Separable Effects of Childhood Maltreatment and Adult Adaptive Functioning on Amygdala Connectivity During Emotion Processing

Supplement 1

Participant Exclusion

Fifty-three individuals from the longitudinal study were unable to participate for the following reasons: size/obesity (19), metal in body (6), incarceration (11), refusal or scheduling conflicts (14), pregnancy (1), or deceased (2). Other exclusion criteria for MRI scanning included current or past history of neurological disorders or trauma, known intellectual impairment, and uncorrected visual or auditory impairments.

Description of Developmental Task Score

The developmental task scores range from 0 through 14, and were generated through a composite of rank scores (0: stalling, 1: maintaining, 2: succeeding) on seven domains of functioning. Participants were ranked in one of three categories for each domain based on their success on the developmental task relative to other participants in the study. This approach was taken to emulate work by Schulenberg et al (2004; 1). Information from each domain was drawn from the Adult Self Report measure (ASR; 2) and a demographics questionnaire. Thus, this score includes both objectively verifiable components (e.g., education attainment, annual income) and subjective components (e.g., friendship quality, family involvement).

For the education domain, 25 individuals did not finish high school (however 13 of those received their GED) and were categorized as stalling, 19 graduated from high school and were categorized as maintaining, and 36 pursued further education (21 earned a vocational technical

diploma or completed part of a collegiate program, 9 earned an associate's degree, 4 earned a bachelor's degree, and 2 earned a master's degree).

Success in work was based on occupational standing according to the Hauser and Warren Socioeconomic Index (SEI) score that considers earnings, education, and prestige associated with occupations (2), and the adaptive functioning job score on the ASR. Scores for the participants' current work and usual work were averaged to create one score. Eighteen people were categorized as stalling in this domain, including individuals who were currently unemployed or disabled. Individuals who reported that they were keeping house or in school, or held a job of mediocre occupational standing (e.g., maid, janitor, construction laborer, kitchen worker), or an adaptive functioning job score of < 1.5 (low job satisfaction and confidence) were considered maintaining in this domain. This group contained 44 individuals. Finally, 18 participants who had a relatively high SEI score (e.g., health aide, teacher or teacher's aide, general office clerk, sales worker) and an adaptive functioning job score greater than 1.5 (medium-high job satisfaction and confidence) were considered succeeding occupationally in this group.

Financial autonomy was based on total family income rank within this sample. The range of family income levels was divided into approximate thirds. Twenty-seven individuals were in the stalling category, which included those earning less than $20k/year. Thirty-three individuals' family income was between $20-40k and were in the maintaining category. Lastly, 20 individuals were in the succeeding category with family earnings of $40-120k. Based on the 2013 Federal Poverty Guidelines, the poverty line is defined as household income of less than $23.5k/year for a family of four (3).

Ranking of success in the romantic involvement domain differed from rankings by Schulenberg and colleagues (2004) to reflect the average age of marriage in New York state (28

years of age, as opposed to 26 years, which was used in Schulenberg's ranking). Unmarried and non-cohabiting individuals who were 28 years old or younger were classified as maintaining. Otherwise, rankings were based on marital status, divorce history, and relationship ratings given on the ASR. To be classified as stalling, individuals had to have been divorced more than twice, single and not cohabiting, or in a low-quality marriage (ASR adaptive functioning Spouse/Partner score < 1). This group contained 25 individuals. The maintaining group, which contained 37 individuals, included divorced but remarried participants, unmarried but cohabiting participants, and married but unsatisfied participants (ASR adaptive functioning spouse/partner score = 1-1.5). Eighteen individuals were classified as succeeding in the romantic involvement domain, which included individuals who had never been divorced and were currently in a high-quality marriage (ASR adaptive functioning Spouse/Partner score > 1.5).

For the peer involvement domain, ranking was based on the ASR adaptive functioning friends scale. This scale encompasses quantity of friendships and contact as well as quality of friendships. Twenty-six participants were stalling (score < 1.75), 24 participants were maintaining (score = 1.75 - 2.25), and 30 participants were succeeding (score > 2.25) in this domain.

Family involvement rankings were also based on the ASR report, using the adaptive functioning family scale, which indexes how well one gets along with family members. These scores were averaged across family members that participants reported having contact with (including parents, siblings, and children), as it may actually be adaptive to not have contact with some family members, particularly if maltreatment was perpetrated by a family member. Thirty participants were categorized as stalling (score < 1.25), while 21 were maintaining (score = 1.25 - 1.75) and 29 were succeeding (score > 1.75).

The last developmental task domain indexed in this sample was related to substance abuse. Rankings were based on ASR Substance Use Scales for tobacco, alcohol, and drugs. Scores on these three subscales (ranging from 50 to 100) were averaged. The sample was nearly evenly divided into thirds, with 27 individuals ranked as stalling (score > 66.67), 27 ranked as maintaining (score = 50 - 66.67), and 26 ranked as succeeding (score = 50).

MRI Acquisition

Structural and functional MRI data were acquired on a Siemens 3-Tesla Trio scanner using a 32-channel head coil. High-resolution, T1 weighted images were acquired for each participant using an MPRAGE sequence (echo time [TE] = 3.44 ms, repetition time [TR] = 2530 ms, field of view = 256 mm, matrix = 256x256, slice thickness = 1 mm, flip angle = 7°, 192 sagittal slices). Functional data were acquired using an echo-planar imaging sequence (TE = 30 ms, TR = 2000 ms, field of view = 224 mm, matrix = 64x64, slice thickness = 3.5 mm with a 29% gap, flip angle = 90°, 30 interleaved oblique axial slices). To correct geometric distortion in the functional data, a fieldmap volume was collected immediately prior to the functional data acquisition using the same slice prescription (TE1 = 5.19 ms, TE2 = 7.65 ms, TR = 400 ms, field of view = 224 mm, matrix = 64x64, slice thickness = 3.5 mm with a 29% gap, flip angle = 60°, 30 interleaved oblique axial slices).

Behavioral Results

Behavioral accuracy and response time were examined using separate 2 x 2 mixed model analyses of variance with condition (shape-matching, emotion-matching) as a within-subjects factor and group (maltreated, comparison) as a between-subjects factor. Trials without a response,

due to either failure to respond or a late response outside of the allotted time window, were coded as inaccurate. For accuracy, there were significant main effects of both condition, $F(1, 78) = 180.95$, $p < .001$, and group, $F(1,78) = 6.80$, $p = .01$, as well as a significant interaction effect, $F(1,78) = 4.04$, $p = .05$. All participants were more accurate in shape matching ($M = 96.63\%$, $SD = 4.96\%$) than in emotion matching ($M = 65.21\%$, $SD = 20.84\%$), $t(79) = 13.71$, $p < .001$. The maltreated group had lower accuracy than the comparison group for emotion matching only ($M_{mal} = 59.97\%$, $SD_{mal} = 21.06\%$, $M_{comp} = 70.73\%$, $SD_{comp} = 19.37\%$), $t(78) = 2.69$, $p < .01$. For response time, there was a significant main effect of condition, $F(1,78) = 630.898$, $p < .001$. Participants were faster to respond when shape matching ($M = 1341.98$ ms, $SD = 360.02$ ms) than when emotion matching ($M = 2550.28$ ms, $SD = 445.48$ ms), $t(79) = -25.28$, $p < .001$. The interaction effect and the main effect of group were not statistically significant for response time.

Resilience and Behavior

Adult adaptive functioning, as measured by developmental task scores, did not differ between groups (CM and non-CM), $t(78) = 1.19$, $p = .24$ ($M_{mal} = 6.61$, $SD_{mal} = 3.06$, $M_{comp} = 7.38$, $SD_{comp} = 2.75$). Resilience also did not relate to accuracy on shape matching trials, $r(79) = -.04$, $p = .70$, accuracy on emotion matching trials, $r(79) = .04$, $p = .72$, response time for shape matching, $r(79) = -.08$, $p = .50$, or response time for emotion matching trials, $r(79) = -.09$, $p = .43$. Resilience was positively associated with age, $r(79) = .31$, $p < .01$, and differed by sex, $t(78) = -1.99$, $p = .05$, with females showing higher developmental task scores than males ($M_{female} = 7.64$, $SD_{female} = 2.90$, $M_{male} = 6.37$, $SD_{male} = 2.84$). Accordingly, all PPI analyses included demeaned age and sex as covariates.

Task Reactivity: Effect of Adaptive Functioning, Controlling for Group

For the emotion-matching > shape-matching contrast, there was one region that was related to adaptive functioning, controlling for group. Activation in the left inferior frontal gyrus (3256mm^3, MNI: -50, 26, 18) was positively correlated with the adult adaptive functioning score.

Internalizing Symptoms and PPI

All participants completed the Beck Depression Inventory, Second Edition (BDI-II; 4) and the Achenbach Self-Report (ASR) in order to provide information on current mood and anxiety disorders. Measures did not differ by group (CM vs. non-CM): BDI-II $t(78) = -0.009$, $p = .993$; ASR Internalizing *t*-score: $t(78) = 0.324$, $p = .747$. As the developmental task score and these internalizing measures are from the same time point and are conceptually related, we examined their statistical relationship. The BDI-II score distribution was skewed because of the high proportion of zero-scores. Therefore we focused on the ASR score, which was related to the developmental task score: $r(78) = -.263$, $p = .019$.

Next, to ensure that CM and/or adult adaptive functioning's effect on amygdala connectivity was not entirely due to adult mental health, we ran an analysis in FSL including an additional nuisance variable: centered ASR internalizing score. Results revealed the same brain regions showed connectivity with the amygdala, but that cluster size changed. The observed change in cluster size may be due to the decreased power of the model with an additional nuisance variable and the collinearity between the Developmental Task Score and internalizing measure.

After including internalizing in the model, the resulting hippocampus region with significant findings was 95% of the size of the significant region when internalizing was not included. The size of the cingulate and DMPFC clusters was reduced by about half, whereas the

parietal regions actually increased in size. As adult functioning and internalizing measures were collected concurrently and are conceptually and statistically overlapping, $r(78) = -.263$, $p = .019$, this change in cluster size is likely an artifact of the shared variance. The consistency of regions that showed amygdala-based emotion-related PPI connectivity after the inclusion of internalizing symptoms demonstrates that these effects are not merely a reflection of mental health status.

PPI Effects Using a More Stringent Cluster Forming Threshold

While there is no standard for selecting a cluster forming threshold in the neuroimaging community, a recent paper has recommended using $p < .001$ as a voxel-wise threshold in order to reduce potential Type I error (5). When applying this threshold to our PPI analyses, group differences in amygdala-based connectivity show a region within the left hippocampus (1456 mm^3, MNI: -36, -22, -26), but no regions survive when controlling for adaptive functioning. One region survives this threshold for the effect of developmental task on functional connectivity (when controlling for group) with the amygdala. Specifically, PPI signal in the dorsomedial PFC (2520 mm^3, MNI: 8, 28, 46) remains negatively associated with adaptive functioning at this threshold. However, we caution that it is possible that use of this threshold inflates Type II error (6-9). The use of $p < .005$ as a voxel-wise threshold is standard practice in our laboratory, specifically because it strikes a balance between Type I and Type II errors.

Supplemental References

1. Schulenberg JE, Bryant AL, O'Malley PM (2004): Taking hold of some kind of life: How developmental tasks relate to trajectories of well-being during the transition to adulthood. *Dev Psychopathol*. 16:1119-1140.
2. Achenbach TM, Rescorla LA (2003): *Manual for ASEBA adult forms and profiles.* Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.
3. US Department of Health and Human Services (2013): Poverty guidelines. *Federal Register* 78: 5182-5183.
4. Beck AT, Steer RA, Brown GK (1996): *Beck depression inventory-II.* San Antonio, TX: The Psychological Corporation.
5. Eklund A, Nichols TE, Knutsson H (2016): Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*.
6. Cunningham WA, Koscik TR (2017): Balancing type I and type II error concerns in fMRI through compartmentalized analysis. *Cogn Neurosci*. 8:147-149.
7. Kang H, Blume J, Ombao H, Badre D (2015): Simultaneous control of error rates in fMRI data analysis. *NeuroImage*. *123*:102–113.
8. Lohmann G, Stelzer J, Mueller K, Lacosse E, Buschmann T, Kumar VJ, *et al.* (2017): Inflated false negative rates undermine reproducibility in task-based fMRI. *bioRxiv*:122788.
9. Lieberman MD, Cunningham WA (2009): Type I and type II error concerns in fMRI research: Re-balancing the scale. *Soc Cogn Affect Neurosci*. 4:423-428.