4-1-2020

Infant behavioral reactivity predicts change in amygdala volume 12 years later

Courtney A. Filippi

Jessica F. Sachs

Dominique Phillips

Anderson M. Winkler
The University of Texas Rio Grande Valley, Anderson.Winkler@utrgv.edu

Andrea L. Gold

See next page for additional authors

Recommended Citation

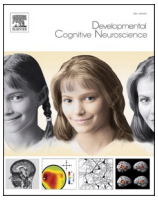
Filippi, C. A., Sachs, J. F., Phillips, D., Winkler, A., Gold, A. L., Leibenluft, E., ... & Fox, N. A. (2020). Infant behavioral reactivity predicts change in amygdala volume 12 years later. Developmental Cognitive Neuroscience, 42, 100776. https://doi.org/10.1016/j.dcn.2020.100776

Developmental Cognitive Neuroscience 42 (2020) 100776

Contents lists available at ScienceDirect

Developmental Cognitive Neuroscience

journal homepage: www.elsevier.com/locate/dcn



Infant behavioral reactivity predicts change in amygdala volume 12 years later

Courtney A. Filippi [a,b,*], Jessica F. Sachs [a], Dominique Phillips [a], Anderson Winkler [a], Andrea L. Gold [c], Ellen Leibenluft [d], Daniel S. Pine [a], Nathan A. Fox [b]

[a] Section on Development and Affective Neuroscience, National Institute of Mental Health, Bethesda, MD, 20892, United States
[b] Department of Human Development and Quantitative Methodology, University of Maryland, College Park, MD, 20742, United States
[c] Department of Psychiatry and Human Behavior, Brown University Warren Alpert Medical School, Providence, RI, 02912, United States
[d] Section on Mood Dysregulation and Neuroscience, National Institute of Mental Health, Bethesda, MD, 20892, United States

ARTICLE INFO

Keywords:
Amygdala
Temperament
Reactivity
Brain
Development

ABSTRACT

The current study examined the link between temperamental reactivity in infancy and amygdala development in middle childhood. A sample (n = 291) of four-month-old infants was assessed for infant temperament, and two groups were identified: those exhibiting negative reactivity (n = 116) and those exhibiting positive reactivity (n = 106). At 10 and 12 years of age structural imaging was completed on a subset of these participants (n = 75). Results indicate that, between 10 and 12 years of age, left amygdala volume increased more slowly in those with negative compared to positive reactive temperament. These results provide novel evidence linking early temperament to distinct patterns of brain development over middle childhood.

1. Introduction

Temperament operates in a developmental context and thus continually shapes how children evoke and respond to their environment (Fox et al., 2001, 2005). In the first few months of life, infants already differ in their reactions to novel stimuli. Some infants are highly aroused and distressed by novelty (Fox, 1989; Kagan et al., 1988; Stifter and Fox, 1990). These infants are described as having a *negative reactive temperament*. Others, who show positive affect in response to novelty, are described as having a *positive reactive temperament*. These differences in early temperament are thought to gradually shape the brain and behavior over time (Kagan et al., 1988)—particularly brain regions associated with emotion processing (e.g., amygdala). However, to date, only one study links negative reactivity measured during infancy to amygdala function in adults (Schwartz et al., 2012) and none examine such associations with amygdala structure. Despite considerable interest in the neurobiology of temperament (Clauss et al., 2014; Schwartz et al., 2012; Sylvester et al., 2016), to date no prospective evidence demonstrates that the brain *develops* differently as a function of early temperament. The current study addresses this gap by longitudinally assessing brain morphometry across middle childhood.

Negative reactive temperament is associated with the development of behavioral inhibition, a temperament characterized by fearful behavior (particularly in response to novel stimuli) in toddlerhood. Negative reactive infants are significantly more likely to develop behavioral inhibition than non-reactive or positive reactive infants (Fox et al., 2015), although not all children follow this trajectory. Amygdala function is hypothesized to moderate progression of negative reactivity to behavioral inhibition (Snidman et al., 1995). Early emerging differences in the amygdala may continuously impact children's environmental responses, potentially reinforcing fearful behavior. However, to date, limited research relates negative reactive temperament to amygdala structure or function. Schwartz et al. (2012) found greater amygdala responsivity to novel faces in *adults* classified as negative reactive in infancy. At present, only one study has linked fearful temperament (measured via retrospective self-report) to larger amygdalae in adults (Clauss et al., 2014). However, this study utilized retrospective self-report of childhood temperament, leaving open the question of whether rigorous behavioral phenotyping in infancy would produce similar results. Moreover, these data focus on brain structure in adults. Thus, it remains unknown whether the same patterns are present in childhood.

The amygdala, like many structures in the brain, develops in a non-linear fashion. Studies in typically-developing children document

* Corresponding author at: 9000 Rockville Pike, Bethesda, MD, 20892, United States.
E-mail address: cfilippi@umd.edu (C.A. Filippi).

https://doi.org/10.1016/j.dcn.2020.100776
Received 17 September 2019; Received in revised form 6 March 2020; Accepted 9 March 2020
Available online 21 March 2020

amygdala volume growth through middle childhood, with volume peaking during preadolescence before shrinking (Durston et al., 2001; Giedd, 2008; Giedd et al., 1996, 2006; Wierenga et al., 2014). While the amygdala is central to fearful behavior (Bauman et al., 2004; LeDoux, 2003; Yang et al., 2008) and theories of infant temperament have speculated that changes in the amygdala are linked to stable patterns of fearful behavior (Kagan et al., 1988), no research, to date, has examined whether changes in amygdala development during childhood are associated with early temperament. Identifying neural markers of fearful behavior requires a comprehensive understanding of how early fear phenotypes (e.g., negative reactivity) are linked to maturation of the amygdala. The current study takes this approach.

The current study evaluates the association between infant reactivity and change in amygdala structure during middle childhood. To do so, we link "gold-standard" behavioral assessments of infant temperament ($n = 291$) with longitudinal imaging data acquired at age 10 and 12 ($n = 117$). Specifically, we compare amygdala volume of children previously classified as negative reactive infants ($n = 116$) to those classified as positive reactive infants ($n = 106$; See Supplement for details on non-reactive infants; $n = 69$). However, no prior studies examine associations between temperament and rates of amygdala growth. We hypothesize that negative reactivity relates to deviations in amygdala development, which could manifest in one of two ways. Negative reactive infants may exhibit consistently larger amygdalae across preadolescence, when amygdala volume peaks after steady growth. Alternatively, negative reactive infants may fail to show the expected normative pattern of volume increase for children of this age. We also provide exploratory analyses for other regions of the brain.

2. Method

2.1. Participants

Seven hundred and seventy-nine infants were recruited from the Washington DC metro area for a study of early temperament with the goal of identifying and recruiting infants who showed distinct temperamental reactivity patterns to novelty. Infants were excluded if they were born premature, had low birth weight, a known developmental disorder, and/or birth complications. Parental consent was obtained prior to all visits and child assent was obtained prior to 10 and 12 year visit. All 779 infants recruited were screened via an assessment of reactivity to novel visual and auditory stimuli at 4 months using the procedure described in Calkins et al., 1996. Following this assessment, infant behavior was coded for the frequency of positive affect, negative affect, and motor reactivity. Of the 779 infants that were screened, the scores of the first 96 infants enrolled were used to generate cut-off criteria to create temperament groups. Infants who were above the median for both positive affect and motor reactivity were classified as "positive reactive" and those who were above the median for both negative affect and motor reactivity were classified as "negative reactive." If infants met criteria for both of these groups, they were placed in either the high positive reactive or high negative reactive groups depending on their affective bias (i.e., the difference between the standardized scores of positive affect and negative affect). Specifically, children whose affect was more positive than negative were classified as high positive reactive, whereas those whose affect was more negative than positive were classified as high negative reactive. The infants who did not meet criteria for either the positive or negative reactive groups were classified as "non-reactive". This screening procedure was used to decide which of the 779 infants were invited to participate in longitudinal follow-up visits. 291 infants ($M = 4$ months, 2 days, range = 3 months, 22 days–4 months, 13 days) were invited to participate in subsequent assessments. The 291 infants were oversampled for positive and negative reactive groups (106 positive reactive, 116 negative reactive, and 69 non-reactive group).

At ages 10 and 12, eligible children were invited to participate in brain imaging visits. Brain imaging visits were conducted at the National Institute of Mental Health as part of the longitudinal assessment. Participants were excluded if they were taking any psychotropic medications at the time of scanning. However, subjects on psychostimulant medications who could tolerate a 24 -h medication-free period prior to scanning were included. Children were also deemed ineligible to participate in imaging visits if they had an MRI contraindication (i.e., metal in their body). 117 children (57 male) came in for at least one fMRI visit. Children who contributed brain imaging data did not differ from children who did not with respect to sex, maternal education, or reactivity group (*ps* >.46). However, individuals who participated in brain imaging visits did have marginally lower anxiety at age 12 (measured via mean parent and child SCARED; $p < .056$) than those who did choose not to participate (Smith et al., 2019). Furthermore, among children who provided imaging data, we found that there was a marginal reactivity group difference between average parent and child SCARED—with negative reactive individuals exhibiting heightened anxiety ($p < .082$). Supplemental Table 1 provides sample size for all classifications for individuals that provided neuroimaging data.

Of the 117 children who participated in both the infant temperament assessment and at least one MRI visit, 21 (18 %) were excluded from analyses because they aborted before MR images were collected ($n = 8$), scanned on a different scanner ($n = 2$), or had excessive motion artifact resulting in poor segmentations ($n = 11$). Thus, a total of 96 individuals provided useable MRI data. 21 individuals were characterized in infancy as non-reactive and were excluded from the focal analyses. Thus, the final sample for the focal analyses was 75 (41 negative reactive, 34 positive reactive; although see Supplemental analyses for data on the 21 non-reactive children). Children whose data were useable after quality control did not differ from children who had un-useable brain imaging data with respect to sex, maternal education, reactivity group, or anxiety at age 12 (*ps* > 0.537). A subset of children came in for MRI visits at both ages 10 and 12 ($n = 49$), but only $n = 38$ children (77.55 %) had useable data at both time points (See Table S1 for a breakdown by reactivity group). Among the children who came in for imaging visits, participant age at the 4-month, 10-year or 12-year visits did not differ as a function of reactivity group (p>.581). Supplemental Tables 2 and 3 summarizes sample demographics and attrition.

2.2. Measures

2.2.1. 4 Month reactivity

Infants were assessed for reactivity to novel visual and auditory stimuli at 4 months using the procedure described in Calkins et al., 1996. After testing, research assistants coded responses to the novel stimuli along three dimensions: motor arousal, positive affect, and negative affect. The frequency of positive affect, negative affect, and motor behavior was observed and summed (Fox et al., 2015). Positive affect reflected the frequency of the infant's smile and positive vocalizations. Negative affect reflected the frequency of cry, fuss, and negative verbal expressions. Motor affect reflected the frequency of arm movements, arm bursts, leg movements, leg bursts, arches, and hyperextensions. Interrater reliability has been previously published on this assessment (See Fox et al., 2001) and was between .78 and .86 for all measures.

2.3. Neuroimaging at age 10 and 12

2.3.1. Imaging data acquisition

High-resolution T1-weighted MRI images were acquired using a MPRAGE sequence on a 3 T MR750 GE scanner with a 32-channel head coil, with the following acquisition parameters: 256 × 176 × 256 voxels, 1 mm^3 isotropic, TR/TI/TE = 7700/425/3.42 ms, flip angle = 7°.

2.3.2. Image processing

Images were processed using FreeSurfer 5.3.0 (Dale et al., 1999; Fischl and Dale, 2000) (http://surfer.nmr.mgh.harvard.edu/),

producing surfaces delineating the inner (white) and outer (pial) limits of the cortical gray matter, and subcortical volume segmentation, which produced amygdala volume estimates. FreeSurfer's automatic segmentation was optimized using *adult* MRI data—not child data (Schoemaker et al., 2016). Thus, all data underwent thorough visual inspection, and several segmentation defects were noted (most often in insular cortex, occipital lobe, temporal lobe). Surfaces were visually inspected for defects and, where necessary, edited and regenerated. To extract reliable volume estimates, all images were processed using FreeSurfer's longitudinal stream (Reuter et al., 2012). This method creates a within-subject template (Reuter and Fischl, 2011) using inverse consistent (Reuter et al., 2010).FreeSurfer's processing steps (i.e., skull stripping, Talairach transforms, atlas registration, etc) are then initialized with common information from the within-subject template. This process has been shown to significantly increase reliability and statistical power (Reuter et al., 2012. Surfaces for each subject were then aligned to a common template fsaverage, resampled to a common resolution, and smoothed with a Gaussian filter of FWHM = 20 mm, then downsampled to a five-times recursively tessellated icosahedron ic5, which comprises 10,242 vertices per hemisphere. Exploratory analyses report measurements of cortical surface area and thickness at each vertex. These measures were calculated in native resolution.

2.4. Analytic approach

2.4.1. Preliminary analyses

Preliminary analyses tested whether any demographic variables differed as a function of reactivity. To do so, we used R v.3.3.2 (R Core Team, 2012). First, we tested whether ICV differed as a function of reactivity group and/or age. Next, we tested whether sex or maternal education differed as a function of reactivity group. To conduct all chi-squared analyses we used simulated *p*-values based on 2000 replicates (due to the small sample cell size). These three factors (ICV, sex, and maternal education) were controlled for in all subsequent models.

2.4.2. Focal analyses

Focal analyses examined longitudinal changes in amygdala volume as a function of infant temperament group (between-subjects factor: positive, negative) and time (within-subjects factor: age 10 and age 12). Given that the non-reactive group was comprised of only 7 individuals that provided longitudinal imaging data and 14 individuals that provided any cross-sectional data (See Table S2), we chose to exclude non-reactive individuals from our focal analysis (See Table S2). Thus, the sample size for our focal analysis was 75. We used R v.3.3.2 (R Core Team, 2012) and the lme4 package (Bates et al., 2012) to perform a linear mixed effects analysis because this method yields more reliable estimates than complete-case analyses when there is significant missing data (Krueger, 2004; Matta et al., 2018). The regression model tested included reactivity group, time, estimated intracranial volume (ICV), sex, maternal education, and the interaction of reactivity group by time as independent variables; to account for repeated measures, we used random intercepts for all subjects. To test the significance of this model, we conducted a likelihood ratio test comparing this model to an identical model without the interaction effect. Amygdala volume was tested separately for each hemisphere. To correct for laterality, significance criterion was set at $\alpha = .025$ (Bonferroni correction $\alpha = .05/2$).

2.4.3. Exploratory analyses

Previous studies have demonstrated a relation between temperament and cortical thickness across a variety of cortical regions associated with cognitive control, we also conducted whole brain exploratory analyses with cortical thickness and surface area as outcome measures, with the aim of facilitating comparison of our sample data with other reports in the literature. Two models were tested: one for cortical thickness and another with surface area as the outcome. The models included the same predictors reported in the focal analyses including reactivity, time, and the interaction between the two, while controlling for ICV, sex and maternal education. These whole brain analyses were conducted utilizing the tool Permutation Analysis of Linear Models (Winkler et al., 2014) (PALM), which, among other features, allows exact control over the familywise error rate (FWER) over the many thousand non-independent tests performed across the brain (one per vertex), as well as the computation, on the surfaces, of the threshold-free cluster enhancement (Smith and Nichols, 2009) (TFCE), a test statistic known to be more powerful for considering the spatial distribution of effects. We used 5000 permutations, along with the fit of a generalized Pareto distribution to the tail of the permutation distribution (Winkler et al., 2016). Correction of the FWER was considered all vertices of both hemispheres jointly and was corrected across both contrasts and modalities (i.e., area and thickness). Results were visualized using Blender version 2.77a (Kent, 2015).

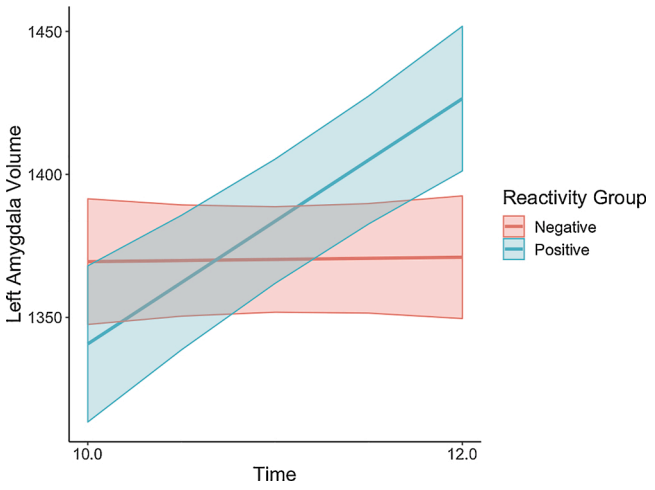


Fig. 1. Linear fit of (left amygdala) model results plotted as a function of group (negative reactive (n = 41) in red; positive reactive (n = 34) in blue). Error bars indicate one standard error.
(a) LEFT HEMISPHERE
(b) RIGHT HEMISPHERE

3. Results

3.1. Preliminary analyses

Prior to testing our focal models, we tested whether age at any visit, ICV, sex, or maternal education differed as a function of reactivity group. Results indicated that ICV did not differ as a function of reactivity group ($p > .270$) but increased as expected with age ($p < .037$). Next, we tested whether the gender distribution was roughly equivalent as a function of reactivity group. Results suggested there were more females in the negative reactive group than the positive reactive group ($\chi^2 = 5.28$, $p < .076$). There were also significant differences in maternal education: more infants in the negative reactive group had low maternal education (i.e., high school degree) relative to both non-reactive and positive reactive infants ($\chi^2 = 28.72$, $p < .0004$). Based on these results (See Table S3) and prior studies demonstrating that gender(Reiss et al., 1996) and social economic status (which is often indexed using maternal education; Noble et al., 2012) predict subcortical volume, we decided to control for these three factors in our focal models: ICV, gender, and maternal education.

3.2. Focal analyses

Focal analyses examined amygdala growth of children with history of negative compared to positive reactivity while controlling for intracranial volume, sex, and maternal education (i.e., to account for overall

Table 1
Fixed effects estimates predicting Left Amygdala volume.

Fixed Effects	Estimate	Std Error	t value	p-value
Intercept	−215.0	2.462e + 02	−0.872	0.386
ICV	.0007	1.121e − 04	6.520	1.49e-08***
Maternal Education -level 2	30.70	3.870e + 01	0.793	0.431
Maternal Education -level 3	40.42	4.281e + 01	0.944	0.349
Maternal Education -level 4	9.10	6.207e + 01	0.147	0.884
Sex-female	−37.88	3.130e + 01	−1.210	0.231
Time	42.90	1.487e + 01	2.885	0.006**
Negative Reactive	450.0	2.126e + 02	2.117	0.039*
Time* Negative Reactive	−42.12	1.892e + 01	−2.227	0.031*

Note: Maternal education level 2 corresponds to college graduates, level 3 corresponds to graduate/professional training, level 4 corresponds to other/no information.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

brain size and demographic differences between the groups). Following Bonferroni-correction there was a marginally significant interaction between reactivity group and time (age 10 & 12) on left amygdala volume (β=-42.12, $p < .031$), and the change in the variance explained was significantly greater than the model without the interaction effect ($\chi^2(1) = 5.27$, $p < .022$; AIC/BIC $_{\text{without interaction}} = 1310.3/1336.8$; AIC/BIC $_{\text{with interaction}} = 1307.0/1336.2$). Specifically, in the positive reactive group, left amygdala volume increased over time, unlike the negative reactive group, for whom left amygdala volume did not change (See Fig. 1 & Table 1 for fixed effects estimates). Supplemental analyses demonstrate that this finding remains marginally significant when estimating infant reactivity using a continuous, rather than dichotomous, reactivity measure. Additionally, in line with the best practices outlined in Vijayakumar, 2018, we replicate our models without controlling for ICV in the Supplement. There was no interaction between time and group on right amygdala volume ($p < .366$).

3.3. Exploratory analyses

As an exploratory follow-up, we examined whether cortical thickness and/or surface area differed as a function of reactivity, time, or their interaction. Replicating large longitudinal studies of childhood brain development, we found widespread decreases in cortical thickness across the entire brain between ages 10 and 12, independent of temperament (Fig. 2). Whole-brain analyses revealed no main effect of reactivity and no time by reactivity interaction.

4. Discussion

The amygdala is central to interpreting and responding to perceived threats (LeDoux, 2003) and has long been hypothesized as a neurobiological substrate of infant negative reactivity (Kagan et al., 1988). This is the first study to demonstrate that both positive and negative reactivity are associated with structural differences in the amygdala—particularly, amygdala growth 12 years after the identification of this behavioral phenotype. This study is among few that have followed children longitudinally to track associations between amygdala volume and behavior (Barnea-Goraly et al., 2014; Mosconi et al., 2009; Nordahl, 2012; Schumann, 2004; Cynthia Mills Schumann, Barnes et al., 2009). These effects could not be explained by differences in intracranial volume or demographic differences in maternal education between reactivity groups.

Brain development is a non-linear process (Goddings et al., 2014; Gogtay et al., 2004; Mills et al., 2016; Wierenga et al., 2014). Early in childhood the amygdala grows rapidly (Nordahl, 2012; Payne et al., 2010). Then, around age 10–12 amygdala growth peaks and subsequently stagnates with slower growth from adolescence into adulthood (Payne et al., 2010; Pechtel et al., 2014; Wierenga et al., 2014). The results of this study demonstrate that these developmental trajectories differ as a function of early temperament, with the negative reactive individuals showing a pattern of relatively little growth in the left amygdala during middle childhood. In contrast, positive reactive

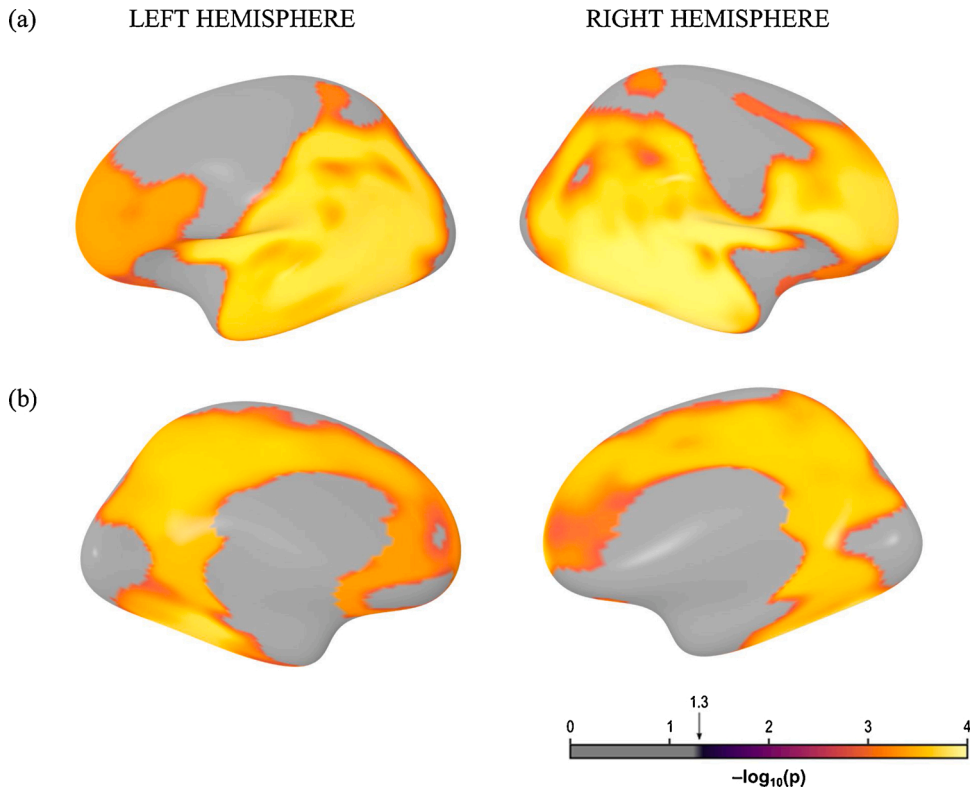


Fig. 2. Cortical thickness decreases as a function of time (lateral view (a); medial view (b)).

individuals exhibited continued amygdala growth over middle childhood; a pattern that is in line with previous reports of amygdala development in typically developing children (Goddings et al., 2014; Pechtel et al., 2014; Wierenga et al., 2014). It remains unknown whether either pattern of growth has behavioral benefits. In the negative reactive group, this finding might suggest a pattern of accelerated maturation (i.e., entering into the less growth early). Paired with data demonstrating that negative reactive infants show an exaggerated stress response (Fox, 1989; Kagan et al., 1988), one possible interpretation is that stress responsivity influences the maturation of emotion circuits in the brain. Future work integrating functional and structural MRI could provide novel insight into this question.

The strengths of this study lie in its longitudinal design, rich behavioral phenotyping, and thorough examination of brain morphometry. However, it is important to note that the sample size in each reactivity group is relatively small, so it is possible that we are overestimating the size of our effects[1] . While we find the same patterns when we operationalize reactivity continuously, rather than categorically (see Supplementary Materials), replication of these effects in larger samples from independent research groups is needed. Second, results did not withstand Bonferroni-correction; thus, it is important that this pattern be replicated in future studies. Third, although temperament is measured in infancy and brain morphometry is measured in middle childhood, these findings are correlational and as such do not provide direct evidence of a cause-effect relation. Possibly, differences in brain morphometry at the beginning of life predict which individuals will demonstrate high negative temperament in infancy. To determine directionality, we would need additional studies that assess both infant brain morphometry and behavior longitudinally. Furthermore, the current study focuses on brain structure only during middle childhood. Further work is needed to determine whether these differences in subcortical volume growth persist into adolescence and adulthood.

In conclusion, we provide novel evidence on the neurobiology of early temperament. These data demonstrate that infant reactivity is related to distinct patterns of brain development 12 years after the emergence of this phenotype. Such findings raise questions about the functional significance of slowed volume growth in regions of the brain sensitive to threat and emotion. Together, this work sheds new light on the neurodevelopment of children with a history of infant reactivity and adds to our understanding of the neural correlates of early temperament.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Acknowledgements

This research was supported by the NIMH Intramural Research Program (ZIAMH002781-15, NCT00018057; CAF, JFS, DP, AW, EL, DSP) and NIH grant U01MH093349 (NAF). We would also like to acknowledge Jennifer Pacheco for consulting on the imaging analyses.

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.2020.100776.

References

Barnea-Goraly, N., Frazier, T.W., Piacenza, L., Minshew, N.J., Keshavan, M.S., Reiss, A. L., Hardan, A.Y., 2014. A preliminary longitudinal volumetric MRI study of amygdala and hippocampal volumes in autism. Prog. Neuropsychopharmacol. Biol. Psychiatry 48, 124–128. https://doi.org/10.1016/j.pnpbp.2013.09.010.

Bates, D., Maechler, M., Bolker, B., 2012. lme4: Linear Mixed-Effects Models Using S4 Classes. R Package Version 0.999375-42. 2011. Reference Source.

Bauman, M.D., Lavenex, P., Mason, W.A., Capitanio, J.P., Amaral, D.G., 2004. The development of social behavior following neonatal amygdala lesions in Rhesus monkeys. J. Cogn. Neurosci. 16 (8), 1388–1411. https://doi.org/10.1162/0898929042304741.

Calkins, S.D., Fox, N.A., Marshall, T.R., 1996. Behavioral and physiological antecedents of inhibited and uninhibited behavior. Child dev. 67 (2), 523–540. https://doi.org/10.2307/1131830.

Clauss, J.A., Seay, A.L., VanDerKlok, R.M., Avery, S.N., Cao, A., Cowan, R.L., Benningfield, M.M., Blackford, J.U., 2014. Structural and functional bases of inhibited temperament. Soc. Cogn. Affect. Neurosci. 9 (12), 2049–2058. https://doi.org/10.1093/scan/nsu019.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. NeuroImage 9, 179–194.

Durston, S., Hulshoff Pol, H.E., Casey, B.J., Giedd, J.N., Buitelaar, J.K., Van Engeland, H., 2001. Anatomical MRI of the developing human brain: what have we learned? J. Am. Acad. Child Adolesc. Psychiatry 40 (9), 1012–1020. https://doi.org/10.1097/00004583-200109000-00009.

Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. 97 (20), 11050–11055. https://doi.org/10.1073/pnas.200033797.

Fox, N.A., 1989. Psychophysiological correlates of emotional reactivity during the first year of life. Dev. Psychol. 25 (3), 364–372.

Fox, N.A., Henderson, H.A., Rubin, K.H., Calkins, S.D., Schmidt, L.A., 2001. Continuity and discontinuity of behavioral inhibition and exuberance: psychophysiological and behavioral influences across the first four years of life. Child Dev. 72 (1), 1–21. https://doi.org/10.1111/1467-8624.00262.

Fox, N.A., Henderson, H.A., Marshall, P.J., Nichols, K.E., Ghera, M.M., 2005. Behavioral inhibition: linking biology and behavior within a developmental framework. Annu. Rev. Psychol. 56 (1), 235–262. https://doi.org/10.1146/annurev.psych.55.090902.141532.

Fox, N.A., Snidman, N., Haas, S.A., Degnan, K.A., Kagan, J., 2015. The relations between reactivity at 4 months and behavioral inhibition in the second year: replication across three independent samples. Infancy 20 (1), 98–114. https://doi.org/10.1111/infa.12063.

Giedd, J.N., 2008. The teen brain: insights from neuroimaging. J. Adolesc. Health 42 (4), 335–343. https://doi.org/10.1016/j.jadohealth.2008.01.007.

Giedd, J.N., Vaituzis, A.C., Hamburger, S.D., Lange, N., Rajapakse, J.C., Kaysen, D., Vauss, Y.C., Rapoport, J.L., 1996. Quantitative MRI of the temporal lobe, amygdala, and hippocampus in normal human development: ages 4-18 years. J. Comp. Neurol. 366 (2), 223–230. https://doi.org/10.1002/(SICI)1096-9861(19960304)366:2<223::AID-CNE3>3.0.CO;2-7.

Giedd, J.N., Clasen, L.S., Lenroot, R., Greenstein, D., Wallace, G.L., Ordaz, S., Molloy, E. A., Blumenthal, J.D., Tossell, J.W., Stayer, C., Samango-Sprouse, C.A., Shen, D., Davatzikos, C., Merke, D., Chrousos, G.P., 2006. Puberty-related influences on brain development. Mol. Cell. Endocrinol. 254–255, 154–162. https://doi.org/10.1016/j.mce.2006.04.016.

Goddings, A.-L., Mills, K.L., Clasen, L.S., Giedd, J.N., Viner, R.M., Blakemore, S.-J., 2014. The influence of puberty on subcortical brain development. NeuroImage 88, 242–251. https://doi.org/10.1016/j.neuroimage.2013.09.073.

Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C., Nugent, T. F., Herman, D.H., Clasen, L.S., Toga, A.W., Rapoport, J.L., Thompson, P.M., 2004. Dynamic mapping of human cortical development during childhood through early adulthood. Proc. Natl. Acad. Sci. 101 (21), 8174–8179. https://doi.org/10.1073/pnas.0402680101.

Kagan, J., Reznick, J., Snidman, N., 1988. Biological bases of childhood shyness. Science 240 (4849), 167–171. https://doi.org/10.1126/science.3353713.

Kent, B.R., 2015. 3D Scientific Visualization with Blender. Morgan & Claypool Publishers.

Krueger, C., 2004. A comparison of the general linear mixed model and repeated measures ANOVA using a dataset with multiple missing data points. Biol. Res. Nurs. 6 (2), 151–157. https://doi.org/10.1177/1099800404267682.

LeDoux, J., 2003. The emotional brain, fear, and the amygdala. Cell. Mol. Neurobiol. 12.

Matta, T.H., Flournoy, J.C., Byrne, M.L., 2018. Making an unknown unknown a known unknown: missing data in longitudinal neuroimaging studies. Dev. Cogn. Neurosci. 33, 83–98. https://doi.org/10.1016/j.dcn.2017.10.001.

Mills, K.L., Goddings, A.-L., Herting, M.M., Meuwese, R., Blakemore, S.-J., Crone, E.A., Dahl, R.E., Güroğlu, B., Raznahan, A., Sowell, E.R., Tamnes, C.K., 2016. Structural brain development between childhood and adulthood: convergence across four longitudinal samples. NeuroImage 141, 273–281. https://doi.org/10.1016/j.neuroimage.2016.07.044.

Mosconi, M.W., Cody-Hazlett, H., Poe, M.D., Gerig, G., Gimpel-Smith, R., Piven, J., 2009. Longitudinal study of amygdala volume and joint attention in 2- to 4-year-old children with autism. Arch. Gen. Psychiatry 66 (5), 509. https://doi.org/10.1001/archgenpsychiatry.2009.19.

Noble, K.G., Houston, S.M., Kan, E., Sowell, E.R., 2012. Neural correlates of socioeconomic status in the developing human brain: neural correlates of socioeconomic status. Dev. Sci. 15 (4), 516–527. https://doi.org/10.1111/j.1467-7687.2012.01147.x.

Nordahl, C.W., 2012. Increased rate of amygdala growth in children aged 2 to 4 years with autism spectrum disorders: a longitudinal study. Arch. Gen. Psychiatry 69 (1), 53. https://doi.org/10.1001/archgenpsychiatry.2011.145.

Payne, C., Machado, C.J., Bliwise, N.G., Bachevalier, J., 2010. Maturation of the hippocampal formation and amygdala in Macaca mulatta: a volumetric magnetic

[1] Furthermore, these results are not corrected for laterality.

resonance imaging study. Hippocampus 20 (8), 922–935. https://doi.org/10.1002/hipo.20688.
Pechtel, P., Lyons-Ruth, K., Anderson, C.M., Teicher, M.H., 2014. Sensitive periods of amygdala development: the role of maltreatment in preadolescence. NeuroImage 97, 236–244. https://doi.org/10.1016/j.neuroimage.2014.04.025.
R Core Team, 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
Reiss, A.L., Abrams, M.T., Singer, H.S., Ross, J.L., Denckla, M.B., 1996. Brain development, gender and IQ in children: a volumetric imaging study. Brain 119 (5), 1763–1774. https://doi.org/10.1093/brain/119.5.1763.
Reuter, M., Fischl, B., 2011. Avoiding asymmetry-induced bias in longitudinal image processing. NeuroImage 57 (1), 19–21. https://doi.org/10.1016/j.neuroimage.2011.02.076.
Reuter, M., Rosas, H.D., Fischl, B., 2010. Highly accurate inverse consistent registration: a robust approach. NeuroImage 53 (4), 1181–1196. https://doi.org/10.1016/j.neuroimage.2010.07.020.
Reuter, M., Schmansky, N.J., Rosas, H.D., Fischl, B., 2012. Within-subject template estimation for unbiased longitudinal image analysis. NeuroImage 61 (4), 1402–1418. https://doi.org/10.1016/j.neuroimage.2012.02.084.
Schoemaker, D., Buss, C., Head, K., Sandman, C.A., Davis, E.P., Chakravarty, M.M., Gauthier, S., Pruessner, J.C., 2016. Hippocampus and amygdala volumes from magnetic resonance images in children: assessing accuracy of FreeSurfer and FSL against manual segmentation. NeuroImage 129, 1–14. https://doi.org/10.1016/j.neuroimage.2016.01.038.
Schumann, C.M., 2004. The amygdala is enlarged in children but not adolescents with autism; the Hippocampus is enlarged at all ages. J. Neurosci. 24 (28), 6392–6401. https://doi.org/10.1523/JNEUROSCI.1297-04.2004.
Schumann, C.M., Barnes, C.C., Lord, C., Courchesne, E., 2009. Amygdala enlargement in toddlers with autism related to severity of social and communication impairments. Biol. Psychiatry 66 (10), 942–949. https://doi.org/10.1016/j.biopsych.2009.07.007.
Schwartz, C.E., Kunwar, P.S., Greve, D.N., Kagan, J., Snidman, N.C., Bloch, R.B., 2012. A phenotype of early infancy predicts reactivity of the amygdala in male adults. Mol. Psychiatry 17 (10), 1042–1050. https://doi.org/10.1038/mp.2011.96.
Smith, S., Nichols, T., 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. NeuroImage 44 (1), 83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061.
Smith, A.R., White, L.K., Leibenluft, E., McGlade, A.L., Heckelman, A.C., Haller, S.P., Buzzell, G.A., Fox, N.A., Pine, D.S., 2019. The heterogeneity of anxious phenotypes: neural responses to errors in treatment-seeking anxious and behaviorally inhibited youths. J. Am. Acad. Child Adolesc. Psychiatry. https://doi.org/10.1016/j.jaac.2019.05.014.
Snidman, Nancy, Kagan, Jerome, Riordan, Linda, Shannon, Daniel C, 1995. Cardiac function and behavioral reactivity during infancy. Psychophysiology 32 (3), 199–207. https://doi.org/10.1111/j.1469-8986.1995.tb02949.x.
Stifter, C.A., Fox, N.A., 1990. Infant reactivity: physiological correlates of newborn and 5-month temperament. Dev. Psychol. 26 (4), 582–588.
Sylvester, C.M., Barch, D.M., Harms, M.P., Belden, A.C., Oakberg, T.J., Gold, A.L., White, L.K., Benson, B.E., Troller-Renfree, S., Degnan, K.A., Henderson, H.A., Luby, J.L., Fox, N.A., Pine, D.S., 2016. Early childhood behavioral inhibition predicts cortical thickness in adulthood. J. Am. Acad. Child Adolesc. Psychiatry 55 (2), 122–129. https://doi.org/10.1016/j.jaac.2015.11.007 e1.
Vijayakumar, N., et al., 2018. Structural brain development: A review of methodological approaches and best practices. Developmental cognitive neuroscience 33, 129–148. https://doi.org/10.1016/j.dcn.2017.11.008.
Wierenga, L., Langen, M., Ambrosino, S., van Dijk, S., Oranje, B., Durston, S., 2014. Typical development of basal ganglia, hippocampus, amygdala and cerebellum from age 7 to 24. NeuroImage 96, 67–72. https://doi.org/10.1016/j.neuroimage.2014.03.072.
Winkler, A.M., Ridgway, G.R., Webster, M.A., Smith, S.M., Nichols, T.E., 2014. Permutation inference for the general linear model. NeuroImage 92, 381–397. https://doi.org/10.1016/j.neuroimage.2014.01.060.
Winkler, A.M., Ridgway, G.R., Douaud, G., Nichols, T.E., Smith, S.M., 2016. Faster permutation inference in brain imaging. NeuroImage 141, 502–516. https://doi.org/10.1016/j.neuroimage.2016.05.068.
Yang, R.J., Mozhui, K., Karlsson, R.-M., Cameron, H.A., Williams, R.W., Holmes, A., 2008. Variation in mouse basolateral amygdala volume is associated with differences in stress reactivity and fear learning. Neuropsychopharmacology 33 (11), 2595–2604. https://doi.org/10.1038/sj.npp.1301665.