Supplementary Methods

Participants

Attrition

Of the 331 participants that completed fMRI at Wave 1 (163 HR and 168 LR), 244 participants (118 HR and 126 LR) completed fMRI at Wave 2 (Figure S1). There was no difference between participants who completed scanning at both waves compared to participants completing only the first scan in age, gender, risk group status, depressive symptoms, CTQ emotional neglect scores, or stressful life events assessed at Wave 1.

Exclusion of fMRI data for quality control

Of the 331 participants who underwent the first scan, 29 were initially excluded due to problems with the scan or raw data, including ending the scan early, no amygdala coverage, anatomical abnormalities, and artifacts, 7 were excluded because they had a diagnosis of major depressive disorder (MDD) before the first scan, and 5 participants in the low risk group were excluded because they had an anxiety disorder before the first scan. Data from the remaining 290 participants underwent pre-processing. Subsequently, 53 participants were excluded from the Time 1 data based on quality control criteria for the processed data (see quality control procedures section below for further details) and 5 participants were removed for not reaching the accuracy criterion on the behavioral task performed during scanning. Participants excluded at Time 1 had higher scores on the emotional neglect subscale of the CTQ (M=8.86, SD=4.4) compared to included participants (M=7.97, SD=3.0), $t(316)=2.08$, $p=.04$, but did not differ on any other measures of interest. Of the 244 participants who underwent the second scan, 9 were excluded due to problems with the scan or raw data, 6 were excluded because they had a diagnosis of MDD before the first scan, and 5 participants in the low risk group were excluded

because they had an anxiety disorder before the first scan. Of the 224 participants that were subsequently pre-processed, 21 participants were removed for quality control and 6 participants were excluded for accuracy. Participants excluded at Time 2 had higher scores on the emotional neglect subscale of the CTQ (M=9.4, SD=5.1) compared to included participants (M=8.01, SD=3.1), $t(232)=2.39$, $p=.02$.

Questionnaire measures of internalizing symptoms, childhood maltreatment, and life stress were obtained every year; for the present paper, only scores obtained at the same wave as neuroimaging data collection were used. Data for participants who completed questionnaires more than one month apart from completing fMRI scanning were excluded from analyses requiring data on depressive symptoms, resulting in the exclusion of an additional 81 participants for the analyses including depressive symptoms as a covariate.

At baseline, 32 participants in the high risk group had an anxiety disorder diagnosis, including specific phobia, social phobia, generalized anxiety disorder, panic disorder, obsessive compulsive disorder, separation anxiety disorder, and posttraumatic stress disorder.

Substance use

Participants were also asked about their use of substances at each wave of the study. Given the relatively young age range of the sample, response rates for the substance use questionnaire were generally low. For instance, only 11 participants reported ever having a drink in the past year at Wave 1, and of these, 4 participants reported having one drink 1-3 times over the past year, 2 reported having 1 or more drink more than 8 times in the past year, and the others did not indicate their drinking frequency. At Wave 2, forty-two participants reported having a drink in the past year and, of these, the majority (n=25) reported having one drink or less 1-3 times in the past 12 months. Thus, given the low response rates for substance use, we did not

examine these further as covariates. None of the subjects met DSM-IV criteria for alcohol use disorder (abuse or dependence) at either assessment.

Life stress procedure

Life stress severity was calculated by squaring the objective severity ratings, summing these ratings, and then dividing by the number of events reported. To illustrate why objective severity ratings are squared, take for example an adolescent that has 4 level 1 (little or no threat) stressors that include moving to a new house, starting a new school, changing grades, and starting a boyfriend/girlfriend relationship, all normal yet potentially stressful developmental adolescent experiences. Simply summing the total objective severity ratings for these would result in an overall score of 4. Contrast this with an adolescent whose parent suddenly dies and he/she finds them which would be a 4 (great threat). If that was their only event and we were simply summing objective severity scores, that adolescent would also get a 4. There are clearly very different stress levels in these two cases. To differentiate them, squaring the objective severity before summing results in the first adolescent still having a 4 and the second adolescent having a 16, more accurately reflecting the differences in stress levels between the two.

fMRI Procedure

fMRI paradigm

Blood oxygen level-dependent (BOLD) functional images were acquired on a 3T Siemens Trio Scanner using a gradient-echo echo planar imaging sequence (TR=2000 ms, TE=25 ms, field of view=20 cm, matrix=64x64, 34 slices, slice thickness=3 mm). The experimental fMRI paradigm consists of 4 blocks of a face-processing task interleaved with 5 blocks of a sensorimotor control task. Participant performance (accuracy and reaction time) is monitored during all scans using an MR-compatible button box. During task blocks, participants

view a trio of faces and select one of two faces (bottom) identical to a target face (top). All three of the faces show the same emotional expression (either angry or fearful); thus the participant's task is to match the identity of the target face to one of the two faces on the bottom row. In the TAOS version of the task, there are four task blocks with angry and fearful facial expressions derived from a standard set of pictures of facial affect (1). During the sensorimotor control blocks, participants perform the same target-matching task with simple geometric shapes (circles and ellipses). Each sensorimotor control block consists of six different shape trios. All blocks are preceded by a brief instruction ("Match faces" or "Match shapes") that lasts 2 seconds. In the task blocks, each of six face trios (three fearful and three angry) is presented for 4 seconds with a variable interstimulus interval (ISI) of 2 to 6 seconds (mean, 4 seconds), for a total block length of 48 seconds. A variable ISI is used to minimize expectancy effects and resulting habituation, and maximize amygdala reactivity throughout the paradigm. In the control blocks, each of the six shape trios is presented for 4 seconds with a fixed interstimulus interval of 2 seconds, for a total block length of 36 seconds. Total task length is 390 seconds.

Preprocessing and quality control procedures

Functional images were slice-timing corrected and then realigned to the first volume in the time series. Images were then normalized into standardized Montreal Neurological Institute space and smoothed with a 6 mm full width at half maximum Gaussian filter. The Artifact Detection Toolbox (ART; http://www.nitrc.org/projects/artifact_detect/) was used to identify images with excessive movement (>2 mm or degrees relative to the previous timeframe) or spiking artifacts (global mean intensity >4 standard deviations from the time series). Participants with >5% of functional images flagged for motion or artifact using ART were excluded from further analyses. For the remaining participants, ART generated regressors for volumes with high

motion or artifact in order to control for these volumes in analyses. A coverage check was performed using the anatomical amygdala region of interest; all participants meeting the ART criteria had coverage of >90% of voxels in the amygdala.

Movement during scanning

In addition to excluding participants who exhibited excessive movement during scanning, motion parameters from the realignment procedure were used to calculate mean head displacement, a summary metric of volume-to-volume translation, using the root-mean-square formula of Van Djik et al. (2): displacement = sqr-rt($x^2 + y^2 + z^2$), expressed in mm. Mean head displacement correlated with age at the second wave of scanning (r=-.18, p=.01). Thus, as an additional control for movement, linear mixed models with neuroimaging measures as the dependent variables were re-run with mean head displacement entered as a time-varying covariate; the addition of group-level motion covariates has been demonstrated to help mitigate the influence of age-related movement in pediatric neuroimaging studies (3).

Extraction of Contrast Values

As in prior research (4, 5), contrast values were extracted from functional clusters within anatomically defined (Automated Anatomical Labeling atlas) amygdala regions of interest (ROIs) exhibiting significant main effects of task (i.e., faces>shapes) at p<.05 family-wise error corrected across the volumes of the ROIs. In order to ensure that BOLD parameter estimates were extracted from identical clusters for each wave, a conjunction analysis was performed to identify all overlapping suprathreshold voxels at the first and second wave. For the contrast of all faces>shapes, the left amygdala functional cluster contained 170 voxels and the right amygdala functional cluster contained 227 voxels. BOLD parameter estimates were extracted for the following contrasts: fearful faces>shapes and angry faces>shapes. To detect outliers in the

extracted parameter values, we examined box and whisker plots in SPSS, which identify any values >3 times the interquartile range as extreme values. Using this approach, one outlier was identified as evidencing extreme values of amygdala reactivity and was removed from subsequent analyses to avoid biasing results, although results were similar regardless of inclusion or exclusion of this outlier.

Statistical Analyses

Linear mixed models

Recommended procedures of Tabachnik and Fidell (6) were followed in testing the linear mixed models. Model fit was tested by comparing models to a null model with no parameters except a random intercept. Chi-square values were calculated using the difference in -2 Log Likelihood scores. In order to test the hypothesis that high risk adolescents would differ from low risk adolescents in development of amygdala reactivity, risk group, age, and an age-by-risk group interaction were entered as parameters in a linear mixed model using the mixed procedure of SPSSv21. A significant age-by-risk group interaction indicates that changes in amygdala reactivity with age differ between the two groups. Continuous predictors were centered before entering these into the model, and maximum likelihood estimation was chosen in order to perform model comparisons (6). Gender was entered as a covariate in all analyses. Model fitting procedures were used to determine whether including the quadratic effect of age improved model fit. When this additional parameter did not significantly improve model fit, the more parsimonious model with fewer parameters is reported in the results. The dependent variable was left or right amygdala reactivity extracted from the functionally defined cluster. We examined left or right amygdala reactivity separately and we extracted amygdala reactivity to fearful and angry face expressions separately. Age and stressful life events were modeled as both fixed and

random effects, unless the model failed to converge, in which case random effects were removed. In the final models, age was modeled as a fixed and random effect for the first hypothesis and both age and stressful life events were modeled as fixed and random effects for the second hypothesis. Note that although family was entered as a level 3 variable for the first hypothesis, the final Hessian matrix was not positive definite when running the linear mixed models to examine effects of life stress. Thus, this third level was removed from these analyses, and only the level 2 variable of participant was included when testing the second hypothesis.

Hypothesis 1

Because this procedure can incorporate participants missing data at one wave, this analysis included all participants with available fMRI data at either wave. Four dependent variables were tested (left or right amygdala reactivity to fearful or angry faces); thus, the significance level for this hypothesis was set at a Bonferroni-corrected $p \leq .0125$ for the F-test of the interaction. Linear mixed models were re-run with the following additional covariates entered as controls: mean head displacement, mean accuracy, and mean RT on the task. We also ran analyses excluding any participants with an internalizing disorder diagnosis, as well as controlled for depressive symptoms at each wave.

Hypothesis 2

Because all effects were tested within one model, no Bonferroni correction was applied. We included the following additional covariates in the linear mixed model analysis for life stress: mean recency of the events (mean time between when the events occurred relative to scanning) and time between the baseline interview and scanning session.

Supplementary Results

Excluding participants with depression or anxiety

When excluding participants with depression or anxiety, the group difference in depressive symptoms at Wave 1 was no longer significant, $t(193)=-.51$, $p=.61$, indicating that in this reduced sample the high and low risk groups had roughly equivalent levels of depressive symptoms at baseline (High risk M=8.7, SD=8.4; Low risk M=8.1, SD=6.4).

Examining differences at baseline and changes between waves within the longitudinal subset

The aggregate effect of age includes both between-person variations in age as well as within-person change with age; thus, we performed follow-up analyses to specifically assess change in amygdala reactivity between waves in the subset of participants that had fMRI data at both waves of scanning. Group differences in amygdala reactivity at Wave 1, differences in residualized change in amygdala reactivity, and differences at Wave 2 were examined. This analysis was conducted using a general linear model in SPSSv21, with age and gender entered as covariates. Residualized change is calculated by computing the difference between observed scores at Wave 2 and predicted scores for Wave 2. Predicted scores for Wave 2 were obtained by conducting a multiple regression with Wave 1 scores as a predictor. Thus, residualized change scores carry the advantage of controlling for the influence of baseline scores on change over time (e.g., regression to the mean), and can be interpreted as a change greater or lesser than expected (7, 8). The residualized change score for left amygdala reactivity to fearful faces is highly correlated with the simple change score (i.e, subtracting Time 1 from Time 2), $r=.82$, $p<.001$.

There were no differences in left amygdala reactivity to fearful faces at baseline (Wave 1) between the groups, $F(1, 152)=.17$, $p=.68$, change $R^2=.001$. However, the high-risk group

evidenced greater than expected change in amygdala reactivity between the waves, $F(1, 152)=5.715$, $p=.018$, change $R^2=.04$, supporting the effect of risk group status on within-person change in amygdala reactivity across development. As a result, at Wave 2 there was a significant difference in amygdala reactivity between the groups, $F(1,152)=5.721$, $p=.018$, change $R^2=.04$. These results suggest that, controlling for age of entry to the study, differences between the risk groups were generally smaller in early adolescence and grew stronger over time.

Moderating effects of gender

Although gender was controlled in the main analyses, we performed a supplementary set of analyses examining moderating effects of gender. For hypothesis 1, including gender as a moderator in the linear mixed model for left amygdala reactivity to fearful faces did not further increase the fit, $\chi^2(3, N=427)=5.14$, $p>.05$. For hypothesis 2, we examined whether gender further moderated the SLES-by-risk group interaction reported in the main results. There was a significant risk group-by-gender-by-stress interaction, $F(1,69)=4.34$, $p=.04$, suggesting the risk group-by-stress interaction was stronger for boys than for girls.

Effects of life stress severity reported at Wave 1 on change in amygdala reactivity

We followed up our significant effects from the linear mixed models by examining the effects of life stress severity at Wave 1 in the subset of longitudinal participants. This was conducted using the PROCESS macro for SPSS (9). We examined the interaction of risk group and life stress severity including age, gender, time between the baseline interview and the first scanning session, time between Scan 1 and Scan 2, and mean recency of events as covariates. The full model was significant, $F(8,145)=2.89$, $p=.005$. The interaction of risk group-by-life stress severity was significant, B=-.10, SE=.04, $t(152)=-2.34$, $p=.02$, R^2 change=.03. Similar to what was found in the linear mixed models, this effect was driven by low risk participants

evidencing decreases in amygdala reactivity under mild levels of life stress and increases in amygdala reactivity under more severe life stress (Figure S3). We did not find a significant moderating effect of gender in this model.

For completeness, we also examined these effects with the other stress measures (child neglect and stress assessed at Wave 2) and the other amygdala variables (right amygdala reactivity to fear, left and right amygdala reactivity to angry). For left amygdala reactivity to fear, we did not see significant effects with the other stress measures. For right amygdala reactivity to fear, the overall model for stress at Wave 1 including covariates approached significance (p=.06), and the risk group-by-stress severity interaction was significant, B=-.12, SE=.04, t(152)=-2.93, p=.004. Results were similar to left amygdala reactivity to fear, with the low risk group evidencing greater than expected changes in reactivity with more severe life stress. Models including covariates were not significant for amygdala reactivity to angry faces (p's>.3).

The effects of different features of life stress on change in amygdala reactivity

We performed a number of post hoc analyses using linear mixed models to determine whether any additional features of stressful life events (besides objective severity) were associated with change in amygdala reactivity with age. Because effects were strongest for stressful life events reported at Wave 1, we focused post hoc analyses on this measure. The first feature examined was number of events, including a count of all events reported, as well as a count by type of event including the following categories: deaths, education and work (i.e., problems at school or work; these two categories were combined), problems with housing or money (i.e., financial difficulties or unstable housing; these two categories were combined), health (i.e., health problems in the participant or a close friend or family member), other

relationships (i.e., problems associated with friends or family such as parental divorce), and romantic relationships (e.g., a breakup). The second feature was the recency of the event, measured as the number of months between when the event occurred and the scan (this was also entered as a covariate in the main analyses and was not significant). Thus, a mean recency score of 4 would indicate that on average events had occurred within 4 months of scanning. The third feature was whether the event was independent of the participant's actions or dependent on the adolescent's behavior; each event was scored on a scale from 1 (totally independent) to 4 (definitely dependent). The fourth feature was the focus of the event, scored on a scale from 1 (participant was the focus of the event) to 4 (pet/possession was the focus of the event). Means and standard deviations for these variables are reported in Supplementary Table 3.

We then proceeded to conduct linear mixed models as described for the main analyses using the aforementioned variables in place of the objective severity of life events. The interaction between total number of life events and age approached significance, $F(1,170)=3.87$, $p=.051$, such that in both groups more life events were associated with greater amygdala reactivity with age. When breaking down life stress by event type, there was a main effect of total deaths, $F(1,266)=5.29$, $p=.02$, indicating that deaths were associated with increased amygdala reactivity at all ages and within both groups. The number of housing and money events interacted with age, $F(1,163)=6.71$, $p=.01$, such that participants in both risk groups who experienced more of these events evidenced increases in amygdala reactivity with age. There was a three-way interaction between the number of health-related events, age, and risk group, $F(1, 399)=8.84$, $p=.003$, such that more health-related events were associated with increases in amygdala reactivity specifically in the low-risk group. Finally, romantic relationship events interacted with age, $F(1, 237)=4.52$, $p=.03$, and risk group, $F(1,284)=8.94$, $p=.003$, such that

more romantic relationship events were associated with decreased amygdala reactivity with age in the low risk group (note that these events can include things like starting a new relationship), but not in the high risk group, which showed increases with age regardless of the number of romantic relationship events reported.

There was a main effect of the mean independence of events, $F(1,55)=7.12$, $p=.01$, indicating that greater mean independence of events was associated with greater amygdala reactivity. There were no significant effects using the measures of recency or self vs. other focus. These results suggest that various types of life events are associated with changes in amygdala reactivity. Given the strong effects observed for objectively-rated life event severity reported in our *a priori* analyses, this suggests that the severity of the life event is an important factor in determining changes in amygdala reactivity, and that more severe events across a range of stressors (e.g., deaths, health problems, housing or money problems) may all potentially be associated with increased amygdala reactivity with age. It is important to note, however, that these post hoc analyses are exploratory and should be examined further in future research.

References

1. Ekman P, Friesen WV. Pictures of Facial Affect. Palo Alto, CA: Consulting Psychologists Press; 1976.

2. Van Dijk KRA, Sabuncu MR, Buckner RL. The influence of head motion on intrinsic functional connectivity MRI. NeuroImage 2012; 59:431-8.

3. Satterthwaite TD, Wolf DH, Ruparel K, Erus G, Elliott MA, Eickhoff SB, et al. Heterogeneous impact of motion on fundamental patterns of developmental changes in functional connectivity during youth. NeruoImage 2013; 83:45-57.

4. White MG, Bogdan R, Fisher PM, Munoz KE, Williamson DE, Hariri AR. FKBP5 and emotional neglect interact to predict individual differences in amygdala reactivity. Genes Brain Behav 2012; 11:869-78.

5. Bogdan R, Williamson DE, Hariri AR. Mineralocorticoid receptor iso/val (rs5522) genotype moderates the association between previous childhood emotional neglect and amygdala reactivity. Am J Psychiatry 2012; 169:515-522.

6. Tabachnik BG, Fidell LS. Multilevel Linear Modeling, in Using Multivariate Statistics. 5th ed. New York: Pearson; 2006.

7. Whittle S, Simmons JG, Dennison M, Vijayakumar N, Schwartz O, Yap MBH, et al: Positive parenting predicts the development of adolescent brain structure: A longitudinal study. Dev Cogn Neurosci 2014; In press.

8. MacKinnon DP: Longitudinal mediation models. Introduction to Statistical Mediation Analysis. New York: Taylor & Francis; 2008. p. 193-236.

9. Hayes AF: Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach. New York: Guilford Press; 2013.

TABLE S1. Bivariate Correlations Between Variables at Each Wave[a]

	Right Amyg Fear T1	Left Amyg Anger T1	Right Amyg Anger T1	Left Amyg Fear T2	Right Amyg Fear T2	Left Amyg Anger T2	Right Amyg Anger T2	MFQ T1	MFQ T2	CTQ Neglect	SLES T1	SLES T2	Age T1	Age T2
Left Amyg Fear T1	r=.44 p<.001	r=.41 p<.001	r=.11 p=.11	r=.18 p=.03	r=.12 p=.13	r=.08 p=.35	r=.08 p=.34	r=-.01 p=.84	r=.07 p=.36	r=-.02 p=.83	r=-.03 p=.61	r=.02 p=.75	r=-.07 p=.29	r=-.04 p=.65
Right Amyg Fear T1		r=.24 p<.001	r=.47 p<.001	r=.10 p=.20	r=.09 p=.27	r=.02 p=.79	r=.09 p=.29	r=-.06 p=.34	r=-.04 p=.58	r=.04 p=.60	r=.08 p=.25	r=-.02 p=.82	r=-.10 p=.13	r=-.10 p=.20
Left Amyg Anger T1			r=.45 p<.001	r=.23 p=.003	r=.16 p=.048	r=.12 p=.14	r=.01 p=.91	r=-.03 p=.64	r=.02 p=.83	r=-.01 p=.87	r=.08 p=.25	r=-.03 p=.70	r=-.13 p=.04	r=-.17 p=.03
Right Amyg Anger T1				r=.17 p=.03	r=.16 p=.05	r=.04 p=.61	r=.08 p=.35	r=-.07 p=.28	r=-.04 p=.55	r=-.09 p=.21	r=.04 p=.59	r=-.03 p=.65	r=-.14 p=.03	r=-.19 p=.02
Left Amyg Fear T2					r=.48 p<.001	r=.45 p<.001	r=.23 p=.001	r=.02 p=.80	r=-.14 p=.06	r=-.02 p=.74	r=.10 p=.16	r=.07 p=.37	r=-.06 p=.47	r=-.01 p=.87
Right Amyg Fear T2						r=.22 p=.002	r=.43 p<.001	r=.03 p=.73	r=.02 p=.84	r=-.01 p=.91	r=.08 p=.26	r=.15 p=.04	r=-.03 p=.70	r=-.03 p=.66
Left Amyg Anger T2							r=.54 p<.001	r=.09 p=.19	r=-.15 p=.03	r=.005 p=.95	r=.08 p=.29	r=.18 p=.02	r=-.03 p=.73	r=-.001 p=.99
Right Amyg Anger T2								r=-.001 p=.99	r=-.02 p=.82	r=-.07 p=.34	r=.04 p=.57	r=.13 p=.07	r=-.07 p=.40	r=-.01 p=.91
MFQ T1									r=.36 p<.001	r=.20 p=.001	r=.15 p=.02	r=.22 p=.001	r=.06 p=.39	r=-.04 p=.57
MFQ T2										r=.14 p=.045	r=.02 p=.81	r=.23 p=.001	r=-.07 p=.36	r=.01 p=.85
CTQ Neglect											r=-.02 p=.79	r=.06 p=.39	r=-.02 p=.77	r=-.04 p=.58
SLES T1												r=.17 p=.02	r=-.04 p=.49	r=.-10 p=.19
SLES T2													r=.01 p=.87	r=.06 p=.44
Age T1														r=.93 p<.001

[a] MFQ = Total scores on the Mood and Feelings Questionnaire; CTQ Neglect=Scores on the Emotional Neglect subscale of the Childhood Trauma Questionnaire; SLES= Stressful Life Events Schedule objective ratings.

TABLE S2. Whole-Brain Results for Risk Group-by-Wave Interaction[a]

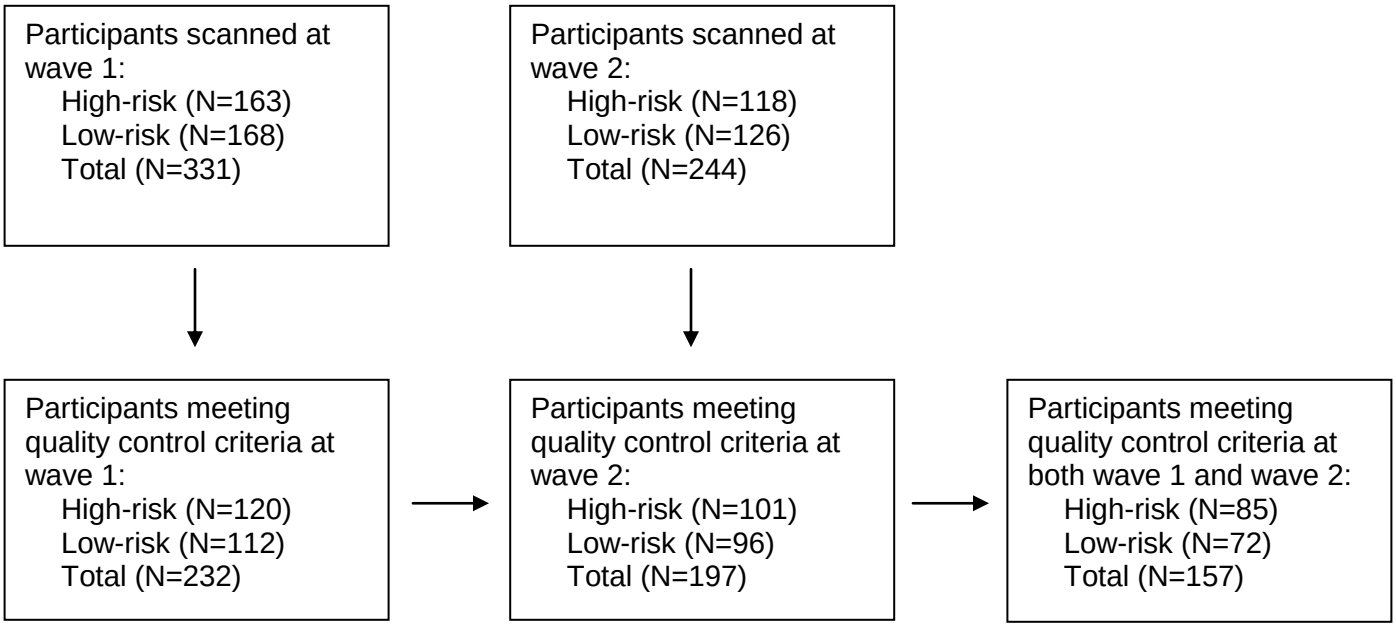
Contrast	Statistic (*p*'s <.005)	Cluster Size	Coordinates (x, y, z)	Region	Direction of Effect
Fearful faces vs. shapes	*F*(1,153)=22.41	120	20, -74, -14	Right lingual gyrus (BA 18)	Increases in HR group; Decreases in LR group
	F(1,153)=14.68	69	-16, -80, -12	Left lingual gyrus (BA 18)	Decreases in LR group
	F(1,153)=14.43	20	-52, -12, -20	Left temporal lobe	Decreases in HR group
	F(1,153)=13.72	81	32, -88, 8	Right middle occipital gyrus	Increases in LR group
	F(1,153)=13.43	28	52, -42, -10	Right temporal lobe	Decreases in HR group
	F(1,153)=13.21	21	-36, -88, 0	Left middle occipital gyrus	Decreases in HR group; Increases in LR group
	F(1,153)=12.56	40	-30, -70, 32	Left occipital lobe	Decreases in HR group; Increases in LR group
	F(1,153)=11.42	21	6, 46, 20	Right anterior cingulate (BA 9)	Increases in LR group
Angry faces vs. shapes	*F*(1,153)=18.97	32	-38, -30, -18	Left parahippocampal gyrus (BA 36)	Decreases in HR group
	F(1,153)=17.29	84	48, -62, -6	Right inferior temporal gyrus	Decreases in HR group
	F(1,153)=15.47	28	-58, -6, -20	Left inferior temporal gyrus (BA 21)	Decreases in LR group
	F(1,153)=12.28	32	4, -52, -26	Right cerebellum	Decreases in LR group
	F(1,153)=12.25	25	22, -62, -28	Right cerebellum	Increases in HR group; Decreases in LR group
	F(1,153)=12.05	35	50, -10, 6	Precentral gyrus (BA 6)	Decreases in HR group
	F(1,153)=11.13	28	22, -74, -14	Right lingual gyrus (BA 18)	Increases in HR group; Decreases in LR group
	F(1,153)=10.03	44	4, 52, -8	Medial frontal gyrus (BA 10)	Increases in LR group

[a] Whole-brain results were examined with a flexible factorial ANOVA in SPM with participants who had data available at both waves of scanning (n=156), and with age entered as a covariate. A risk group-by-wave interaction was evaluated. Whole-brain results were evaluated at $p<.005$ uncorrected, with minimum cluster size of 20. The Wake Forest University Pickatlas was used to determine the location of each set of coordinates. Direction of effect indicates which group is driving the interaction. BA=Brodmann Area; HR=High Risk; LR=Low Risk.

TABLE S3. Characteristics of Stressful Life Events Reported at Baseline

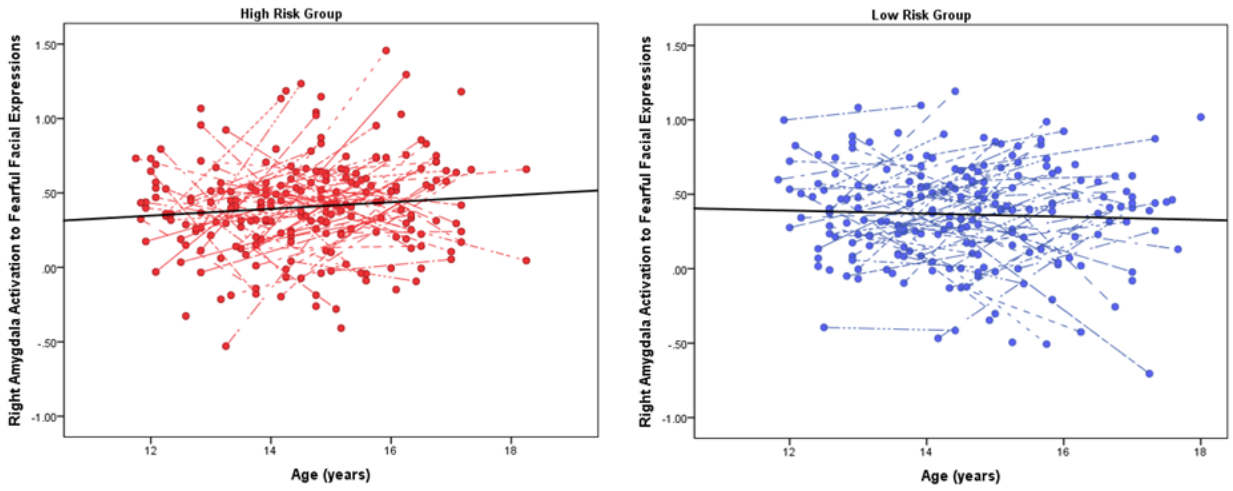
	High-Risk Group		Low-Risk Group	
	M	**SD**	**M**	**SD**
Total number of events	5.4	3.27	4.9	3.21
Number of deaths	.41	.63	.39	.64
Number of school/work events	1.47	1.18	1.39	1.10
Number of housing/money events	.40	.64	.43	.73
Number of health events	.99	.91	.96	1.04
Number of other relationship events	1.39	1.53	.97	1.18
Number of romantic relationship events	.55	1.06	.49	1.03
Mean recency (months)	4.62	2.4	5.12	2.8
Mean independence	2.14	.80	2.19	.82
Mean self vs. other focus	2.05	.48	2.02	.54

FIGURE S1. Participant Inclusion/Exclusion Procedure[a]



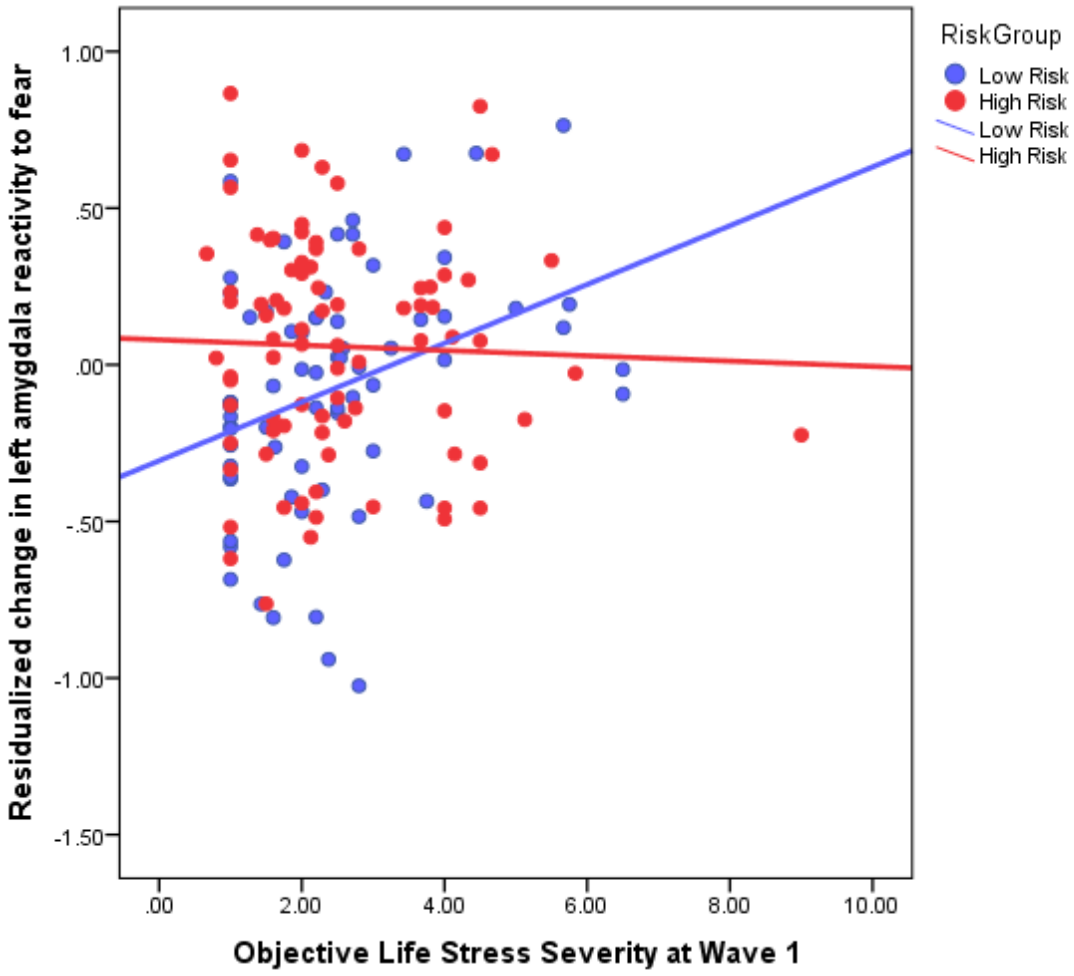
[a] Participants undergoing fMRI were excluded based on quality control criteria, including problems with raw data, excessive motion or artifact, amygdala coverage <90%, and task accuracy <70%. For analyses using questionnaire measures, participants were excluded if they completed the questionnaires >1 month before scanning.

FIGURE S2. Risk Group-by-Age Interaction on Right Amgydala Reactivity to Fearful Facial Expressions[a]



[a] The risk group-by-age interaction for right amygdala reactivity to fearful faces was not significant ($F(1,177)=2.10$, $p=.15$), but in the same direction as the effect for left amygdala reactivity.

FIGURE S3. Risk Group-by-Stress Interaction Predicts Residualized Change in Left Amygdala Reactivity to Fearful Faces[a]



[a] Risk group and objective life stress severity reported at Wave 1 interact to predict residualized change in left amygdala reactivity to fearful faces, B=–.10, SE=.04, $t(152)$=-2.34, p=.02. As seen in the scatterplot, greater life event severity is associated with greater than expected increases in left amygdala reactivity between waves for the low-risk group.