. *Studies Examining Reward.*

Study	Task	Contrast / condition	Finding
Goff et al., 2013	Emotional Faces task	Happy and Fear Face conditions	NAcc is blunted with higher ELS
Mehta et al., 2010	Monetary Incentive Delay Task	Anticipation of Reward vs. Anticpation of Non-Reward	NAcc is blunted with higher ELS across middle and high rewards
Boecker et al., 2014	Monetary Incentive Delay Task	Monetary vs. Verbal contrast	NAcc is blunted with higher ELS
Hanson et al., 2015	Card-guessing task	Win (positive) vs. Loss (negative) feedback	Males with greater ELS exhibited lower activation in the ventral striatum
Hanson, Hariri, & Williamson 2015	Card-guessing task	Positive vs. Negative feedback contrast	Decreased ventral striatum activation mediated relation between emotional neglect and internalizing symptoms
Gotlib et al., 2010	Monetary Incentive Delay Task	Anticipation of Reward vs. Anitcipation of Non-Reward and Receipt of Reward vs. Receipt of Non-Reward	High-risk youth have decreased ventral striatum activation compared to controls.
Luking et al., 2016	Card-Guessing Task	Reward vs. Neutral contrast	High-risk youth have decreased ventral striatum activation compared to controls.
Bjork et al., 2010	Monetary Incentive Delay Task	Receiving vs. not Receiving Rewards	Adolescents with externalizing disorders showed greater NAcc activation compared to controls.
Yau et al., 2012	Monetary Incentive Delay Task	Anticipation of Reward vs. Anticipation of Non-Reward and Anticipation of Loss vs. Anticipation of Non-Loss	Children with a parent with alcoholism exhibited a positive association between externalizing problems and NAcc activation.
Braams et al., 2015	Gambling Task	Wins vs. Losses	NAcc activation peaks in adolescence

Schreuders et al., 2018	Gambling Task	Wins vs. Losses	NAcc activation peaks in adolescence
Insel & Somerville, 2018	Magnitude Tracking Task	High Gains vs. Low Gains	Negative association between gain magnitude tracking and age
Plichta et al., 2009	Monetary intertemporal choice paradigm	Immediate and Delayed Choices	Adults with ADHD had blunted NAcc activation
Scheres et al., 2007	Monetary Incentive Delay Task	Anticipation of Reward vs. anticipation of non-Reward	Adolescents with ADHD had blunted NAcc activation
Strohle et al., 2008	Monetary Incentive Delay Task	Anticipation of Reward vs. anticipation of non-Reward	Adults with ADHD had blunted NAcc activation

Additional Background

While several studies have shown that ELS is associated with aberrant reward circuitry activation cross-sectionally, no researchers have examined *trajectories* of activation, which is important for advancing our understanding of the effects of ELS on neurodevelopment through childhood and adolescence. In this context, Novick and colleagues (2018) highlighted the neurobiological overlap of stress and reward processing, postulating that reward circuitry is involved in mediating the association between ELS and subsequent psychopathology. Indeed, Dillon et al. (2009) and Goff et al. (2013) provided evidence that decreased activation in the putamen and NAcc, respectively, are associated with greater symptoms of depression. Further, Hanson, Hariri, and Williamson (2015) reported that the relation between emotional neglect and depressive symptoms is partially mediated by decreased ventral striatum activation. These researchers speculate that treatment targeting reward processing and positive affect will be particularly effective (Hanson, Hariri, and Williamson 2015).

It is important to note that aberrant activation in reward circuitry may precede the development of internalizing and externalizing symptoms, given that high-risk youth have been found to have decreased activation in the reward response compared to control participants

(Gotlib et al., 2010; Luking et al., 2016); however, we do not yet understand the timing and developmental course of these systems. Indeed, researchers have generally used cross-sectional study designs to examine relations among ELS, psychopathology, and reward circuitry activation. Consequently, we know little about the onset of internalizing and externalizing symptoms as a consequence of having experienced and reacted to ELS and whether the aberrant reward processing documented in cross-sectional studies persists or changes over adolescence.

Recruitment and Exclusion Criterion

Children were recruited to participate in the study through media advertisements and flyers. Children were excluded from participating in the study if they had a history of a neurological disorder or major illness, had contraindications for participating in an MRI scan (e.g., non-removable metal), were not fluent English speakers, had cognitive or physical challenges that interfered with their ability to participate in study procedures, or, for females, had the onset of menses prior to the baseline assessment. Participants' fMRI data were excluded for a variety of reasons, including excessive motion (Time 1: *n*=44; Time 2: *n*=21; Time 3: *n*=10), errors in the raw timing files (Time 1: *n*=5; Time 2: *n*=3; Time 3: *n*=0), and poor engagement measured via a hit rate <30% (Time 1: *n*=7; Time 2: *n*=0; Time 3 *n*=2). In total, 146, 123, and 107 participants had usable fMRI data at Times 1, 2, and 3, respectively. Three participants withdrew from the study at baseline and, therefore, were excluded. Regarding behavioral data, those missing the ELS interview (*n*=2), a stress sensitivity score (*n*=5), or the YSR at baseline (*n*=7) were excluded.

Income-to-Needs Ratio

We measured socioeconomic status by calculating the income-to-needs ratio for each participant. Participant's parents reported on their annual household income in bins (1: ≤$5,000,

2: $5,001-$10,000, 3: $10,001-$15,000, 4: $15,001-$25,000, 5: $25,001-$35,000, 6: $35,001-$50,000, 7: $50,001-$75,000, 8: $75,001-$100,000, 9: $100,001-$150,000, 10: ≥$150,000). We divided the midpoint of their reported income bin by the low-income value for Santa Clara County. Importantly, this considers the number of people in the home and the time period in which the study occurred (https://www.huduser.gov/portal/datasets/il/il2017/2017summary.odn; (1)).

Pubertal Assessment

Pubertal status was examined via the Tanner Staging Questionnaire (2). Stage 1 is pre-pubertal, stage 2 marks the onset of puberty, stages 3 and 4 indicate more pronounced development, and stage 5 indicates full pubertal development. We averaged the two self-reported measures to yield an average index of pubertal development. Boys and girls were pre-pubertal at baseline on average and did not differ significantly in their pubertal stage (see Table 1).

MRI Scan Acquisition

We collected whole-brain functional scans throughout the task using a T2*-weighted gradient pulse sequence (43 oblique slices, TR=2 seconds, TE=30 ms, flip angle=77º, resolution=3.2 x 3.2 x 3.0 mm voxels, interleaved acquisition, total scan time=8:20 minutes). In addition, we acquired a high-resolution spoiled gradient echo (SPGR) T1-weighted anatomical scan (TR=6.3 ms, TE=2.4 ms, flip angle=12º, resolution=0.90 mm isotropic, scan time=5:16 minutes) for registration of the functional data.

Table 2S. *Prevalence of Objective ELS in the Study Sample.*

Type of ELS	Prevalence
Experience accident	22%
Witness accident	24%
Natural disaster	5%
Experience illness	29%
Witness illness	52%
Death of someone close	32%
Separation due to work / travel	9%
Moving	48%
Divorce	30%
Separation due to rehab / prison	16%
Witness mental illness	23%
Witness self-harm	7%
Physical abuse	5%
Threats of physical abuse	2%
Bullying	34%
Kidnapping	2%
Animal attack	6%
Domestic violence (non-physical)	40%
Domestic violence (threat)	3%
Domestic violence (physical)	9%
Legal trouble	11%
Witness stranger arguments	5%
Witness stranger fights	11%
Witness war	3%
Victim of sexual abuse	2%

Witness sexual abuse	0.005%
Emotional abuse	6%
Neglect	6%
Financial troubles	14%
Financial (other)	8%
Community instability	13%
Mugging	4%

Note: Domains from the Traumatic Events Screening Inventory for Children.

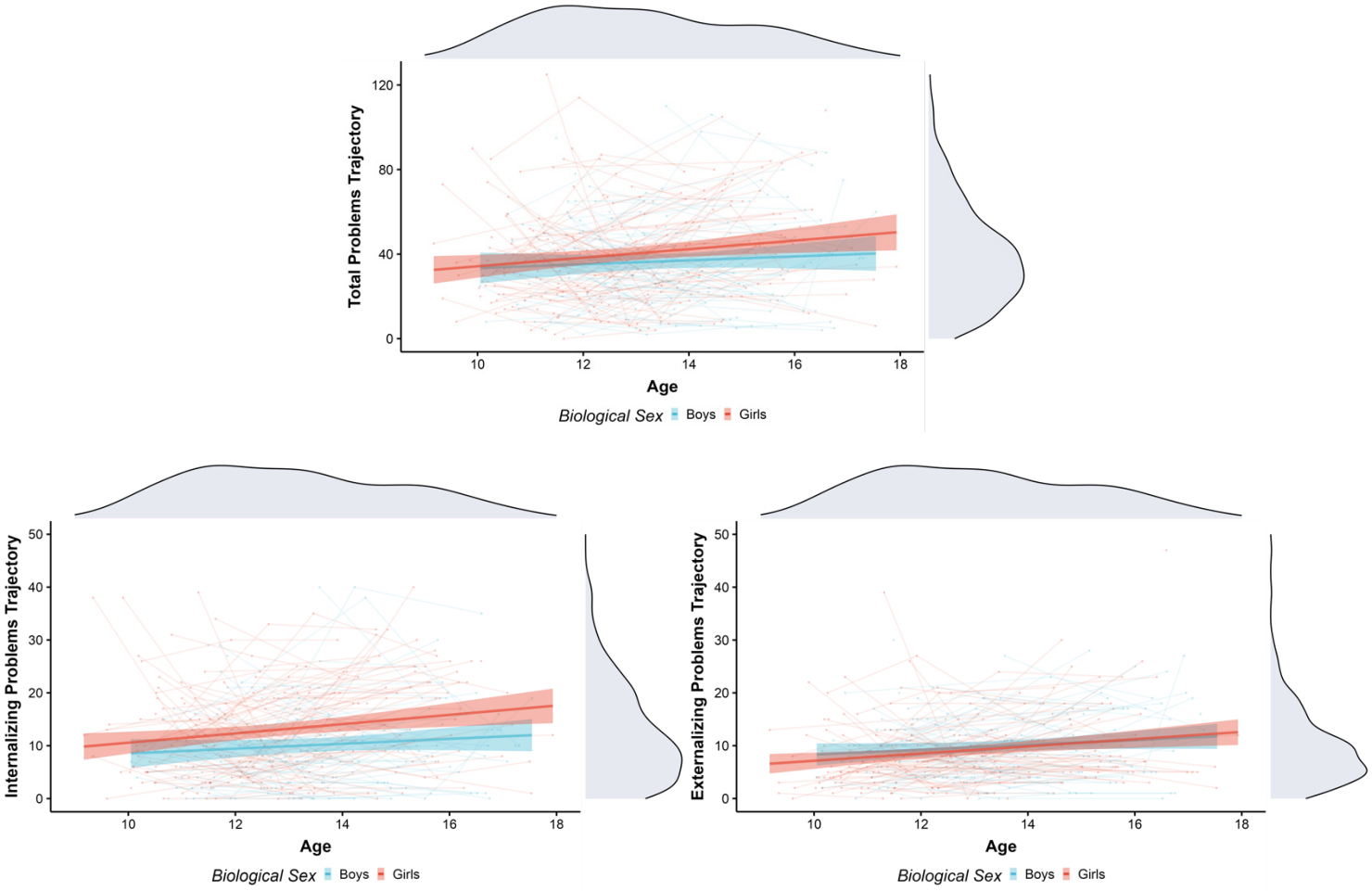
Scanner Upgrade/COVID-19 Status

The 3T Discovery MR750 scanner was upgraded to an Ultra High-Performance System on March 16th, 2020. This coincided with the COVID-19 pandemic, in which in-person research sessions were temporarily halted. Twenty-six participants (18 of whom had usable scan data) completed their Time 3 session after the scanner upgrade and following COVID-19 shelter-in-place orders. Therefore, we controlled for this in all analyses assessing neural activation at the third timepoint using a binary dummy-coded variable referred to as “collection status”.

Demographics and Stress Variables

Approximately half of the sample was female and the majority were right-handed (n=156 right-handed; n=13 left-handed; n=1 ambidextrous; n=3 missing data). Although there were no sex differences in objective ELS (p=.339), girls were more sensitive to stressors than were boys (t=-2.32, p=.022). Objective ELS and stress sensitivity were not related to pubertal stage at baseline (all ps>.053). There was a negative association between objective ELS and income-to-needs ratio (r=-0.31, p<.001). There was no relation between objective ELS and age at baseline (p=.670). Further, stress sensitivity was not related to income-to-needs ratio (p=.980) or age at baseline (p=.085).

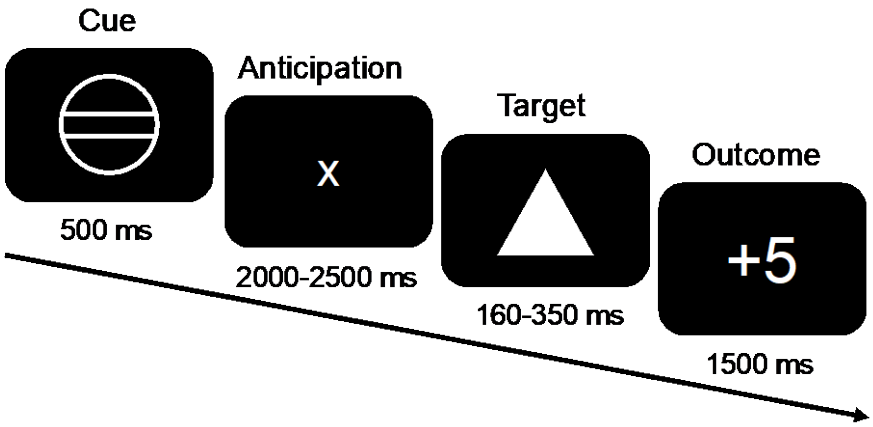
Figure 1S. *Distribution of Youth-Self-Report Scores Across Adolescence.*



Note. x-axis is age across 3 timepoints; y-axis is raw YSR problems. Solid lines reflect each participant measured across time. The red line is the trajectory for girls and the blue line is the trajectory for boys. The density plots on the top are for age and the density plots on the right are for the measure of interest (i.e., total problems, internalizing, externalizing).

16.4% of the sample had T-scores of 65 or greater on internalizing or externalizing problems at Time 3. These figures are consistent with base rates reported in the general population (Merikangas et al., 2009). Further, 67.6% of the sample increased in internalizing or externalizing problems across adolescence, as indexed by a positive slope.

Figure 2S. *Child-Friendly Monetary Incentive Delay Task.*



Note. Example of a successful reward trial.

KIDMID Procedure

Participants completed two practice rounds of the KIDMID task outside of the scanner to determine the difficulty of the task (20 trials to determine the difficulty for the second practice, 32 trials to determine the difficulty for the scanner) and to ensure they were familiar with the task prior to scanning. Participants were given a button box and responded with their dominant hand. There were four different conditions: reward, loss, non-reward, and non-loss. Reward and non-reward trials were indicated by circle cues, and loss and non-loss trials were indicated by square cues. If participants responded to a circle-cue target in time, they would successfully win points (+5; +0 for non-reward condition). If participants did not respond to a square-cue target in time, they would lose points (-5; -0 for non-loss condition). Reaction time was adjusted using a staircase procedure to ensure that participants' hit rate was 66% across all conditions.

Performance on the KIDMID Task

We examined hit rate (number of hits/total number of trials) and reaction times for all trials. In addition, we assessed participant's overall performance and affect towards cues across

all timepoints. We used participant's hit rate as a proxy for engagement. Participants performed well on the KIDMID task on average. Because of the staircase procedure used to determine the difficulty for participants, there were no differences in their hit rates across reward or loss trails at any of the three timepoints (all *p*s>.050); however, participants responded faster to reward trials than to loss trials at all three timepoints (Time 1: $t(146)=2.06$, $p=.041$; Time 2: $t(122)=2.49$, $p=.014$; Time 3: $t(107)=2.44$, $p=.016$).

Table 3S. *Behavior on the KIDMID Task.*

Variable	***M* (*SD*)**
Hit Reward Trials T1	70.89% (15.42%)
Hit Loss Trials T1	66.91% (16.20%)
Hit Reward Trials T2	71.41% (12.40%)
Hit Loss Trials T2	68.26% (13.33%)
Hit Reward Trials T3	73.61% (11.65%)
Hit Loss Trials T3	68.58% (13.35%)
Reaction Time Reward T1	224.99 ms (35.41 ms)
Reaction Time Loss T1	229.07 ms (36.07 ms)
Reaction Time Reward T2	221.18 ms (27.13 ms)
Reaction Time Loss T2	225.61 ms (29.04 ms)
Reaction Time Reward T3	223.47 ms (26.97 ms)
Reaction Time Loss T3	227.96 ms (24.83 ms)
Total Performance T1	46.02 points (33.37 points)
Total Performance T2	48.98 points (25.44 points)
Total Performance T3	52.36 points (24.85 points)

Note. T1=Time 1, T2=Time 2, and T3=Time 3; ms=milliseconds.

Preprocessing

We implemented Advanced Normalization Tools (ANTs v.2.2.0; Avants et al., 2008) to adjust for intensity non-uniformity and to skull-strip the T1w anatomical scan. Next, brain tissue was segmented into gray matter, white matter, and CSF with fast from the FMRIB Software Library (FSL v.5.0.9; Smith et al., 2004; Zhang et al., 2001). We applied *recon_all* in Freesurfer (v.6.0.1; Dale et al., 1999) to reconstruct the brain surface and applied volume-based spatial normalization to the MNIPediatricAsym age-appropriate cohort (Cohort 4 age range 7.5-13.5 for Time 1 and Cohort 6 age range 13.5-18 for Time 2 and Time 3; Fonov et al., 2011). We applied normalization using antsRegistration. Next, an unwarped reference volume was created, and the fMRI data were skull-stripped. The BOLD reference was co-registered to the T1w reference with 6 degrees of freedom using *bbregister* from FreeSurfer (Greve & Fischl, 2009). We estimated head motion with FSL's MCFLIRT (Jenkinson et al., 2002) and slice-time corrected BOLD runs using 3dTshift from Analysis of Functional Neuroimages (AFNI v.20160207; Cox & Hyde, 1997). Motion outliers were identified as frames with more than 0.5 mm framewise displacement, or 1.5 standardized derivative of root mean square variance over voxels (Power et al., 2014; Smyser et al., 2011). Participants were excluded if over 20% of their volumes were marked as outliers. Further, we conducted motion censoring. Specifically, if any task volume was flagged as a motion outlier, the value was dropped from the analysis (i.e., percent signal change values with a motion outlier were replaced with "N/A" in the statistical analysis). In addition to identifying motion outliers, we rigorously quality-checked all raw and preprocessed data at all three timepoints to ensure that the scans included in the analysis did not have excessive motion or artifacts.

fMRIPrep (version 20.2.1)

Anatomical Data Preprocessing

A total of 1 T1-weighted (T1w) images were found within the input BIDS dataset. The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al. 2010), distributed with ANTs 2.3.3 (Avants et al. 2008, RRID:SCR_004757), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a *Nipype* implementation of the antsBrainExtraction.sh workflow (from ANTs), using MNIPediatricAsym:cohort-4 as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.9, RRID:SCR_002823, Zhang, Brady, and Smith 2001). Volume-based spatial normalization to two standard spaces (MNIPediatricAsym:cohort-4, MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.3.3), using brain-extracted versions of both T1w reference and the T1w template. The following templates were selected for spatial normalization: *MNI's unbiased standard MRI template for pediatric data from the 4.5 to 18.5y age range* [RRID:SCR_008796; TemplateFlow ID: MNIPediatricAsym:cohort-4], *ICBM 152 Nonlinear Asymmetrical template version 2009c* [Fonov et al. (2009), RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym],

Functional Data Preprocessing

For each of the 1 BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Susceptibility distortion correction (SDC) was omitted. The BOLD reference was then co-registered to the T1w reference using flirt (FSL 5.0.9, Jenkinson and Smith 2001) with the boundary-based registration (Greve and Fischl 2009) cost-function. Co-registration was configured with nine degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters with respect

to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9, Jenkinson et al. 2002). BOLD runs were slice-time corrected using 3dTshift from AFNI 20160207 (Cox and Hyde 1997, RRID:SCR_005927). The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying the transforms to correct for head-motion. These resampled BOLD time-series will be referred to as *preprocessed BOLD in original space*, or just *preprocessed BOLD*. The BOLD time-series were resampled into standard space, generating a *preprocessed BOLD run in MNIPediatricAsym:cohort-4 space*. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, Power et al. (2014)) and Jenkinson (relative root mean square displacement between affines, Jenkinson et al. (2002)). FD and DVARS are calculated for each functional run, both using their implementations in *Nipype* (following the definitions by Power et al. 2014). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (*CompCor*, Behzadi et al. 2007). Principal components are estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, the aCompCor masks are subtracted a mask of pixels that likely contain a volume fraction of GM. This mask is obtained by thresholding the corresponding partial volume map at 0.05, and it ensures components are not

extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the *k* components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al. 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos 1964). Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer).

Many internal operations of *fMRIPrep* use *Nilearn* 0.6.2 (Abraham et al. 2014, RRID:SCR_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in *fMRIPrep*'s documentation.

Copyright Waiver

The above boilerplate text was automatically generated by fMRIPrep with the express intention that users should copy and paste this text into their manuscripts *unchanged*. It is released under the CC0 license.

fMRIPrep References

Abraham, Alexandre, Fabian Pedregosa, Michael Eickenberg, Philippe Gervais, Andreas Mueller, Jean Kossaifi, Alexandre Gramfort, Bertrand Thirion, and Gael Varoquaux. 2014. "Machine Learning for Neuroimaging with Scikit-Learn." *Frontiers in Neuroinformatics* 8. https://doi.org/10.3389/fninf.2014.00014.

Avants, B.B., C.L. Epstein, M. Grossman, and J.C. Gee. 2008. "Symmetric Diffeomorphic Image Registration with Cross-Correlation: Evaluating Automated Labeling of Elderly and Neurodegenerative Brain." *Medical Image Analysis* 12 (1): 26–41. https://doi.org/10.1016/j.media.2007.06.004.

Behzadi, Yashar, Khaled Restom, Joy Liau, and Thomas T. Liu. 2007. "A Component Based Noise Correction Method (CompCor) for BOLD and Perfusion Based fMRI." *NeuroImage* 37 (1): 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042.

Cox, Robert W., and James S. Hyde. 1997. "Software Tools for Analysis and Visualization of fMRI Data." *NMR in Biomedicine* 10 (4-5): 171–78. https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L.

Esteban, Oscar, Ross Blair, Christopher J. Markiewicz, Shoshana L. Berleant, Craig Moodie, Feilong Ma, Ayse Ilkay Isik, et al. 2018. "FMRIPrep." *Software*. Zenodo. https://doi.org/10.5281/zenodo.852659.

Esteban, Oscar, Christopher Markiewicz, Ross W Blair, Craig Moodie, Ayse Ilkay Isik, Asier Erramuzpe Aliaga, James Kent, et al. 2018. "fMRIPrep: A Robust Preprocessing Pipeline for Functional MRI." *Nature Methods*. https://doi.org/10.1038/s41592-018-0235-4.

Fonov, VS, AC Evans, RC McKinstry, CR Almli, and DL Collins. 2009. “Unbiased Nonlinear Average Age-Appropriate Brain Templates from Birth to Adulthood.” *NeuroImage* 47, Supplement 1: S102. https://doi.org/10.1016/S1053-8119(09)70884-5.

Gorgolewski, K., C. D. Burns, C. Madison, D. Clark, Y. O. Halchenko, M. L. Waskom, and S. Ghosh. 2011. “Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python.” *Frontiers in Neuroinformatics* 5: 13. https://doi.org/10.3389/fninf.2011.00013.

Gorgolewski, Krzysztof J., Oscar Esteban, Christopher J. Markiewicz, Erik Ziegler, David Gage Ellis, Michael Philipp Notter, Dorota Jarecka, et al. 2018. “Nipype.” *Software*. Zenodo. https://doi.org/10.5281/zenodo.596855.

Greve, Douglas N, and Bruce Fischl. 2009. “Accurate and Robust Brain Image Alignment Using Boundary-Based Registration.” *NeuroImage* 48 (1): 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060.

Jenkinson, Mark, Peter Bannister, Michael Brady, and Stephen Smith. 2002. “Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images.” *NeuroImage* 17 (2): 825–41. https://doi.org/10.1006/nimg.2002.1132.

Jenkinson, Mark, and Stephen Smith. 2001. “A Global Optimisation Method for Robust Affine Registration of Brain Images.” *Medical Image Analysis* 5 (2): 143–56. https://doi.org/10.1016/S1361-8415(01)00036-6.

Lanczos, C. 1964. “Evaluation of Noisy Data.” *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis* 1 (1): 76–85. https://doi.org/10.1137/0701007.

Power, Jonathan D., Anish Mitra, Timothy O. Laumann, Abraham Z. Snyder, Bradley L. Schlaggar, and Steven E. Petersen. 2014. “Methods to Detect, Characterize, and Remove Motion Artifact in Resting State fMRI.” *NeuroImage* 84 (Supplement C): 320–41. https://doi.org/10.1016/j.neuroimage.2013.08.048.

Satterthwaite, Theodore D., Mark A. Elliott, Raphael T. Gerraty, Kosha Ruparel, James Loughead, Monica E. Calkins, Simon B. Eickhoff, et al. 2013. “An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data.” *NeuroImage* 64 (1): 240–56. https://doi.org/10.1016/j.neuroimage.2012.08.052.

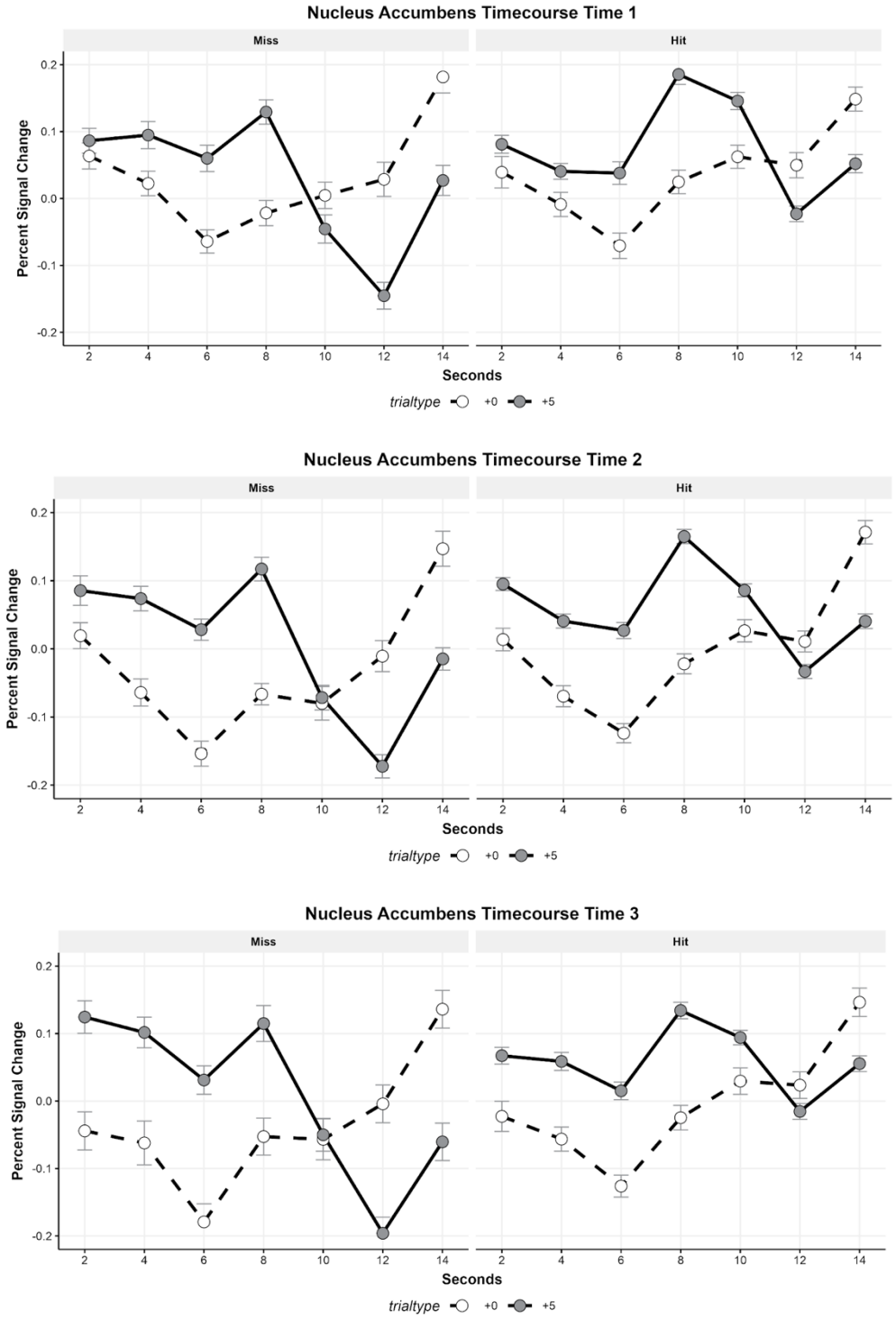
Tustison, N. J., B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee. 2010. “N4ITK: Improved N3 Bias Correction.” *IEEE Transactions on Medical Imaging* 29 (6): 1310–20. https://doi.org/10.1109/TMI.2010.2046908.

Zhang, Y., M. Brady, and S. Smith. 2001. “Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm.” *IEEE Transactions on Medical Imaging* 20 (1): 45–57. https://doi.org/10.1109/42.906424.

AFNI Preprocessing

We trimmed the lead-in (19 volumes) and lead-out (15 volumes) volumes (total of 216 task volumes), corrected for motion in six different directions, spatially smoothed our images using a 4mm kernel (3), normalized the signal by converting scanner units to raw percent signal change, and removed signals slower than 90 seconds with a high-pass filter.

Figure 3S. *Timecourse of the NAcc Across Missed and Hit Trials.*



Note: Second 6 is the anticipation phase; second 10 is the receipt phase. Reward trials are indicated by solid black lines and non-reward trials by dashed black lines. Missed trials (left panel) indicate when participants did not respond to the triangle target in time (+0 points) whereas hit trials (right panel) indicate when participants responded to the triangle target in time (+5 points); (Time 1 *n*=146, Time 2 *n*=123, Time 3 *n*=107). Error bars are shown in gray.

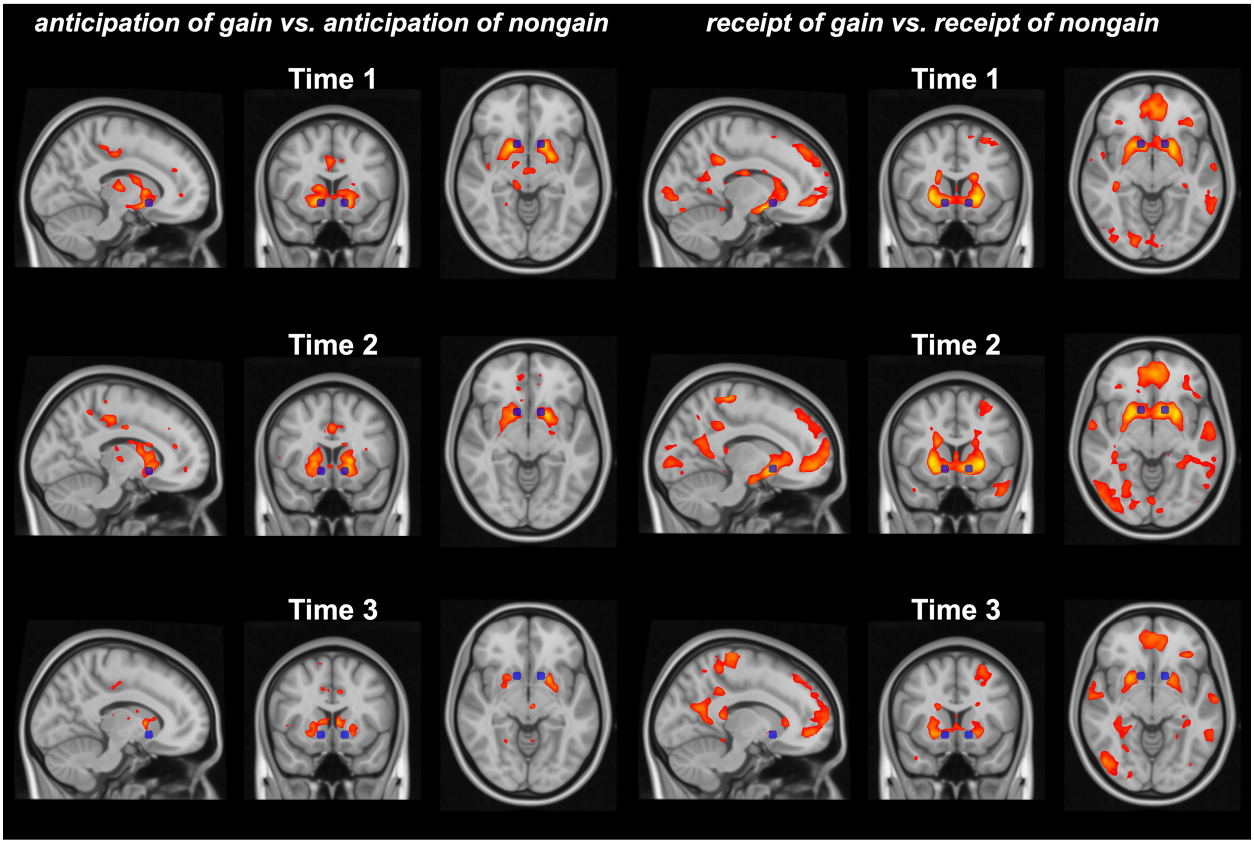
Whole Brain Analysis

We conducted whole-brain analysis in FSL (v. 6.0.1) with *FEAT* for two contrasts of interest: anticipation of reward vs. anticipation of non-reward and receipt of reward vs. receipt of non-reward (MRI Expert Analysis Tool; (4,5)). We trimmed the first five volumes, applied a 5 mm spatial smoothing, and a 90 second high-pass filtering to the fMRI data. The general linear model (GLM) includes regressors for the MID task in addition to 25 regressors from fMRIPrep (e.g., head motion, temporal derivatives and quadratic terms, three global signals, and the six physiological CompCor components). We censored motion outliers as described by Siegal and colleagues (6). We applied a double-gamma convolution to model the hemodynamic response function.

We conducted a whole-brain cluster analysis to test for main effects of the two reward contrasts (i.e., anticipation of reward vs. anticipation of non-reward and receipt of reward vs. receipt of non-reward) during the MID task. We conducted a data-driven cluster analysis to identify voxels that displayed significant activation during these clusters for each participant using FSL's FLAME I. The cluster threshold was set to Z=3.1, with a multiple-comparisons-corrected cluster threshold of α=0.025. We controlled for biological sex at all timepoints and for collection status at Time 3.

The cluster maps show NAcc activation during the anticipation of reward vs. the anticipation of non-reward and the receipt of reward vs. the receipt of non-reward at Time 1 and Time 2; however, this activation is reduced at Time 3.

Figure 4S. *Group-Level Results Indicating Increased Activation During the KIDMID Task at each Timepoint.*



Note: The bilateral blue spheres visualize the bilateral NAcc ROI. Significant increased activation is indicated in red. We modeled both missed and hit trials for the receipt of reward vs. receipt of non-reward contrast.

Whole-Brain Analysis

Lower activation in the NAcc at Time 3 may be due to the smaller sample size (Time 1=146, Time 2=123, Time 3=107), to developmental effects, and/or to task adaptation (7). Further, we should consider the design of the KIDMID task. Participants were told they can win (+5 points) or avoid losing points (-5 points) to receive a prize. While this format was developmentally appropriate when children entered our study at baseline, these outcomes may have been less rewarding to participants at older ages. Analogous tasks administered in adolescent and adult populations often provide a monetary stimulus (i.e., +$5, -$5; (8)).

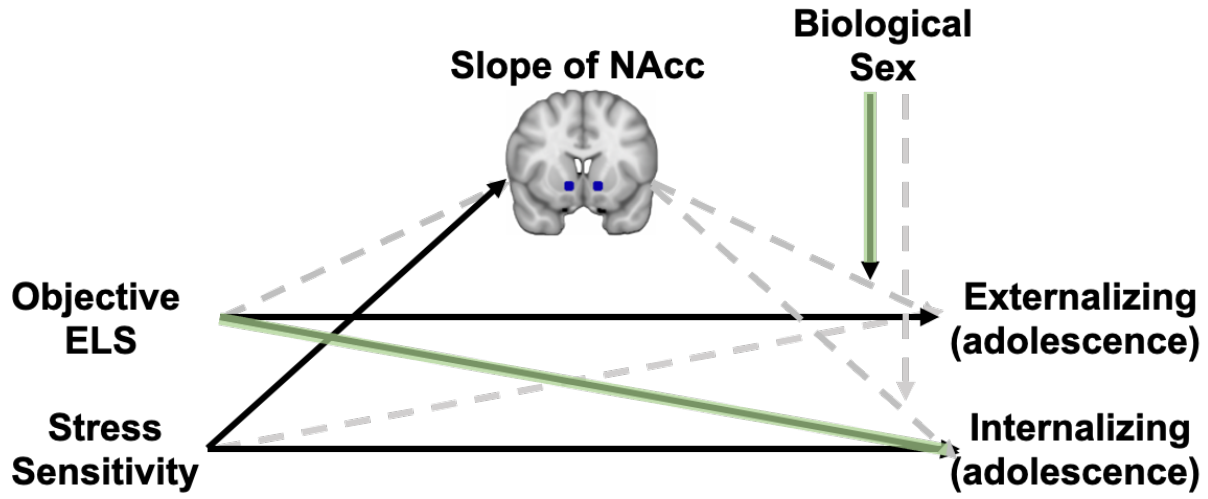
Post-Scan Questionnaires

It is possible that changes in NAcc activation across time were due to differences in participants' affect while they viewed the reward cues. For example, it is possible that participants will feel less positive after seeing a circle cue (reward trial) with repeated administrations. We asked participants "*How did you feel when you saw [image of circle cue with two lines]?"* after completing the KIDMID task. A 1 would indicate they felt negative toward the cue and a 7 would indicate they felt positive toward the cue. On average, participants felt positively toward the circle cue at all three timepoints (*M* Time 1=6.22; *M* Time 2=6.06; *M* Time 3=6.01). Participants who completed the post-scan questions felt less positively toward the circle cue over time (β=-0.11, *p*=.023).

Table 4S. *Stability of Contrasts and Conditions Across All Trials.*

Variable	**ICC**
Anticipation of reward vs. anticipation of non-reward	0.32
Receipt of reward vs. receipt of non-reward	0.08
Anticipation of reward	0.34
Anticipation of non-reward	0.20
Receipt of reward	0.33
Receipt of non-reward	0.00

Figure 5S. *Primary Findings Schematic.*



Note: Solid black lines indicate significant p-values, lines highlighted in green indicate p-values that survived Bonferroni correction, and gray dashed lines indicate non-significant p-values.

Attrition

63 participants had ELS measures and usable fMRI data at all three timepoints. We compared these participants to participants who were missing at least one timepoint of data due to motion, attrition, etc (n=110). These two groups did not differ in the proportion of males and females (p=.458), race (p=.151), age at baseline (p=.118), pubertal stage at baseline (p=.516), severity of objective ELS (p=.160), stress sensitivity (p=.517), income-to-needs ratio (p=.101), baseline total problems (p=.659), or baseline NAcc activation during the anticipation of reward vs. anticipation of non-reward contrast (p=.184).

Sensitivity Analysis: Time 3 YSR Data

We obtained a consistent pattern of results when conducting a parallel analysis using Time 3 YSR data (n=134) as the outcome variable rather than intercepts centered at age 15 obtained from lme analysis (n=173). There was a significant interaction between NAcc trajectories and biological sex when predicting psychopathology at Time 3 (β=0.08, p=.004);

findings were unique to externalizing (β=0.37, *p*=.030) and not internalizing problems (*p*=.273). Follow-up simple-slope analyses indicated that boys (β=-0.28, *p*=.010), but not girls (*p*=.490), with decreasing NAcc trajectories had greater externalizing problems at Time 3. The same pattern of findings held when using T-scores at Time 3 rather than raw scores. Specifically, there was an interaction of biological sex and NAcc trajectories in predicting internalizing and externalizing T-scores (β=0.08, *p*=.004) even after controlling for baseline total problems T-scores. The interaction was significant for externalizing (β=0.37, p=.028), but not for internalizing (β=-0.26, p=.328) problems T-scores. Follow-up simple slope analysis indicated that boys (β=-0.28, p=.010), but not girls (β=0.08, p=.560), with decreasing NAcc trajectories had greater externalizing problems T-scores in adolescence.

Sensitivity Analysis: Non-Winsorized Values

Estimated NAcc trajectories were winsorized to 3 standard deviations above or below the mean to reduce the impact of extreme cases (*n*=1). When examining non-winsorized NAcc values, we also obtained a similar pattern of results reported in the main text. Specifically, there was an interaction between NAcc trajectories and biological sex (β=0.08, *p*=.001); the interaction was significant for externalizing (β=0.37, *p*=.014) but not internalizing problems (*p*=.497). Boys (β=-0.22, *p*=.010), but not girls (*p*=.210), with decreasing NAcc trajectories had more externalizing problems in adolescence.

Exploratory Analyses

Exploratory analyses demonstrated there was no relation of objective ELS with *trajectories* of internalizing (β=-0.11, *p*=.149) or externalizing problems (β=-0.01, *p*=.933), or of stress sensitivity with *trajectories* of internalizing (β=0.01, *p*=.875) or externalizing problems (β=0.00, *p*=.962). Stress sensitivity (β=0.16, *p*=.042), but not objective ELS (β=-0.09; *p*=.246), was related to the NAcc intercept. Further, there was an interaction of biological sex and NAcc

trajectories in predicting *trajectories* of psychopathology (β=0.08, *p*=.001). Specifically, the interaction was present for externalizing (β=0.42, *p*=.005), but not for internalizing (β=-0.09, *p*=.566) problems. Boys (β=-0.28, *p*=.010), but not girls (β=0.14, *p*=.200), with decreasing NAcc trajectories had increasing *trajectories* of externalizing problems.

Equations

Relation of objective ELS and stress sensitivity on NAcc trajectory:

$$\text{NAcc trajectory}=\beta_{\text{Objective ELS}}+\beta_{\text{Stress Sensitivity}}+\beta_{\text{Collection Status}}+\beta 0$$

Relation of objective ELS and stress sensitivity on psychopathology in adolescence:

$$\text{Internalizing \& Externalizing}=\beta_{\text{Objective ELS}}+\beta_{\text{Stress Sensitivity}}+\beta_{\text{Collection Status}}+\beta 0$$

Interaction between NAcc trajectories and biological sex on psychopathology in adolescence:

$$\text{Internalizing \& Externalizing}=\beta_{\text{NAcc Trajectory}}*\beta_{\text{Biological Sex}}+\beta_{\text{Collection Status}}+\beta_{\text{Baseline Total Problems}}+\beta 0$$

Sensitivity Analysis: Missing at Random

We conducted an analysis including only the 63 participants with usable KIDMID data at all three timepoints, also controlling for pre-post COVID-19 lockdown data collection status. The findings were mostly replicated, although some associations were reduced to a trend level. This is not surprising given the power difference and reduction of our sample from 173 to 63 participants. Specifically, stress sensitivity and NAcc trajectories were correlated at *r*=-0.19, *p*=.132, and the interaction between NAcc trajectories and biological sex in predicting externalizing problems trended at β=2.31, *p*=.059. The trajectory was significantly negative in boys (β=-0.41, *p*=.030) and nonsignificant in females (*p*=.660). Aside from the sample size reduction, our results are robust to the employed missing data assumption: missing at random (MAR; (9)).

Typical NAcc Development

Few studies have examined repeated assessments of neural activation and even fewer have described trajectories of NAcc activation. Braams et al., (10) examined participants aged 8-27 years old and found that NAcc activation in a contrast of wins vs. losses during a gambling task peaks in mid-adolescence. Schreuders and colleagues (11) extended these findings in the same sample of participants with the addition of a third wave of fMRI data covering the ages of 8-29. Further, Insel and Somerville (12) found that gain magnitude tracking (i.e., heightened response to large gains compared to small gains) demonstrated a linear negative relation across time in a sample of 13- to 20-year-old participants. Importantly, Insel and Somerville (12) examined only one timepoint of neural activation.

Hyperactivation in the NAcc

In contrast to our results, Yau et al. (13) and Bjork et al. (14) found that adolescents with externalizing disorders were characterized by greater NAcc activation. Aspects of the designs of these studies, including adopting a categorical approach to examining externalizing problems and only examining one timepoint of NAcc activation, may help explain these discrepancies.

Neural Stability Considerations

The receipt of reward vs. receipt of non-reward contrast had poor stability regardless of whether we modeled all trials or only hit trials, findings that are consistent with Bach et al.'s (15) suggestion that the reliability of conditions are stronger than the reliability of contrasts. One reason for the poor stability within the receipt of reward vs. receipt of non-reward contrast is that both the active conditions (i.e., receipt of reward) and the neutral conditions (i.e., receipt of non-reward) have some degree of noise. In this context, we found that whereas the ICC for the condition of the receipt of reward was modest, the ICC for the receipt of non-reward was zero (Table 3S). As a related point, Baranger and colleagues (16) reported that contrasts of one

active state were preferable to contrasts of two active states during reward tasks measured across 2-year intervals. Given the ICC values in our study and the desired specificity of measuring an active condition, we decided that it was optimal to examine neural activation during the contrast of anticipation of reward vs. anticipation of non-reward.

Suggestions for Improving Stability in Longitudinal Neuroimaging Studies

With respect to study design, it is ideal to use the same scanner throughout. There are large effects of scanners on neural activation; indeed, even scanner upgrades can affect the BOLD signal. As a related point, it is important to keep acquisition parameters and sequence order consistent throughout the study. Altering repetition time or echo time can affect the stability of the signal. Further, ensuring a consistent protocol throughout the study is helpful, making sure that participants receive the same set of instructions at each timepoint and checking their comprehension prior to and after scanning. As with any fMRI study, it is important to instruct participants to refrain from drugs and alcohol, such as caffeine, nicotine, and melatonin.

A more nuanced consideration is the task presentation. Keeping the task consistent throughout is preferable; however task adaptation is a risk.

When measuring subcortical reward-related activation, it is recommended that researchers use single-band sequences. Recently, Srirangarajan and colleagues (2021) found that multi-band acquisition does not detect mesolimbic reward activation as well as single-band acquisitions.

Maintaining engagement in the task is also important. Personalizing the task by adjusting reaction time is one way to do this. Another way to keep engagement high is to incentivize good performance, especially with MID-like tasks.

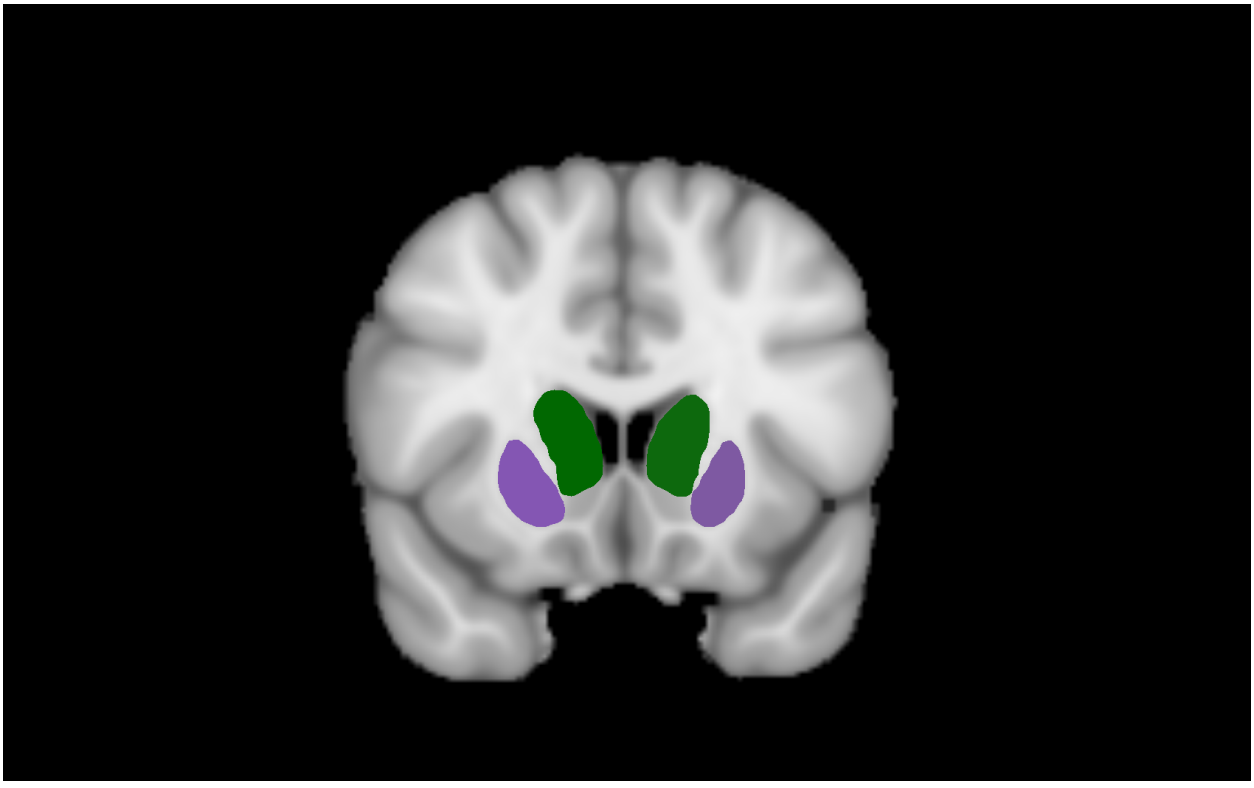
After data are collected, it is important to have stringent motion thresholds. Especially when scanning younger populations, stringent motion correction is imperative. Further, it is

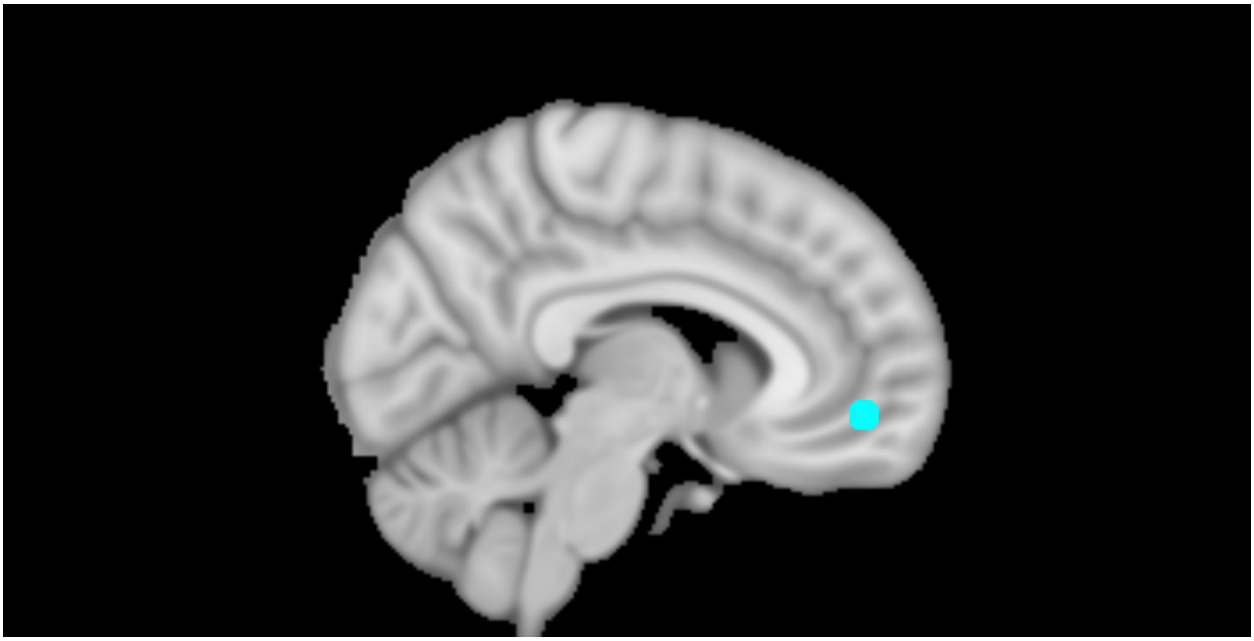
important to preprocess data consistently at each timepoint. Finally, the selection of contrasts is important. Contrasts of two active states (e.g., anticipation of reward vs. anticipation of loss) have less stability do contrasts of baseline activation or neutral conditions (e.g., anticipation of reward vs. anticipation of non-reward). Notably, there is an advantage of contrasting to neutral conditions, which yields better discriminate validity.

Reward-Related ROIs

We extracted the left and right putamen and left and right caudate using the Harvard-Oxford subcortical Atlas, thresholded at 25%. We averaged the left and right hemispheres in all analyses. In addition, we examined the bilateral medial prefrontal cortex (mPFC) using a 8-mm sphere with the center coordinates at x=±6, y=49, z=-8 (Wu et al., 2014). The mPFC was warped to MNI space using the icm2mni function (https://www.brainmap.org/icbm2tal/).

Figure 6S. *Additional Reward-Related Regions.*





Note: putamen (green), caudate (purple), and mPFC (teal).

Analogous to our procedure with the NAcc, we extracted trajectories of activation in these regions using a series of linear mixed effects models. We winsorized trajectories to minimize the impact of extreme cases. The stability of the putamen, caudate, and mPFC during the anticipation of reward vs. anticipation of non-reward was 0.13, 0.10 and 0.22, respectively (compared to the higher stability of the NAcc (0.32)).

Trajectories of activation in the putamen (β=-0.16, p=.031) were related to stress sensitivity, but not to objective ELS (putamen: p=.998). Similarly, trajectories of activation in the caudate were related to stress sensitivity (β=-0.17, p=.027), but not to objective ELS (p=.814). Trajectories of activation in the mPFC were not related to stress sensitivity (p=.120) or objective ELS (p=.872).

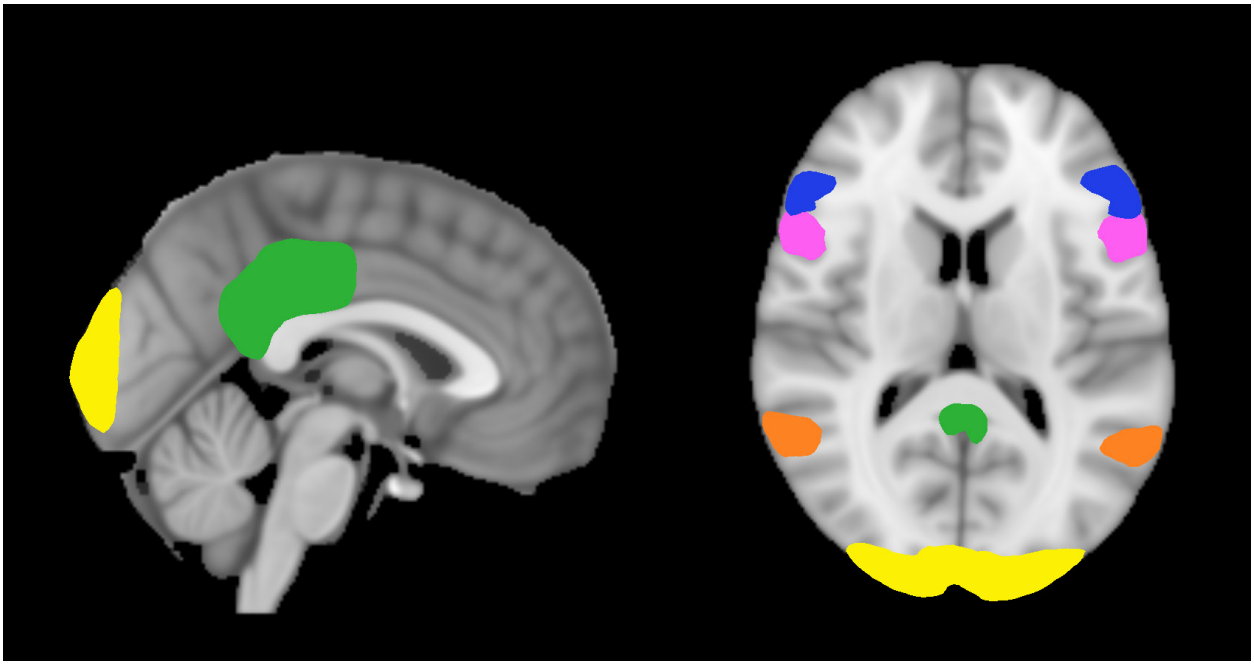
Trajectories of activation in the putamen interacted with biological sex to predict symptoms of psychopathology in adolescence (β=0.11, p<.001); decreasing activation of the putamen was related to externalizing (β=0.39, p<.001), but not internalizing problems (p=.306). Simple-slope analysis revealed a negative slope in boys (β=-0.25, p=.010), but not in girls

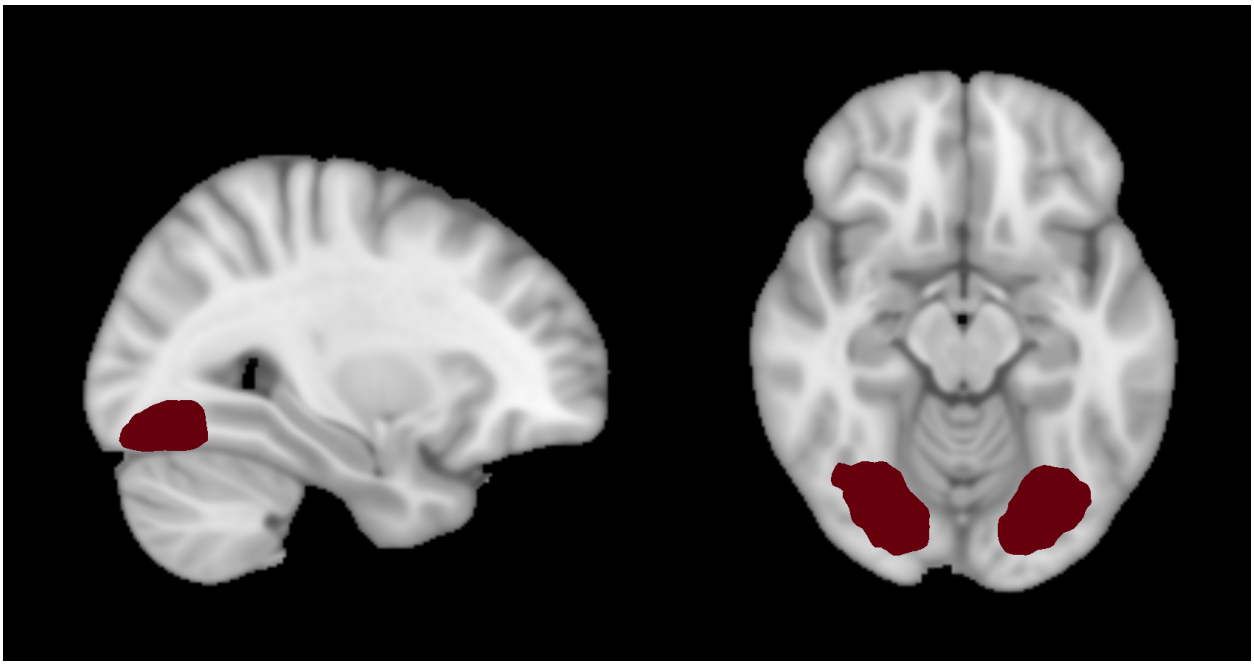
(*p*=.140). Trajectories of activation in the caudate interacted with biological sex to predict symptoms of psychopathology (β=0.09, *p*<.001). Specifically, activation in the caudate was related to externalizing (β=0.32, *p*=.019), but not internalizing problems (β=-0.16, *p*=.198; simple-slope analysis: boys β=-0.19, *p*=.070; girls *p*=.120). There was no main effect or interaction between trajectories of activation in the mPFC and biological sex with symptoms of psychopathology (main effect: β=0.01, *p*=.432; interaction: β=0.03, *p*=.074).

Control Regions

We examined 6 additional control regions to examine the specificity of our findings. Specifically, we extracted the bilateral angular gyrus, cingulate gyrus (posterior division), occipital fusiform gyrus, the occipital pole, inferior frontal gyrus (IFG) opercularis and triangularis. All control regions were extracted from the Harvard-Oxford structural Atlas, thresholded at 25%.

Figure 7S. *Control Regions.*





Note: occipital pole (yellow), angular gyrus (orange), IFG opercularis (pink), IFG triangularis (dark blue), cingulate gyrus (posterior division; green), and occipital fusiform gyrus (maroon).

Trajectories of activation were estimated in a series of linear mixed effects models. We winsorized trajectories to 3 standard deviations above or below the mean to minimize the impact of extreme cases. The stability of each control region is presented below:

Table 5S. *Stability of Control Regions.*

Variable	ICC
Angular gyrus	0.17
Cingulate gyrus	0.19
Occipital fusiform gyrus	0.00
Occipital pole	0.05
IFG opercularis	0.06
IFG triangularis	0.25

Control regions were not related to objective ELS (all *p*s>.200) or stress sensitivity (all *p*s>.121). There were no main effects or biological sex interactions during the anticipation of reward vs. anticipation of non-reward contrast in control regions when predicting symptoms of psychopathology. Specifically, the angular gyrus (main effect: β=0.03, p=.123; interaction: β=0.02, p=.203), cingulate gyrus (posterior division; main effect: β=0.01, p=.481; interaction: β=0.01, p=.376), IFG opercularis (main effect: β=0.00, p=.836; interaction: β=0.01, p=.577), IFG triangularis (main effect: β=0.03, p=.069; interaction: β=0.00, p=.806), occipital fusiform gyrus (main effect: β=0.03, p=.081; interaction: β=0.00, p=.907), and occipital pole (main effect: β=0.02, p=.240; interaction: β=0.00, p=.945) did not predict symptoms of psychopathology. Therefore, we did not conduct follow-up tests.

Finally, to ensure our findings in the NAcc were not driven by general neural activation, we conducted a MANOVA examining all control regions and trajectories of NAcc activation simultaneously. Importantly, the significant interaction between biological sex and activation in the trajectories of NAcc activation persisted even when considering control regions (β=0.06, p=.010).

References

1. King LS, Dennis EL, Humphreys KL, Thompson PM, Gotlib IH (2020): Cross-sectional and longitudinal associations of family income-to-needs ratio with cortical and subcortical brain volume in adolescent boys and girls. *Dev Cogn Neurosci* 44: 100796.

2. Marshall WA, Tanner JM (1968): Growth and Physiological Development During Adolescence. *Annu Rev Med* 19: 283–300.

3. Leong JK, Ho TC, Colich NL, Sisk L, Knutson B, Gotlib IH (2021): White-matter tract connecting anterior insula to nucleus accumbens predicts greater future motivation in adolescents. *Dev Cogn Neurosci* 47: 100881.

4. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, *et al.* (2004): Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 23: S208–S219.

5. Woolrich MW, Ripley BD, Brady M, Smith SM (2001): Temporal Autocorrelation in Univariate Linear Modeling of FMRI Data. *NeuroImage* 14: 1370–1386.

6. Siegel JS, Power JD, Dubis JW, Vogel AC, Church JA, Schlaggar BL, Petersen SE (2014): Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. *Hum Brain Mapp* 35: 1981–1996.

7. Telzer EH, McCormick EM, Peters S, Cosme D, Pfeifer JH, van Duijvenvoorde ACK (2018): Methodological considerations for developmental longitudinal fMRI research. *Dev Cogn Neurosci* 33: 149–160.

8. Knutson B, Greer SM (2008): Anticipatory affect: neural correlates and consequences for choice. *Philos Trans R Soc B Biol Sci* 363: 3771–3786.

9. Little R, Rubin D (1989): The Analysis of Social Science Data with Missing Values. *Sociol Methods Res* 18: 292–326.

10. Braams BR, Duijvenvoorde ACK van, Peper JS, Crone EA (2015): Longitudinal Changes in Adolescent Risk-Taking: A Comprehensive Study of Neural Responses to Rewards, Pubertal Development, and Risk-Taking Behavior. *J Neurosci* 35: 7226–7238.

11. Schreuders E, Braams BR, Blankenstein NE, Peper JS, Güroğlu B, Crone EA (2018): Contributions of Reward Sensitivity to Ventral Striatum Activity Across Adolescence and Early Adulthood. *Child Dev* 89: 797–810.

12. Insel C, Somerville LH (2018): Asymmetric neural tracking of gain and loss magnitude during adolescence. *Soc Cogn Affect Neurosci* 13: 785–796.

13. Yau W-YW, Zubieta J-K, Weiland BJ, Samudra PG, Zucker RA, Heitzeg MM (2012): Nucleus Accumbens Response to Incentive Stimuli Anticipation in Children of Alcoholics: Relationships with Precursive Behavioral Risk and Lifetime Alcohol Use. *J Neurosci* 32: 2544–2551.

14. Bjork JM, Chen G, Smith AR, Hommer DW (2010): Incentive-elicited mesolimbic activation and externalizing symptomatology in adolescents. *J Child Psychol Psychiatry* 51: 827–837.

15. Bach P, Reinhard I, Koopmann A, Bumb JM, Sommer WH, Vollstädt-Klein S, Kiefer F (2022): Test–retest reliability of neural alcohol cue-reactivity: Is there light at the end of the magnetic resonance imaging tube? *Addict Biol* 27: e13069.

16. Baranger DAA, Lindenmuth M, Nance M, Guyer AE, Keenan K, Hipwell AE, *et al.* (2021): The longitudinal stability of fMRI activation during reward processing in adolescents and young adults. *NeuroImage* 232: 117872.

17. Srirangarajan, T., Mortazavi, L., Bortolini, T., Moll, J., & Knutson, B. (2021). Multi-band FMRI compromises detection of mesolimbic reward responses. *NeuroImage*, *244*, 118617.

18. Wu, C. C., Samanez-Larkin, G. R., Katovich, K., & Knutson, B. (2014). Affective traits link to reliable neural markers of incentive anticipation. *NeuroImage*, 84, 279-289.