Longitudinal changes in amygdala, hippocampus and cortisol development following early caregiving adversity

Michelle VanTieghem a,*, Marta Korom a, Jessica Flannery c, Tricia Choy d, Christina Caldera e, Kathryn L. Humphreys f, Laurel Gabard-Durnam g, Bonnie Goff h, Dylan G. Gee h, Eva H. Telzer c, Mor Shapiro i, Jennifer Y. Louie j, Dominic S. Fareri k, Niall Bolger l, Nim Tottenham a

a Department of Psychology, Columbia University, New York, NY, USA
b Department of Psychological and Brain Sciences, University of Delaware, Newark, DE, USA
c Department of Psychology and Neuroscience, University of North Carolina Chapel Hill, Chapel Hill, NC, USA
d Graduate School of Education, University of California Riverside, Riverside, CA, USA
e Department of Psychology, University of California Los Angeles, Los Angeles, CA, USA
f Department of Psychology and Human Development, Vanderbilt University, Nashville, TN, USA
g Northeastern University, Department of Psychology, Boston, MA, USA
h Department of Psychology, Yale University, New Haven, CT, USA
i Family Medicine, Boca Raton, FL, USA
j Child Mind Institute, San Francisco Bay Area, San Mateo, CA, USA
k Gordon F. Derner School of Psychology, Adelphi University, Garden City, NY, USA

ABSTRACT

Although decades of research have shown associations between early caregiving adversity, stress physiology and limbic brain volume (e.g., amygdala, hippocampus), the developmental trajectories of these phenotypes are not well characterized. In the current study, we used an accelerated longitudinal design to assess the development of stress physiology, amygdala, and hippocampal volume following early institutional care. Previously Institutionalized (PI; N = 93) and comparison (COMP; N = 161) youth (ages 4–20 years old) completed 1–3 waves of data collection, each spaced approximately 2 years apart, for diurnal cortisol (N = 239) and structural MRI (N = 156). We observed a developmental shift in morning cortisol in the PI group, with blunted levels in childhood and heightened levels in late adolescence. PI history was associated with reduced hippocampal volume and reduced growth rate of the amygdala, resulting in smaller volumes by adolescence. Amygdala and hippocampal volumes were also prospectively associated with future morning cortisol in both groups. These results indicate that adversity-related physiological and neural phenotypes are not stationary during development but instead exhibit dynamic and interdependent changes from early childhood to early adulthood.

1. Introduction

Early caregiving adversity (ECA) is associated with increased risk for internalizing disorders, such as anxiety and depression (Green et al., 2010; Kessler et al., 2010). Decades of research suggests that ECA alters key biological systems implicated in psychopathology risk: stress physiology (e.g., hypothalamic-pituitary-adrenal (HPA) axis; Gunnar and Quevedo, 2008; Miller et al., 2007) and limbic brain regions (e.g., amygdala and hippocampus; Kribakaran et al., 2020; McEwen et al., 2016). However, the developmental trajectories and interactions that lead to these adult phenotypes following ECA have not been well characterized. Although the HPA-axis bi-directionally interacts with the amygdala and hippocampus (Herman et al., 2012), these variables are rarely examined in the same study, leaving open the question as to how they influence each other across development. Secondly, these systems change substantially across the first two decades of life (Flannery et al., 2017; Herting et al., 2018; Ostby et al., 2009), yet developmental effects of ECA are rarely examined, in part because longitudinal data are less common and cross-sectional designs often treat age as a nuisance variable (although see Flannery et al., 2017; Gunnar et al., 2019; King et al., 2017; Luby et al., 2019). These approaches may artificially suggest that ECA-related phenotypes are static, limiting our ability to understand the

* Corresponding author. E-mail address: mv2115@columbia.edu (M. VanTieghem).
neurodevelopmental sequelae of ECA exposure. In the current study, we address these challenges with an accelerated-longitudinal design to characterize developmental patterns of diurnal cortisol and limbic brain structure following early institutional care and examine their developmental interactions over time.

1.1. Diurnal cortisol

The HPA-axis is one of the primary mammalian physiological systems involved in regulating responses to environmental stressors (Gunnar and Quevedo, 2007). Diurnal cortisol provides an index of HPA-axis regulation, characterized by high morning levels that declines across the day in a circadian fashion (Gunnar and Quevedo, 2007), a slope that becomes more negative with increasing age (Flannery et al., 2017). ECA exposure is associated with alterations in daily rhythm, typically due to blunted morning levels in childhood (Bernard et al., 2015; Fisher and Stoolmiller, 2008; Koss et al., 2014; Pitula et al., 2019; Zalewski et al., 2016). Initial studies have suggested that adolescence represents a period of recalibration of stress physiology, with normalized (Flannery et al., 2017) or higher morning cortisol levels (Weems et al., 2017) and/or awakening responses (King et al., 2017) in ECA-exposed samples. These findings have been primarily cross-sectional studies or those longitudinally limited to narrow age-ranges (i.e., 2 year span; King et al., 2017), making it difficult to discern whether the literature reflects developmental changes in adversity-related cortisol phenotypes, or methodological differences (e.g. sample composition, analysis methods). The current longitudinal study addresses these questions and characterizes the developmental effects of ECA exposure on diurnal cortisol from early childhood to early adulthood.

1.2. Limbic brain volume

ECA exposure is associated with amygdala and hippocampal volumetric alterations (Tottenham and Sheridan, 2009)—regions critically involved in the development of healthy emotion regulation (Silvers et al., 2016). Amygdala and hippocampal volume reductions are frequently observed in adults with ECA exposure (Butterworth et al., 2012; Calem et al., 2017; Dannlowski et al., 2012; Riem et al., 2015; van Velzen et al., 2016) and have been linked with greater risk of internalizing psychopathology (Gorka et al., 2014; Rao et al., 2010). However, associations between ECA and limbic regions during development have been less clear. While hippocampal differences, if observed, have consistently shown hypotrophy in developmental samples (Hanson et al., 2011, 2014; Hodel et al., 2014; Humphreys et al., 2019; King et al., 2018; Luby, 2013; Noble et al., 2012; Piccolo and Noble, 2018), the literature on amygdala volume is mixed, with some studies showing larger (Lupien et al., 2011; Mehta et al., 2009; Roth et al., 2018; Tottenham et al., 2010), smaller (Edmiston, 2011; Hanson et al., 2014; Luby, 2013; Noble et al., 2012) or no differences (Hodel et al., 2014; King et al., 2018; Noble et al., 2015; Sheridan et al., 2012) between ECA and comparison youth. Although the amygdala and hippocampus undergo significant age-related changes (Herting et al., 2018; Oxby et al., 2009; Uematsu et al., 2012; Wierenga et al., 2018), the majority of prior studies used cross-sectional samples that control for age (although see Ellwood-Lowe et al., 2018; Keding and Herringa, 2015; Merz et al., 2018; Paquola et al., 2017; Whittle et al., 2017), leaving the possibility that existing discrepancies may be due to developmental effects. We address this gap by characterizing ECA-related developmental changes in amygdala and hippocampal volume across a wide age-range.

1.3. Cortisol-brain interactions

Animal studies have demonstrated that limbic brain development and the HPA axis are highly coupled during development (Myers et al., 2012). Amygdala and hippocampus are rich with cortisol-binding receptors, particularly early in development (Avishai-Eliner et al., 1996; Gilmore et al., 2012; Payne et al., 2010; Vázquez et al., 2012), and they provide regulatory feedback to the HPA-axis (Herman et al., 2012). However, there is a paucity of human studies that examine how these systems interact during development. Some longitudinal studies have tested this relationship in one direction, from child cortisol to adolescent brain (Burghy et al., 2012; Pagliaccio et al., 2015), leaving open the question whether there are also reciprocal influences. We leverage a longitudinal dataset with diurnal cortisol and subcortical brain volume measurements to investigate their bidirectional predictions across development.

1.4. Current study

In the current study, we used an accelerated longitudinal design to characterize developmental changes in stress physiology and limbic brain development in internationally adopted youth with a history of previous institutional (PI) caregiving. Institutional care is characterized by abandonment and sparse, unstable caregiving (Gunnar et al., 2000), a potent stressor for the developing infant (Tottenham, 2012). However, PI youth are subsequently adopted into stable caregiving settings, which allows for the rare opportunity to examine development following temporally defined adversity exposure. Here, we assessed developmental changes in diurnal cortisol, amygdala, and hippocampal volume following institutional care. Rather than focus on group differences, we hypothesized that ECA would be associated with altered rates of developmental change across childhood and adolescence, leading to different adversity-related phenotypes at different ages. Next, we probed the relationships between cortisol and limbic brain development using cross-lagged structural equation modeling (CL-SEM), in order to explore longitudinal and bidirectional associations between these systems across development and as a function of ECA.

2. Materials and methods

2.1. Participants

Previously institutionalized (PI; N = 93) and comparison (COMP; N = 161) youth between the ages of 4–20 years were included in the study (Table 1). PI youth were internationally adopted into the United States, from a range of countries and institutional settings (see Table 1). Although information on specific preadoption variables are unknown in PI populations in general, we obtained information regarding age of placement and duration of institutional care from families when available, and most youth had exposure to institutional care during the first 3 years of postnatal life. PI youth were recruited via local international adoption agencies, adoption family networks, posted flyers, and friend

| PI COMP Group difference |
|--------------------------|
| N | 93 | 161 |
| Sex (M/F) | 31 / 62 | 78 / 83 | p = 0.017 |
| IQ at T1, mean (SD) | 101.72 (16.68) | 111.78 (16.7) | p < 0.0001 |
| Age in years, mean (SD), range | 9.98 (3.35), 0–12 | 9.1 (4.12), 0–12 | p = 0.065 |
| Age adopted (months), median (SD), range | 16.5 (25.7), 0–72 | 4.08–17.58 | p = 0.065 |
| Country of Origin | Asian | 38 (41 %) | p = 0.001 |
| Country of Origin | Eastern European | 50 (54 %) | p = 0.001 |
| Country of Origin | Unknown/Other | 5 (5 %) | p = 0.001 |
referral. Healthy COMPS (always raised by biological parents in the United States) were recruited via birth records, posted flyers, and friend referral. COMPS were pre-screened for prior diagnoses of behavioral/psychological/learning difficulties. Participants who completed MRI sessions were screened for contraindications. The protocol was approved by the Institutional Review Board at the University of California, Los Angeles. Participants/parents provided informed assent/consent.

2.2. Study procedure

An accelerated longitudinal design was used, with baseline visits occurring at ages 4–16 years old (Fig. 1). By design, this study was not a full longitudinal study. First, participants were over-enrolled at baseline in order to ensure that target follow-up recruitment goals were reached, given the anticipated difficulties of scanning this age range (e.g., braces, motion artifact). Next, depending on entry date, participants were invited to complete either 1 or 2 follow-up visits in order to complete data collection within a 5-year period (e.g., participants enrolled in Year 1 were invited to 2 follow-up visits in years 3 and 5, while participants enrolled in year 2 were invited to a single follow-up visit in year 4). At each wave, participants completed structural MRI scans, and diurnal cortisol and parent-reported questionnaires were obtained (Table 2). For additional information on missingness patterns, see Supplemental Analyses Section 1. Follow-up retention for each sub-sample (cortisol, brain volume) did not vary by group, age or sex (Supplemental Analyses 2.1.1 and 3.1.1).

2.3. Cortisol sampling procedure

Participants provided salivary cortisol samples over 2 days (4 time points per day: wake up, 45 min after waking, 5 pm, and 8 pm). Families were instructed to collect samples before eating/drinking, or at least 15 min after eating/drinking, and not to collect samples on days youth felt ill. Saliva diaries (date, bedtime, illness, medication use, and/or unusual levels of activity) were included for each day of sampling and were used to account for psychotropic and oral steroid medication use (see Table S1 for medication rates by group). Full details on the processing protocol for cortisol samples are reported elsewhere (Flannery et al., 2017). Data were assayed in singlet (no an intra-assay coefficient is available), but cortisol values at each time of day were highly stable within-subject across day 1 and day 2 (p < 0.001 for all 4 cortisol sample time points, see Table S2).

2.4. Structural neuroimaging procedure

High resolution T1-weighted scans were acquired on a Siemens’ 3 T Trio scanner for waves 1 and 2 (TR = 2170 ms, TE = 4.33 ms, flip angle = 7°, 192 slices, 1 × 1 × 1 mm voxels, FOV = 256 mm, scan time = 8.08 min). Wave 3 was collected on a different Siemens’ 3 T Trio scanner (TR = 1900, TE = 3.26, flip = 9°, slices = 176, FOV = 250, scan time = 3.83 min).

2.5. Quality control procedures

2.5.1. Cortisol data

The four daily cortisol samples were categorized into morning and evening time points for analysis due to compliance and/or sample quality concerns (21% of diurnal cortisol samples had missing points or

---

**Table 2**

Demographic information by group and wave for sub-samples of participants who completed diurnal cortisol, structural MRI, and symptom assessments. PI = Previously Institutionalized, COMP = comparisons.

| Group     | Wave | N (M/F) | Sex (M/F) | Age (years) Mean | SD | Range |
|-----------|------|---------|-----------|------------------|----|-------|
| **Diurnal cortisol sample** |      |         |            |                  |    |       |
| COMP 1    | 141  | 71 / 70 | 9.15      | 4.13             | 4.08–17.58 |
| COMP 2    | 56   | 22 / 34 | 11.50     | 4.32             | 5.25–20.33 |
| COMP 3    | 30   | 11 / 19 | 11.49     | 3.75             | 6.67–19.08 |
| PI 1      | 82   | 29 / 53 | 9.76      | 3.41             | 3.92–17.17 |
| PI 2      | 37   | 10 / 27 | 12.08     | 3.42             | 6.25–18.25 |
| PI 3      | 34   | 9 / 25 | 13.66     | 3.56             | 7.08–18.83 |
| **Structural MRI sample** |      |         |            |                  |    |       |
| COMP 1    | 70   | 35 / 35 | 10.67     | 3.91             | 4.25–18.58 |
| COMP 2    | 69   | 29 / 40 | 11.88     | 4.18             | 4.83–20.33 |
| COMP 3    | 41   | 14 / 27 | 12.31     | 3.99             | 6.67–21.08 |
| PI 1      | 145  | 18 / 77 | 10.67     | 2.86             | 4.58–16.58 |
| PI 2      | 45   | 15 / 30 | 12.41     | 3.15             | 6.75–18.25 |
| PI 3      | 36   | 12 / 24 | 13.90     | 3.32             | 7.08–18.83 |

---

**Fig. 1.** Age of sampling for accelerated longitudinal design. Participants completed 1-3 waves of data collection for diurnal cortisol and/or structural scans. COMP = comparison, PI = Previously Institutionalized.
time logs). Rates of missingness in cortisol data at each time point did not differ by group (Table S3). Of those samples that included time logs, 24% of morning samples were collected greater than 45 min apart. In the absence of accurate time stamps or sleep data, we did not model the cortisol awakening response (CAR). Instead, we assessed change in cortisol concentration from morning to evening. Per Salimetrics guidelines for cortisol ranges in developmental samples, cortisol values > 75 nmol/L were excluded as biologically implausible (Salimetrics, 2018). The final sample included 239 participants (151 comparisons, 88 PIs) with a total of 380 diurnal cortisol samples collected across 1–3 waves (Table 2). A subsample of these participants was previously published in a cross-sectional analysis (Flannery et al., 2017). Raw values were used since linear mixed effects models are robust to non-linearity in the first level (Maas and Hox, 2004). We controlled for batch effects on cortisol values (Table S4) as a covariate in all models.

2.5.2. Structural MRI data
Freesurfer v6.0 (Fischl et al., 2002) identified subject-specific segmentations of the amygdala and hippocampus. We used the cross-sectional Freesurfer stream, as the longitudinal pipeline is not recommended for developmental studies with a wide age-range (Reuter, 2016). Scan quality was assessed by two independent raters to identify (a) motion artifact and (b) subcortical segmentation quality on a scale of 1 (good) to 4 (poor), with inter-rater reliability (Cronbach’s alpha) of 0.80 and 0.85, respectively. MRI images with substantial motion artifact (rating of 4 by at least one rater, N = 13), substantial segmentation errors (rating of 4, N = 4), or failed Freesurfer processing (N = 2) were excluded. Outliers, defined as values beyond 3SD of the mean within each group and wave, were excluded (ICV = 1, amygdala = 3, hippocampus = 0). The final sample (N = 156; 66 PIs, 90 comparisons) had a total of 306 MRI scans collected across 1–3 waves (Table 2).

Volumes were extracted for hippocampus, amygdala, and intracranial volume (ICV; intracranial brain volume). Amygdala and hippocampal volumes were scaled to 100 mm and ICV to 100,000 mm for all analyses. Due to no a priori hypotheses on laterality and high correlations between hemispheres for amygdala (r = 0.73, p < .001) and hippocampus (r = 0.88, p < .001), bilateral volumes were used for analyses. To account for the scanner/protocol differences in wave 3, a dummy covariate for scanner was included in all analyses (Noble et al., 2015). Sensitivity analyses with only the sub-sample of data collected on the same scanner (i.e., omitted wave 3) and right and left hemispheres are provided in Supplemental Analyses 3.1.3 and 3.1.4, respectively.

2.6. Modeling age-related change
Linear mixed effects models tested for Group X Age interactions on brain volume (amygdala, hippocampus) and Group X Age X Time of day interactions on cortisol concentration. Any significant interactions were interrogated with non-linear Age X Group effects using quadratic and piecewise models. For piecewise models, we conducted a data-driven iterative search (using the optimize function in R) to determine the optimal breakpoint in age that minimized the deviance of the model fit. The optimal piecewise model was then compared with quadratic and linear models using Akaike Information Criterion (AIC) to determine the best model fit. All modeling was performed using the lme4 and lmerTest package in R (Bates et al., 2015; Kuznetsova et al., 2017), which uses maximum likelihood estimation and Satterthwaite method of degrees of freedom (Satterthwaite, 1946), and all model results are reported with 95% confidence intervals (CI). Significant piecewise Age X Group interactions were visualized with a regions of significance plot (Preacher et al., 2006) which shows significant group differences across the entire age range. For each model, we also calculated the intraclass correlation coefficient (ICC) using the R sjstats package (Lüdecke, 2018) to assess the reliability of within-subject estimates for each outcome variable.

All analyses controlled for sex due to the higher incidence of females in the PI group (Table 1), reflecting the broader demographic that females tend to be more common in PI populations (Hellerstedt et al., 2008). For cortisol models, we also controlled for day of collection, wave of collection, batch of sample processing (conducted in 5 batches over 5 years), and medication status (1 = medication, 0 = no medication for that wave). Random intercepts per subject and random effects of wave were included to account for repeated measures within subject. For structural brain volume models, we included covariates of ICV, motion ratings, and scanner. Random intercepts per subject were included to account for repeated measurements; random effects of wave were not included due to too few longitudinal data points for model convergence.

2.7. Cross-lag autoregressive SEM models
To examine bidirectional effects between cortisol and limbic brain volume across development, we used cross-lag structural equation modeling (CL-SEM) using the Lavaan package in R (Rosseel, 2012). To maximize the sample size, we used a sub-sample of participants with usable cortisol and structural MRI data from at least one wave (Table S8) and used Full information maximum likelihood (FIML) to account for missing data. T1 represented the first wave, and T2 represented the final wave of data available; T1-T2 intervals ranged from 2 to 4 years (mean = 2.73, SD = 1.03). Data from participants with usable data from all time points were included across the maximal time delay available (e.g., data at wave 1 and wave 3). Based on results showing significant group and age-related differences in morning but not evening cortisol, morning cortisol values were used for these analyses (see Supplemental Analyses for sensitivity analyses using diurnal cortisol slopes). Amygdala and hippocampal volumes highly covaried (r = 0.62, p < .001). To avoid suppression effects, they were modeled separately.

We tested two cross-lag paths: Path 1 tested the effect of morning cortisol T1 on hippocampal/amygdala volume at T2, controlling for hippocampal/amygdala volume at T1 and covariates; Path 2 tested the effect of hippocampal/amygdala volume T1 on morning cortisol at T2, controlling for morning cortisol at T1 and covariates. Separate cross-sectional regressions controlled for the effects of covariates on T1 measures of hippocampal volume and morning cortisol. Covariance between the two T1 variables and T2 errors are also reported. In a parallel set of models, group interactions were added to each cross-lagged path (Supplemental Analyses 4.2). Model fit was assessed using nested model comparisons (Supplemental Analyses 4.1).

3. Results

3.1. Developmental changes in morning cortisol following institutional caregiving
A significant Group X Age X Time of Day interaction was detected for diurnal cortisol (nmol/L) using linear mixed effects modeling (b = 0.78, t(2,187) = 4.27, p < 0.001, CI = 0.42, 1.14). We then probed non-linear age effects using quadratic and piecewise age terms. A non-linear piecewise age model showed the best model fit (Table 3), with an optimal breakpoint at 13.1 years (CI = 11.7, 15.6) with a significant Group X Age X Time of Day interaction for ages 13 and older, but not before age 13 (Table S9). To further interrogate these interactions, follow-up models tested piecewise age effects for morning and evening separately.

3.1.1. Morning cortisol
Modeling morning cortisol separately with piecewise Age X Group effects verified an optimal breakpoint at 13.1 years old (Fig. 2A).
Table 3
Cortisol and amygdala model comparisons. Piecewise age models provided the best model fit for diurnal cortisol, morning cortisol, and amygdala volume. AIC = Akaike information criterion.

| Model                          | DF  | AIC          |
|-------------------------------|-----|--------------|
| Diurnal Cortisol              |     |              |
| Linear Age x Group x Time of Day | 17  | 17,955.98    |
| Quadratic Age x Group x Time of Day | 21  | 17,970.32    |
| Piecewise Age x Group x Time of Day | 21  | 17,942.41    |
| Morning cortisol              |     |              |
| Linear Age x Group            | 13  | 9,900.62     |
| Quadratic Age x Group         | 15  | 9,906.29     |
| Piecewise Age x Group         | 15  | 9,892.28     |
| Evening cortisol              |     |              |
| Linear Age x Group            | 13  | 7,380.11     |
| Quadratic Age x Group         | 13  | 7,390.83     |
| Piecewise Age x Group         | 13  | 7,379.13     |
| Amygdala volume               |     |              |
| Linear Age x Group            | 10  | 1049.75      |
| Quadratic Age x Group         | 12  | 1023.08      |
| Piecewise Age x Group         | 12  | 1015.98      |

Piecewise Age X Group effects showed differences in age-related change in morning cortisol between groups. Before age 13.1, age moderated the effects of group (b = −0.68, t (301) = −2.12, p = 0.034, CI = −1.29, −0.08); COMPS showed a significant age-related increase in morning cortisol during this age-range (b = 0.73, t (321) = 4.04, p < 0.001, CI = 0.37, 1.08) but the PI group did not (b = 0.05, t (321) = 0.20, p = 0.841, CI = −0.46, 0.56). After age 13.1, age also moderated the effects of group (b = 2.03, t (349) = 2.89, p < 0.01, CI = 0.65, 3.40), such that age-related increases in morning cortisol were observed in the PI group (b = 2.38, t (313) = 4.07, p < 0.001, CI = 1.23, 3.53) but not the COMPS (b = 0.34, t (313) = 0.84, p = 0.404, CI = −0.47, 1.16). The sensitivity plot in Fig. 2B shows that PI group was estimated to have lower morning cortisol relative to COMPS between ages 9 and 14, and higher morning cortisol after age 17.5 (Table S10). For follow-up analyses on the effects of 1) IQ, race, income and education, 2) sex as a moderator, and 3) testosterone on morning cortisol, see Supplemental Analyses 2.1.2, 2.2.2, and 2.2.3, respectively. Secondary analyses also detected a dose-dependent relationship on morning cortisol in the PI group, such that later age of adoption was associated with lower morning cortisol (controlling for age; Supplemental Analyses 2.2.1). Finally, intra-class coefficients (ICC) of the random effects of the model were relatively low (ICC = 0.22), indicating high variability of morning cortisol across waves at the within-subject level.

3.1.2. Evening cortisol
We repeated the piecewise model with evening cortisol only, using the same age-breakpoint of 13.1 years old. Group did not moderate the piecewise age effects before or after 13.1 years old (Table S11). In contrast to morning cortisol, which showed the best model fit with piecewise age terms (i.e., lower AIC by > 2; Table 3), using piecewise age terms for evening cortisol did not provide a better model fit relative to a linear age term (Table 3).

3.2. Age-dependent associations between institutional caregiving and amygdala volume

Figure 2. Effects of PI status on morning cortisol depend on age. (A) Fitted results of piecewise Age X Group effects on morning cortisol are depicted. Raw cortisol values (lines connecting within-subject observations) are shown with a 95 % CI band around the fitted regression lines. (B) The region of significant plot is shown, depicting the magnitude of group differences in morning cortisol (PI – COMP) across the entire age-range. When the 95 % CI band is above zero, morning cortisol is significantly higher in PI group than the comparison group, and when the 95 % CI is below zero, morning cortisol is significantly lower in the PI group than the comparison group. * p < 0.05, ** p < 0.01.
ages 11 and 16.5 and larger estimated amygdala volumes before age 6.5, although estimations in the youngest and oldest age range should be interpreted with caution due to imbalanced observations between groups. Further sensitivity analyses are provided in Supplemental Analyses Section 3. ICC estimates of amygdala volume within-subject from the piecewise model are 0.77, indicating high within-subject reliability across waves of the study.

3.3. Age-invariant associations between institutional caregiving and hippocampal volume

We observed a significant main effect of group on hippocampus volume (b = -1.78, t(129.82) = -3.24, p < 0.01, CI = -2.86, -0.71), such that PI youth had smaller hippocampal volumes relative to COMPS (Fig. 4). We also detected a main effect of age (b = 0.23, t (290.15) = 4.03, p < 0.001, CI = 0.12, 0.34) on hippocampal volume (Table S13). In comparison to amygdala volume, PI status did not moderate the effects of age on hippocampal volume (b = -0.10, t (252.55) = -1.09, p = 0.277, CI = -0.29, 0.08). See Supplemental Analyses Section 3 for sensitivity and secondary analyses.

3.4. Prospective associations between limbic brain volume and morning cortisol

Bidirectional associations between morning cortisol and limbic brain volumes over time were assessed with CL-SEM. Amygdala volume at T1 predicted morning cortisol at T2 (Fig. 5), over and above the effects of baseline cortisol, such that T2 morning cortisol was better predicted by
T1 amygdala volume than T1 morning cortisol (Table S14). Importantly, morning cortisol at T1 did not predict amygdala volume at T2. Parallel results were found when modeling bidirectional effects between hippocampus and cortisol (Fig. 5), such that T1 hippocampus volume predicted T2 morning cortisol, but T1 cortisol did not predict T2 hippocampal volume (Table S15). To visualize these cross-lagged relationships for both amygdala and hippocampal models, separate linear regressions were performed including only the significant paths (with the same covariates included). As shown in Fig. 6, T2 morning cortisol was predicted by T1 amygdala volume ($b = 2.730, t (63) = 2.803, p < 0.01$) and T1 hippocampal volume ($b = 0.931, t (63) = 2.256, p = 0.027$). Group (PI vs. COMP) did not significantly moderate these cross-lag paths (Amygdala: $z = 1.21, p = 0.228, b = 1.85$; Hippocampus: $z = -0.20, p = 0.839, b = -0.20$; Supplemental Analyses 4.2). Sensitivity analyses using cortisol slope instead of morning values are provided in Supplemental Analyses 4.3.

4. Discussion

The present study demonstrates that adversity-related alterations in stress physiology and limbic brain volume change dynamically across development. Specifically, associations between early institutional care and amygdala volume and morning cortisol depended on age, whereas adversity-related hippocampal reductions were age-invariant from childhood to early adulthood. Moreover, amygdala and hippocampal volumes were prospectively associated with higher future morning cortisol levels in both groups, as shown by longitudinal cross-lagged models, suggesting a feed-forward relationship from amygdala and hippocampus to the HPA-axis during childhood and adolescence. These findings emphasize the importance of longitudinal studies to characterize and interpret adversity-related phenotypes.

We identified a developmental shift in ECA-related cortisol phenotypes, such that PI group showed bluntened morning cortisol during childhood and heightened levels by late adolescence. These non-linear age effects provide a developmental framework to integrate prior discrepancies in the literature of both bluntened and heightened morning cortisol phenotypes (King et al., 2017; Quevedo et al., 2012). The observed pattern of bluntened morning cortisol in PI children is consistent with numerous prior studies (Bernard et al., 2015; Fisher and Stoolmiller, 2008; Koss et al., 2014; Pitula et al., 2019; Zalewski et al., 2016) and is hypothesized to emerge over time as a result of excess cortisol in response to a chronic stressor (Miller et al., 2007). In contrast, the marked increase in morning cortisol observed in PI adolescents corresponds with recent work suggesting pubertal recalibration in stress physiology, observed in both morning cortisol levels (Flannery et al., 2017; King et al., 2017; Quevedo et al., 2012) and reactivity to stressors (Gunnar et al., 2019). Consistent with this idea, supplemental analyses showed effects of puberty, independent of age, on morning cortisol in both PI and comparison groups (see Supplement). This adolescent-specific plasticity may allow for positive influences (e.g. time with adoptive family) to recalibrate the HPA-axis for more adaptive functioning following early caregiving adversity, as indicated by prior studies (DePasquale et al., 2019). However, further within-person longitudinal research is needed to determine whether the observed developmental shift in cortisol is associated with adaptive or maladaptive outcomes in PI youth.

The PI group showed altered amygdala volume growth, which resulted in differing group effects depending on the age of measurement. COMPS showed age-related amygdala growth during childhood which plateaued around age 9–10, while the PI group showed relatively stable amygdala volumes from early childhood to early adulthood. Therefore, although the PI group had larger estimated amygdala volumes at the youngest ages, due to subsequent lack of age-related growth during childhood, they exhibited smaller amygdala volumes relative to COMPS by adolescence. These findings suggest that prior conflicting findings of smaller, larger, or no differences in amygdala volume following ECA may reflect differences in age sampling or analysis (i.e., controlling for age). For example, the majority of cross-sectional studies reporting larger amygdala included younger participants (Buss et al., 2012; Lupien et al., 2011; Roth et al., 2018; Tottenham et al., 2010) and a recent dense-sampling study between ages 4–6 also showed larger amygdala volume in the context of parental insensitivity (Lee et al., 2019). In contrast, the majority of studies reporting smaller amygdala volume following ECA include participants in late adolescence and young adulthood (Edmiston, 2011; McLaughlin et al., 2016; Noble et al., 2012; Saxbe et al., 2018 [but see Hanson et al., 2014; Luby, 2013]) or only detected smaller amygdala volume in adolescence (Korcanek et al., 2013; Merz et al., 2018). Although the extant literature suggests that amygdala volumetric differences have been observed most often in the context of physical threat (see McLaughlin et al., 2019 for review), the

![Fig. 5](image_url)
current findings extend our understanding by showing that the adversity that accompanies institutional caregiving is also followed by amygdala volumetric differences.

However, it remains unclear whether the observed age-related differences in amygdala volume reflect developmental effects or temporal proximity to adverse experiences (Gerritsen et al., 2015; Tottenham and Sheridan, 2009). Prior hypotheses have suggested that early increases in amygdala volume, as shown in animal models (Guadagno et al., 2018; Vyas et al., 2002), may sensitize the amygdala to future stressors, resulting in later-occurring amygdala atrophy (Teicher et al., 2016).

However, the apparent reduced amygdala volume in the current study of PI adolescents is not due to decrease of amygdala volume over time, but instead reflects a lack of age-related growth in PI children relative to comparison children. As such, it is possible instead that the period of amygdala growth observed in childhood is shifted earlier in time by exposure to ECA and could result in a lower ceiling of possible maximum volume. Further research is needed to characterize amygdala volume changes across earlier in life, closer to the time of adversity exposure, to further determine how ECAs alter developmental timing of amygdala growth.

PI status was also associated with reduced hippocampal volume across all ages. These results are consistent with prior research showing adversity-related reductions in hippocampal volumes in samples of varying age-ranges (Calem et al., 2017 for meta-analysis). That these reductions are stable by 4 years and persist throughout development suggests that ECA exposure has an enduring impact on hippocampal volume, in line with work suggesting that early life is a sensitive period for hippocampal development (Andersen et al., 2008; Humphreys et al., 2019). These data are also consistent with animal models, showing hippocampal volumetric reductions and reduced dendritic complexity following chronic stress (Magarínos and McEwen, 1995; Vyas et al., 2002).

Not all human developmental studies have identified smaller hippocampal volume following ECAs (Raineki et al., 2019) and in preschoolers (Pagliaccio et al., 2001; Kumsta et al., 2017), later age of adoption was associated with outcomes of interest (see Supplemental Analyses). Consistent with previous studies (Gunnar et al., 2001; Kumsta et al., 2017), late age of adoption was associated with lower morning cortisol, but we did not detect dose-dependent relationships with amygdala or hippocampal volume. However, it is important to note that dose-dependent effects of ECA on stress physiology and limbic neurobiology may change over development or wane over time, and in the current study, we were under-powered to test such interactions. Additional longitudinal research is needed to further assess how timing and chronicity of ECA exposure might influence the developmental trajectories of limbic brain development and stress physiology.

This study has limitations to be noted. Although institutional care is an important means of studying the role of adversity that is restricted to...
very early life, it is a relatively extreme and rare form of ECA, which may limit the generalizability of these findings to other forms of ECA (e.g., abuse). Second, the data in this study were collected as part of a large-scale study, whose design intentionally did not include longitudinal follow-ups for all participants. As a result, we were not able to model within-subject changes in cortisol or brain volume (Madhyastha et al., 2018) which may reveal different associations at the single-subject level. We also note the imbalance between PI and comparison groups at younger and older ages in the MRI data. Secondary analyses conducted in a restricted age range (6–19 years) yielded similar results for both hippocampal and amygdala volume, providing confidence that the observed effects were not driven by leverage points in the younger age-ranges. However, given evidence that discrepancies in volume studies may also arise from methodological choices such as processing pipelines (Lyden et al., 2016), future research is needed to extend/replicate these findings, particular with higher rates of sampling at the extreme ages. Similarly, the limited sample size of participants with two usable data points of both cortisol and limbic brain volume prevented addressing whether their bidirectional relationships interacted with age. It should be noted that while the CL-SEM showed longitudinal associations between limbic brain volume and morning cortisol, we cannot conclude causality from these models. Instead, they provide evidence to suggest directional coupling between these stress-responsive systems over development. Finally, because we do not have information about prenatal/developmental histories for PI youth. This is a common issue for investigators studying this population and prevents us from concluding that these phenotypes are purely the result of institutional care. Continued translational work in animal models are needed to assess whether manipulations in amygdala and hippocampus development at specific ages influence future diurnal cortisol and vice versa.

5. Conclusion

The present study demonstrated that ECA-related changes in stress physiology and limbic neurobiology depend on developmental stage. Although we observed smaller amygdala and hippocampal volumes in PI adolescents, these volume reductions emerged via different developmental mechanisms (i.e., smaller volume vs. reduced growth). Further, although the effects of ECA on amygdala and hippocampal volume were relatively stable after age 10, stress physiology continued to show dramatic developmental shifts during adolescence. By capitalizing on longitudinal data across wide age-ranges, we can gain greater resolution into the developmental sequelae of ECA, with critical implications for identifying sensitive windows of development when potential recalibration of stress physiology could occur.

Funding

This work was supported by the National Science Foundation GRFP [grant number DGE-1644869]; American Psychological Foundation; the Dana Foundation; and the National Institute of Mental Health of the National Institutes of Health [Award Numbers F31MH115686 and 5R01MH091864-08]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

CRediT authorship contribution statement

Michelle VanTieghem: Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data curation, Visualization, Writing - original draft. Marta Korom: Resources, Data curation, Writing - review & editing. Jessica Flannery: Project administration, Investigation, Data curation. Tricia Choy: Project administration, Investigation, Data curation. Christina Caldera: Project administration, Investigation, Data curation. Kathryn L. Humphreys: Investigation, Writing - review & editing. Laurel Gabard-Durnam: Investigation, Writing - review & editing. Bonnie Goff: Investigation. Dylan G. Gee: Investigation. Eva H. Telzer: Investigation. Mor Shapiro: Investigation. Jennifer Y. Louie: Investigation. Dominic S. Fareri: Investigation. Niall Bolger: Formal analysis, Supervision, Writing - review & editing. Nim Tottenham: Conceptualization, Methodology, Funding acquisition, Project administration, Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Jasmine Feliciano and Liliana Varman for assisting with Freeesurfer data quality control, and Paul Bloom, Anna Vannucci, and Andrea Fields for providing feedback on the data analysis.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.dcn.2021.100916.

References

Andersen, S.L., Tomada, A., Vincow, E.S., Valente, E., Polcari, A., Teicher, M.H., 2008. Preliminary evidence for sensitive periods in the effect of childhood sexual abuse on regional brain development. J. Neuropsychiatry Clin. Neuropsy. 20 (3), 293–301. https://doi.org/10.1176/jnp.2008.20.3.292.

Avishai-Eliner, S., Yi, S.J., Baram, T.Z., 1996. Developmental profile of messenger RNA for the corticotropin-releasing hormone receptor in the rat limbic system. Brain Res. Dev. Brain Res. 91 (2), 159–163. https://doi.org/10.1016/0165-0173(95)01581-z.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67 (1), 1-28. https://doi.org/10.18637/jss.v067.i01.

Bernard, K., Zwerling, J., Dozier, M., 2015. Effects of early adversity on young children’s diurnal cortisol rhythms and externalizing behavior. Dev. Psychobiol. 57 (8), 935–947. https://doi.org/10.1002/dev.21324.

Bradbury, M.J., 1994. Role of type I and II corticosteroid receptors in regulation of basal activity in the hypothalmo-pituitary-adrenal axis during the diurnal trough and the peak: evidence for a nonadditive effect of combined receptor occupation. Endocrinology 134 (3), 1286–1296. https://doi.org/10.1210/en.134.3.1286.

Burghy, C.A., Stodola, D.E., Ruttle, P.L., Molloy, E.K., Armstrong, J.M., Oler, J. A., Fox, M. E., Hayes, A.S., Kalin, N.H., Enzwe, M.J., Davidson, R.J., Birm, R.M., 2012. Developmental pathways to amygdala-prefrontal function and internalizing symptoms in adolescence. Nat. Neurosci. 15 (12), 1736–1741. https://doi.org/10.1038/nn.3257. PMCID: PMC3509229.

Buss, C., Davis, E.P., Shahbaba, B., Pruessner, J.C., Head, K., Sandman, C.A., 2012. Maternal cortisol over the course of pregnancy and subsequent child amygdala and hippocampus volumes and affective problems. Proc. Natl. Acad. Sci. 109 (20), E1312-E1319. https://doi.org/10.1073/pnas.1209510.

Butterworth, P., Chorbin, N., Sachdev, P., Anstey, K.J., 2012. The association between financial hardship and amygdala and hippocampal volumes: results from the PATH through life project. Soc. Cogn. Affect. Neurosci. 7 (5), 548–556. https://doi.org/10.1093/scan/nst027.

Calem, M., Bronis, K., McGuire, P., Morgan, C., Kempton, M.J., 2017. Meta-analysis of associations between childhood adversity and hippocampus and amygdala volume in non-clinical and general population samples. Neuroimage Clin. 14, 471–479. https://doi.org/10.1016/j.nicl.2017.02.016.

Dahmen, B., Puetz, V.B., Schark, W., von Polier, G.G., Herpertz-Dahlmann, B., Konrad, K., 2017. Effects of early-life adversity on hippocampal structures and associated hva axis functions. Dev. Neurosci. https://doi.org/10.1159/000484238.

Dannulovski, U., Stuhrmann, A., Beutelmann, V., Zwanzger, P., Lenzen, T., Grotegerd, D., Domschke, K., Hohoff, C., Ohrmann, P., Bauer, J., Lindner, C., Postert, C., Konrad, K., Arot, V., Heindel, W., Suslow, T., Kugel, H., 2012. Limbic scars: long-term consequences of childhood maltreatment revealed by functional and structural magnetic resonance imaging. Biol. Psychiatry 71 (4), 286–293. https://doi.org/10.1016/j.biopsych.2011.10.021.
