Title

Rhesus infant nervous temperament predicts peri-adolescent central amygdala metabolism & behavioral inhibition measured by a machine-learning approach

Journal

Translational Psychiatry, 14(1)

ISSN

2158-3188

Authors

Holley, D

Campos, LJ

Drzewiecki, CM

et al.

Publication Date

2024

DOI

10.1038/s41398-024-02858-3

Peer reviewed

ARTICLE OPEN

Check for updates

Rhesus infant nervous temperament predicts peri-adolescent central amygdala metabolism & behavioral inhibition measured by a machine-learning approach

D. Holley[1,2,4], L. J. Campos[1,2,4], C. M. Drzewiecki[2], Y. Zhang[3], J. P. Capitanio[1,2] and A. S. Fox[1,2]✉

Anxiety disorders affect millions of people worldwide and impair health, happiness, and productivity on a massive scale. Developmental research points to a connection between early-life behavioral inhibition and the eventual development of these disorders. Our group has previously shown that measures of behavioral inhibition in young rhesus monkeys (*Macaca mulatta*) predict anxiety-like behavior later in life. In recent years, clinical and basic researchers have implicated the central extended amygdala (EAc)—a neuroanatomical concept that includes the central nucleus of the amygdala (Ce) and the bed nucleus of the stria terminalis (BST)—as a key neural substrate for the expression of anxious and inhibited behavior. An improved understanding of how early-life behavioral inhibition relates to an increased lifetime risk of anxiety disorders—and how this relationship is mediated by alterations in the EAc—could lead to improved treatments and preventive strategies. In this study, we explored the relationships between infant behavioral inhibition and peri-adolescent defensive behavior and brain metabolism in 18 female rhesus monkeys. We coupled a mildly threatening behavioral assay with concurrent multimodal neuroimaging, and related those findings to various measures of infant temperament. To score the behavioral assay, we developed and validated *UC-Freeze*, a semi-automated machine-learning (ML) tool that uses unsupervised clustering to quantify freezing. Consistent with previous work, we found that heightened Ce metabolism predicted elevated defensive behavior (i.e., more freezing) in the presence of an unfamiliar human intruder. Although we found no link between infant-inhibited temperament and peri-adolescent EAc metabolism or defensive behavior, we did identify infant nervous temperament as a significant predictor of peri-adolescent defensive behavior. Our findings suggest a connection between infant nervous temperament and the eventual development of anxiety and depressive disorders. Moreover, our approach highlights the potential for ML tools to augment existing behavioral neuroscience methods.

Translational Psychiatry (2024)14:148 ; https://doi.org/10.1038/s41398-024-02858-3

INTRODUCTION

Anxiety disorders are among the most prevalent psychiatric conditions, affecting an estimated one in four people during their lifetime [1–3]. These disorders are frequently comorbid with a wide range of other psychopathologies, including depression, as well as alcohol- and substance-abuse disorders [4–7], and are considerably more prevalent in women than in men [8]. Although a complete understanding of the etiology of these disorders remains elusive, researchers have begun to characterize the risk factors that predict their onset. Identifying and investigating these risk factors promises to yield an improved understanding of anxiety disorders and will likely contribute to their treatment and prevention.

An extremely inhibited or anxious temperament during childhood increases the risk of developing an anxiety disorder later in life [9–13]. Developmental researchers often evaluate inhibited or anxious temperaments by measuring a child's behavioral inhibition (BI)—that is, their reactivity to novel stimuli, unfamiliar situations, and strangers [10, 14–16]. Some aspects of BI emerge early in life and are trait-like and stable; for instance, a 4-month-old infant's aversion to unfamiliar stimuli predicts composite BI measured years later [17, 18]. Although high BI often predicts the eventual development of anxiety disorders [9, 19], researchers do not fully understand how infant temperament relates to childhood or adolescent BI, or its associated brain function. Because nonhuman primates (NHP) have a protracted development period, they are well-suited to build this understanding.

Thanks to our relatively recent evolutionary divergence, NHPs and humans share a variety of socioemotional, anatomical, and genetic similarities that facilitate high-impact translational research, notably including an elaborated prefrontal cortex [20–23] Because of this, NHPs are excellent models for studying the mechanisms of early-life risk inherent to a range of disorders [24–30]. To support such studies, researchers at the California National Primate Research Center (CNPRC) have, over the past 2 decades, evaluated over 5000 infant (i.e., 3- to 4-month-old) NHPs as part of its *BioBehavioral Assessment program* (BBA)—a 25-hour battery that catalogs each animal's physiological reactivity, emotionality, and temperament [31]. One of the temperament

[1]University of California, Department of Psychology, Davis, CA, USA. [2]California National Primate Research Center, Davis, CA, USA. [3]Columbia University, Department of Statistics, New York, NY, USA. [4]These authors contributed equally: D. Holley, L. J. Campos. ✉email: dfox@ucdavis.edu

Received: 27 October 2022 Revised: 21 February 2024 Accepted: 1 March 2024
Published online: 15 March 2024

measures is based on behavior; four others are based on human handlers' ratings of trait-like qualities [32, 33] and are similar to evaluations of BI in children [15, 16, 34]. These infant measurements complement measures of BI and anxious temperament in adult and adolescent rhesus monkeys (*Macaca mulatta*) and are thought to reflect a trait-like inhibited temperament defined by an enduring tendency to avoid novel and potentially threatening stimuli and situations [13, 35–40].

Investigations into the neural substrates of anxiety disorders and BI in humans [11, 41–46], as well as inhibited temperament and BI in NHPs [35, 38, 47–52], have implicated a distributed network of brain regions. Notably, this network includes the central extended amygdala (EAc): a neuroanatomical concept that encompasses the central nucleus of the amygdala (Ce) and the bed nucleus of the stria terminalis (BST). The EAc is central to threat processing [53–57] and is well-positioned to orchestrate adaptive defensive physiology and behavior [45, 48, 52, 53, 58, 59]. A range of sensory, evaluative, and contextual inputs converge on the EAc, which projects downstream effector regions to initiate these defensive responses [13, 53, 59, 60]. The EAc plays a role in the integration of emotion-relevant signals and produces scaled behavioral responses to a variety of stimuli—including uncertain threat stimuli, which reliably elicit adaptive defensive responses like freezing [61, 62]. Neuroimaging studies highlight the EAc's role in threat responding: A study of 592 peri-adolescent rhesus monkeys from the Wisconsin National Primate Research Center (WNPRC) and the Harlow Center for Biological Psychology, for example, linked individual differences in anxious temperament to variation in glucose metabolism in both the Ce and BST during exposure to an uncertain threat assay, such that more anxiety-like behaviors predicted increased metabolism in those regions [37]. Additionally, this study found metabolism within different components of the EAc to be differentially sensitive to heritable and non-heritable influences. Metabolism in the BST was co-inherited with individual differences in freezing in response to a potential uncertain threat, whereas metabolism in the Ce was not [37, 63]. This raises the intriguing possibility that Ce metabolism may be especially plastic and represent the environmental contributions to the risk of developing anxiety disorders. Notably, WNPRC animals are raised in small, indoor groups. By comparison, CNPRC animals are raised in large, outdoor, naturalistic colonies, and thus can experience a broader range of socioemotional contexts. To maximally advance our understanding of inhibited temperament, its neural substrates, and its relation to the progression of BI across different early-life environments, it is critical to standardize the methods for cross-facility replication. The current gold standard used to measure defensive behaviors in NHPs is hand scoring, during which trained researchers watch video recordings of animals placed in mildly threatening contexts and denote the time, type, and duration of behaviors of interest, such as freezing episodes. Although hand scoring has been instrumental to our understanding of NHP behavior, it presents challenges to replicability and can demand large time commitments from expert-trained behavioral coders. The rise of computing speed, power, and availability presents an opportunity to develop tools that scale easily and improve study replicability. To aid in the replicable assessment of inhibited temperament in NHPs, we developed and validated *UC-Freeze*, a semi-automated machine-learning (ML) approach that scores freezing behavior via unsupervised clustering (code available upon request).

Here, we assessed brain metabolism and used UC-Freeze to objectively score freezing in 18 peri-adolescent female rhesus monkeys during exposure to an uncertain threat (i.e., a human intruder). We analyzed the relationship between infant measures of BI (and, in exploratory analyses, temperament), and concurrent measures of peri-adolescent BI (i.e., freezing) and brain metabolism (Fig. 1a). We hypothesized that alterations in the EAc would be associated with infant and peri-adolescent BI.

METHODS

Animals and selection procedure

Twenty peri-adolescent female rhesus monkeys (*M. mulatta*, *M* [SD] = 2.71 years [.44]) that previously underwent BBA testing during infancy (3–4 months) were selected from a pool of ninety-eight potential animals using a stratified sampling procedure, in which one animal was selected from each of 20 uniformly distributed bins based on BBA inhibited temperament scores (Fig. 1b). The stratified selection procedure yielded a subject pool that captured the full spectrum of variation in 3–4-month-old inhibited temperament. Because females are at increased risk of developing anxiety and depressive disorders as they transition to adolescence [8, 64, 65], in this study, we focused specifically on females. We subjected each animal to the NEC-FDG paradigm (described in detail below) and scored their behavior with UC-Freeze. Two subjects were excluded from our analyses due to problems with video capture that rendered their videos unusable, making the final number of subjects $n = 18$. A power analysis revealed that, in our $n = 18$ subjects, we had ~80% power to identify a correlation that accounted for ~36% variance (R Studio version 1.0.153's *pwr* package). All housing and experimental procedures were conducted per guidelines set by the UC Davis Institutional Animal Care and Use Committee.

Infant BioBehavioral Assessment

The BBA is a 25-hour battery of emotional, cognitive, and biological assessments that evaluate qualities like resilience to mild challenges, willingness to interact with novel objects, memory, hypothalamic-pituitary-adrenal system regulation, and hematology [33, 66]. CNPRC animals undergo BBA testing during infancy (i.e., between 3 and 4 months of age), and most live the majority of their lives in large, outdoor colonies of roughly 100 conspecifics. This approach has been described in detail elsewhere [33] and has enabled CNPRC researchers to investigate relationships between various infant measures and the eventual emergence of disorder-relevant phenotypes in naturalistic socio-environmental settings [30, 67–69].

Relevant to this study, the BBA yields an inhibited temperament (IT) score for each animal (described in [31, 33, 47]) based on four factors: *Activity* in the first 15 minutes of day 1 and a period during day 2, and *Emotionality* during those same time points. These factors were previously identified through the factor analysis of roughly 2000 animals [70]. *Activity* includes time locomoting; time NOT hanging from the top or side of the cage; rate of environmental exploration; and whether the animal drank water*, ate food*, or was observed crouching in the cage* (* = dichotomized due to rarity). *Emotionality* includes the animal's rates of cooing and barking, as well as whether the animal lipsmacked*, displayed threats*, or scratched* (* = dichotomized due to rarity). Each animal's early-life inhibition score was calculated as the mean of its *z* scored day 1 and day 2 *Activity* and *Emotionality*.

At the end of BBA testing, before each animal was returned to its mother, the technician who administered testing rated each animal on four composite measures of trait-like infant temperament: *vigilance*, *nervousness*, *confidence*, and *gentleness* (see [33], for a full description of the BBA's temperament ratings). These measures are intended to accumulate across the full 25-hour testing period, and reflect an expert primatologist's cumulative assessment of the animal, akin to teacher or experimenter ratings in studies of children.

NEC-FDG paradigm: measuring peri-adolescent behavior & brain metabolism

To evaluate the relationship between infant measures and peri-adolescent defensive behaviors, we used the well-validated no-eye-contact condition (NEC) of the human intruder paradigm [37, 71]. In the NEC context, a human intruder enters the room and presents their profile to the animal while making no-eye contact. Integrated brain metabolism during the NEC was measured using [^{18}F]fludeoxyglucose (FDG) positron emission tomography (PET). Specifically, each animal was injected with FDG immediately prior to behavioral testing and then placed in a test-cage for exposure to the 30-minute NEC context. Immediately after exposure, animals were anesthetized for PET scanning (Fig. 1a). Because of the timecourse of FDG uptake, this paradigm is ideal for identifying integrated brain metabolic differences between individuals during threat processing.

FDG-PET and MRI acquisition. Animals received an intravenous injection (IV) of [^{18}F]fludeoxyglucose ($M = 7.449$ mCi, sd = 1.512 mCi) immediately before their 30-minute exposure to the NEC context, during which FDG

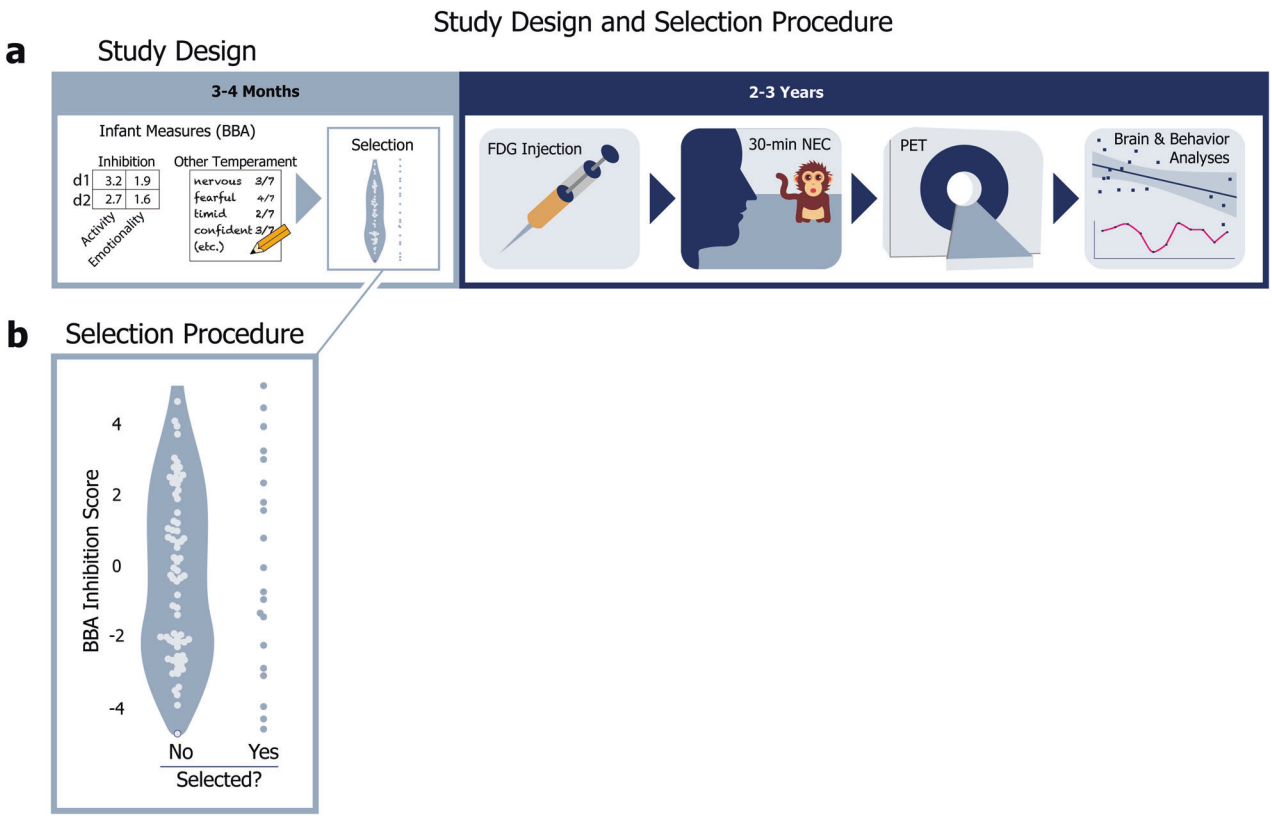

Fig. 1 Study design and selection procedure. a Study design: At 3–4 months old, all candidate animals were evaluated for a range of infant measures during the BBA. Relevant to our study, the BBA yields objective inhibition scores and other temperament ratings for each animal. At 2–3 years old, animals selected for our study were removed from their home colonies, injected with the radiotracer [^{18}F]fludeoxyglucose (FDG), and behaviorally assessed via a 30-minute no-eye-contact (NEC) condition of the human intruder paradigm, after which PET scans were administered to evaluate glucose metabolism during the NEC. **b** Selection procedure: The selection procedure for our study: 20 of 98 candidate peri-adolescent animals were initially selected based on a stratified sampling of 1 animal from each of 20 bins defined by *z* scored inhibited temperament scores, assessed during infancy as part of the BBA.

uptake occurred. After behavioral testing, animals were anesthetized, intubated, and transported to undergo a PET scan. Anesthesia was maintained using 1–2% isoflurane gas. FDG and attenuation scans were acquired using a piPET scanner (Brain Biosciences) located within the Multimodal Imaging Core at the CNPRC. Approximately 1 week after exposure to the NEC-FDG paradigm, anatomical 3D T1-weighted scans were obtained using a 3 T Siemens Skyra scanner, a dedicated rhesus 8-channel surface coil, with inversion-recovery, fast gradient echo prescription (TI/TR/TE/Flip/FOV/Matrix/Bandwidth:1100 ms/2500.0 ms/3.65 ms/7°/154 mm/512 × 512/240 Hz/Px) with whole brain coverage (480 slice encodes over 144 mm) reconstructed to 0.3 × 0.3 × 0.3 mm on the scanner).

FDG-PET and T1-MRI processing. All T1-weighted images were manually masked to exclude non-brain tissue by LJC. A study-specific T1 anatomical template was created using an iterative procedure with Advanced Normalization Tools [72, 73] (ANTS) in order to standardize our study-specific template for cross-facility replication, we first aligned each subject's T1 anatomical image to the National Institute of Mental Health Macaque Template (NMT) using a rigid body registration. The NMT template provides a common platform for the characterization of neuroimaging results across studies [74]. A non-linear registration was then performed using a symmetric diffeomorphic image registration and a .25 gradient step size; a pure cross-correlation with cost-function with a window radius 2 and weight 1; the similarity matrix was smoothed with sigma = 2; and the process was repeated at four increasingly fine levels of resolution with 30, 20, 20, and 5 iterations at each level, respectively. The average of all 20 individual subjects' T1 images in NMT space was computed and taken to be the study-mean. Similarly, the non-linear deformation field was also averaged and taken to be the deformation mean. The deformation mean was then inverted, and 15% of the deformation was applied to the study mean, to obtain the first iteration of the study-specific template. This process of averaging was repeated four times to obtain a final study-specific T1-weighted MRI template that matched the NMT template, and optimally reflected the brain morphometry of subjects of this study.

To get each subject's FDG-PET scan into this template space, each animal's FDG-PET image was aligned to its respective T1 anatomical image using a rigid body mutual information warp, and the transformation matrices from T1 to the study-specific NMT template space was then applied to the FDG-PET image to obtain PET images in NMT template space.

Once in standard space, the FDG-PET images were grand mean scaled to the average metabolism across the brain. To facilitate cross-animal comparisons, images were spatially smoothed using a 4-mm FWHM Gaussian kernel.

A priori regions of interest (ROI) were drawn for the motor cortex, as well as the two major components of the EAc, the Ce and the BST. All ROIs were manually drawn on the study-specific T1 template according to the Paxinos atlas [75] and verified by members of the team (LJC, DH, ASF).

UC-Freeze: an unsupervised-clustering approach to measuring freezing

To accurately and reproducibly assess freezing behavior during the NEC, we developed *UC-Freeze*: a semi-automated ML approach that uses unsupervised clustering to score freezing behavior. We targeted the definition of freezing used in previous NEC NHP studies [35, 37, 63]; that is, any period of 3 or more seconds during which the animal displayed a tense body posture and no movement, other than slow movements of the head.

UC-Freeze assesses freezing by first de-composing 30-fps full-motion video collected from each subject into individual frames. It then converts the frames to grayscale and vectorizes them such that a two-dimensional array of numeric values corresponding to various shades of gray represents

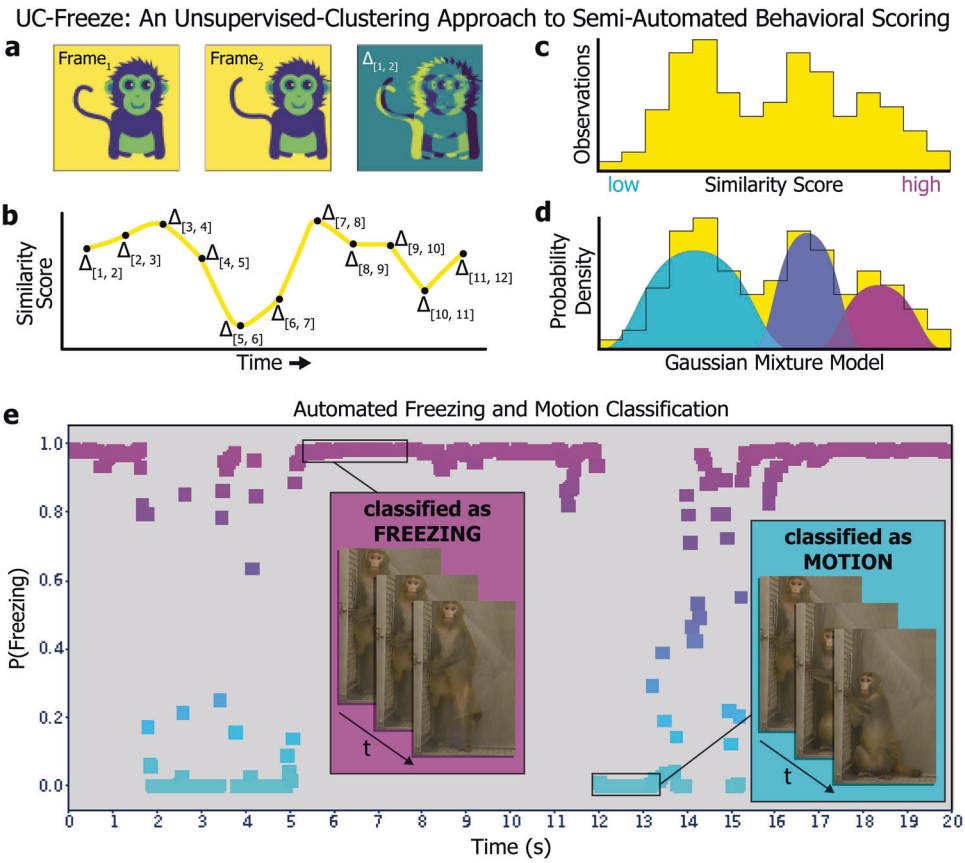

Fig. 2 UC-Freeze: an unsupervised-clustering approach to semi-automated behavioral scoring. a UC-Freeze decomposes 30-fps video into individual frames, converts each frame to grayscale, and computes the coefficient of determination value (i.e., r^2, or *similarity score*) for pairs of consecutive frames. **b** UC-Freeze next filters the similarity scores and arrays them along the timecourse of the video, so that the full timecourse is described as a series of similarity scores. The similarity scores are then passed into our unsupervised-clustering algorithm, which first **c** arranges them as a histogram before **d** computing a probability density function for the similarity scores by iterating over a randomly seeded one-dimensional GMM 300 times. In edge cases, the output of UC-Freeze can be manually overridden (see Methods). **e** Example output: UC-Freeze generates a unique model for each subject. Our program then queries each subject's similarity scores against the model's putative freezing distribution and recapitulates the timecourse of the video as a series of posterior probabilities indicating each similarity score's likelihood of belonging to that distribution. Finally, UC-Freeze uses a combination of anomaly-detection and thresholding operations to find 90-frame sequences during which the posterior probability of every score's membership in the freezing distribution's rightmost 25% density is 95% or greater, and classifies those events as freezing.

each frame's pixels. UC-Freeze next computes the coefficient of determination (r^2) between each pair of consecutive frames in order to quantify the degree of change between frames. We henceforth refer to these r^2 values as *similarity scores*. Lower similarity scores correspond to larger differences between frames, which suggest the animal is in motion (Fig. 2a). To ensure robustness against dropped video frames and video aliasing that can occur as a function of lighting, UC-Freeze then denoises the signal by substituting outlier similarity scores (thresholded as any score at or below an r^2 of .93) with the modal similarity score before passing the corrected vector through an adjustable median filter. (A 3-frame kernel was used in our analyses and is recommended as a default setting.) This process maintains sensitivity to the behavior of interest while buffering against frame-to-frame variation. These filtered similarity scores comprise the dataset that is passed into UC-Freeze's unsupervised clustering algorithm (Fig. 2b), which leverages one-dimensional Gaussian mixture modeling (GMM).

GMM is a form of unsupervised machine learning that assumes non-normal datasets are a mixture of standard normal distributions [76]. An advantage of GMM is its ability to cluster effectively by estimating probability densities of one-dimensional data, such as our subjects' similarity scores, before making probabilistic assignments to clusters based on probability-density estimates. GMM uses expectation maximization (EM) to estimate the underlying Gaussian distributions that comprise a dataset. UC-Freeze adds similarity scores to the model one at a time. Before each new similarity score is added, EM's *expectation* step estimates the model's probability distributions. After a new similarity score has been added, EM's *maximization* step refines the model's distributions based on the inclusion of the new data. These processes are repeated until the model is stable; that is, until the expectation step correctly predicts the maximization step. UC-Freeze iterates over each subject's similarity scores 300 times, each time randomizing the order of its input, to converge on a highly stable model unique to each subject (Fig. 2c, d). Here, we have calibrated UC-Freeze to cluster each animal's similarity scores into three Gaussian distributions: the lowest of which is assumed to reflect freezing; the highest of which is assumed to represent motion; and, between them, a third distribution captures similarity scores that are too ambiguous to confidently classify as either freezing or motion, which makes UC-Freeze more robust against spurious classifications (Fig. 2e). If the animal does not display a range of movement (i.e. both freezing and activity), the GMM approach can yield incorrect estimates for the distribution of each class of behavior. To address this, we use a semi-automated approach, where a human confirms the accurate classification of freezing/activity by watching video clips, and manually setting the point estimates for the 3 distributions, if necessary. Critically, the manual setting of point estimates does not need to be exact because it is only required in edge cases, when the animal does not freeze/locomote, and only requires that all frames be classified as freezing/locomotion. Thus, this approach is semi-automated, and ensures reliable data from all videos. This represents a significant improvement over previous approaches in mice, which require manually setting thresholds for all videos (e.g. ezTrack [77], VideoFreeze [78], and FreezeFrame (https://actimetrics.com/products/freezeframe/).

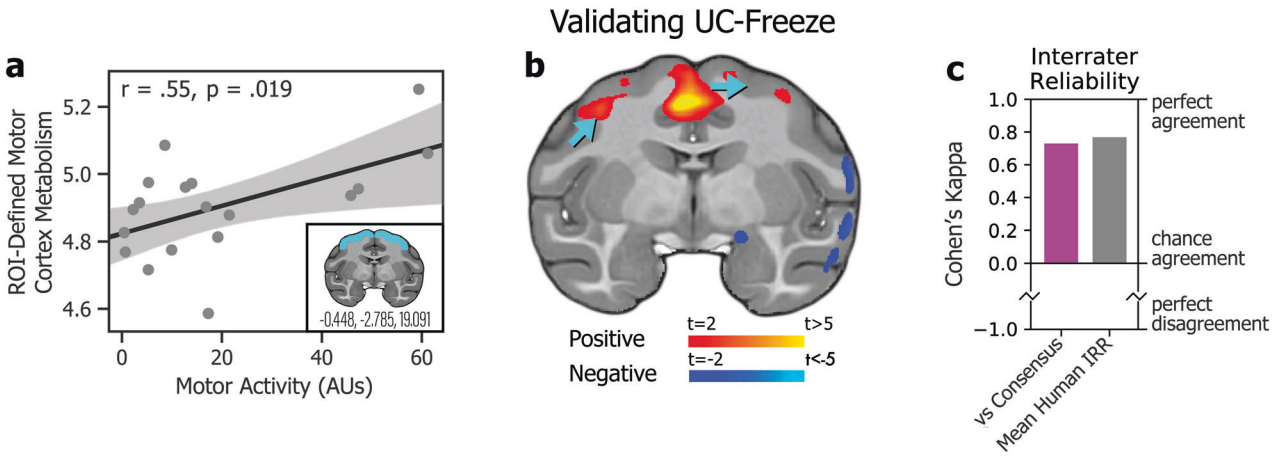

Fig. 3 Validating UC-Freeze. a Subjects' motor activity, as coded by UC-Freeze, significantly predicted integrated motor cortex metabolism in a brain region of interest defined a priori by the coordinates $x = -0.448$, $y = -2.785$, and $z = 19.091$ (inset). **b** A posteriori voxelwise analyses revealed subjects' motor activity as a significant predictor of integrated metabolism in regions of motor cortex (blue arrows). Together, these findings validate UC-Freeze's ability to recapitulate well-established brain-behavior relationships. **c** Measures of interrater reliability (IRR), Cohen's kappa [84], computed in the scoring of 80 3-second video segments for freezing, showed that UC-Freeze had "moderate to substantial" interrater agreement with each of three human raters; performed best when compared to human raters' consensus (magenta; kappa = 0.73, $p < 0.001$); and approximated mean human-vs-human IRR (gray; kappa = 0.77, $p < 0.001$), calculated by round-robin comparison. Together, these findings validate UC-Freeze as a reliable tool for scoring freezing in rhesus.

Once a model has been created for a subject, UC-Freeze recursively queries the model to determine the posterior probability of every similarity score's membership in its putative freezing and motion distributions. The posterior probabilities of every score's membership in the motion distribution are summed to compute an objective measure of an animal's *motor activity*. To objectively measure freezing, UC-Freeze then combines Tukey's anomaly-detection [79] with a thresholding operation to identify similarity scores that have a 95%-or-greater chance of belonging to the rightmost 25% of the freezing distribution's probability density. If 90 or more consecutive frames (i.e., 3 or more seconds) meet this criterion, UC-Freeze automatically classifies that segment as freezing (Fig. 2f). Lack of movement for <3-seconds is not classified as freezing. Freezing bouts are coded as uninterrupted freezing for more than 3 seconds. If an animal freezes for 4 seconds, moves, and then freezes for another 10 seconds, this would be classified as two freezing bouts (i.e., freezing bouts of 4 seconds and 10 seconds, respectively). Importantly, because this approach can fail if an animal very rarely or almost always freezes, this approach is not fully automated. Each video was reviewed, and the thresholding operation was manually adjusted in two cases to ensure edge cases did not disrupt the data (DH). In these cases, neither animal moved sufficiently for UC-Freeze to create a GMM with meaningfully dissimilar distributions.

Freezing established by human raters

To verify UC-Freeze, human raters coded a series of 80 video clips as freezing based on established criteria: any period of 3 or more seconds during which the animal displayed a tense body posture and no movement, other than slow movements of the head. Because UC-Freeze cannot distinguish between "tense" body posture, human raters also coded a series of 80 videos clips and classified animals as having "tense body posture" or not. As in previously published work, human raters watched videos at regular speed.

Statistical analyses

Pearson correlation coefficient (r) values describing the relationships between infant measures, and peri-adolescent measures were performed in Python v3.8.3 using the statsmodels module [80]. Freezing data were log-transformed to ensure data were normally distributed. Results of all relationships tested have been reported in the text and/or in Figs. 3–5. Analyses of interrater reliability (IRR) used to validate UC-Freeze were performed in Python v3.8.3 using the sklearn.metrics module [81]. An independent-samples t test to check for significant differences in animals' freezing behavior between the first and second halves of the NEC context was performed in Python v3.8.3 using the scipy.stats module [82].

Relationships between brain metabolism and other phenotypic measures were performed based on a priori ROIs in the motor cortex, Ce, and BST. FDG-PET values were extracted from each ROI (bilaterally), and the mean metabolism within each region was computed. To ensure our results were robust to a voxelwise approach, exploratory voxelwise analyses were also performed using FSL's nonparametric permutation inference tool *randomize* [83]. Voxelwise analyses were thresholded at $p < 0.05$, uncorrected.

RESULTS

Validation: comparing UC-Freeze to human raters

To validate UC-Freeze's ability to accurately and reliably score freezing behavior, we compared its semi-automated classifications to the manual classifications of three raters who had observed rhesus behaving in experimental and naturalistic conditions, and who were instructed on how to identify freezing in rhesus using criteria from previous publications [35, 37, 63]. We intentionally chose raters with a diversity of hand-scoring experience in order to model the challenges labs are likely to face as they seek to implement, or scale, studies that require hand scoring (i.e., the situations in which UC-Freeze would be most valuable). We randomly selected four animals for our analysis. From each animal's NEC video, we randomly generated 20 3-second video segments, 10 of which were classified by UC-Freeze as freezing, to yield a total of 40 freezing segments and 40 non-freezing segments. The raters were not given any information about the proportion of freezing vs non-freezing segments, and worked in isolation to score every segment as either freezing or non-freezing. We evaluated IRR by calculating Cohen's kappa values for UC-Freeze and each rater. In all three cases, UC-Freeze demonstrated "moderate to substantial" agreement with the rater, well above chance levels, and approximated human-level IRR (Fig. 3c). Next, we estimated UC-Freeze's *sensitivity* and *specificity*—that is, its ability to detect true positives (freezing) and negatives (non-freezing), respectively. Because there was variation between raters' scoring, we used consensus among raters for each video segment, calculated as the mode of scores, as a proxy for "true outcomes" (e.g., if Rater 1 scored freezing in a given video but Raters 2 and 3 did not, the "true outcome" was coded as non-freezing). Using this approach, UC-Freeze exhibited 84% sensitivity and 89% specificity. Finally, to evaluate pairwise internal reliability we calculated the mean Cohen's kappa derived from pairs of raters

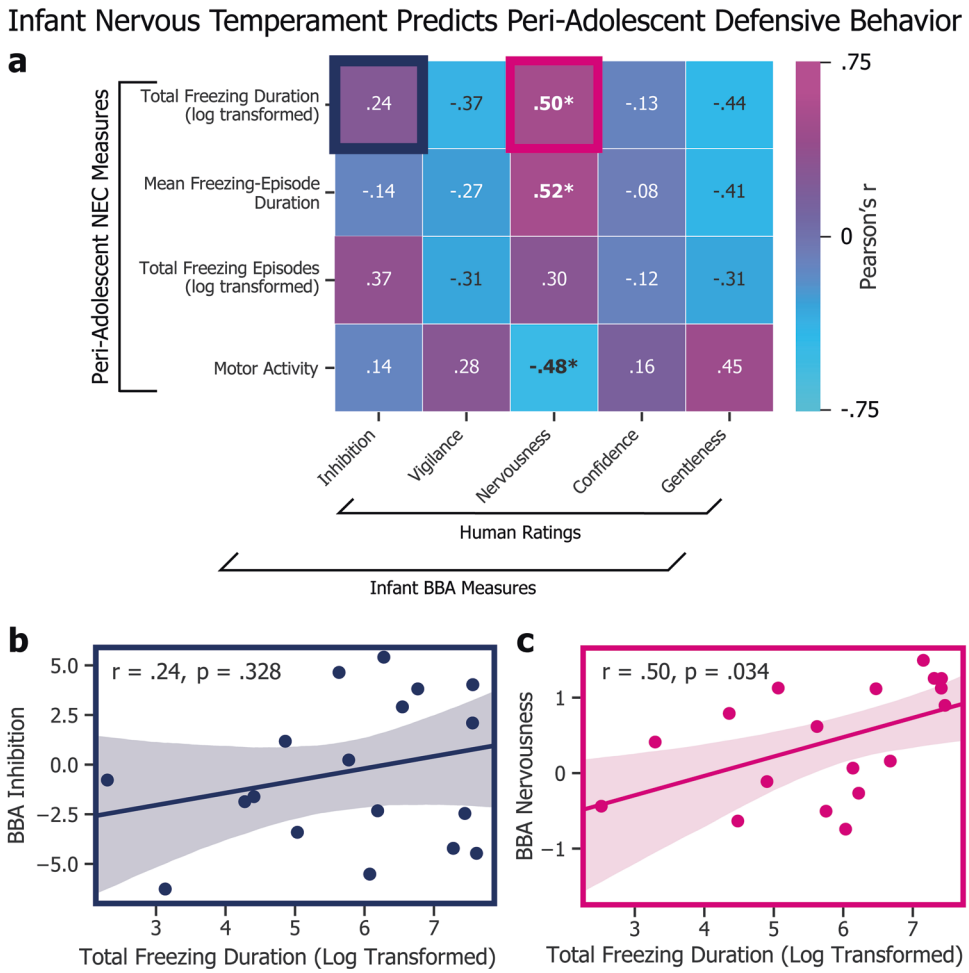

Fig. 4 Infant nervous temperament predicts peri-adolescent defensive behavior. a Heatmap of associations between infant BBA measures and peri-adolescent NEC measures (*$p < 0.05$). **b** We found no association between BBA inhibition and total NEC freezing duration ($r = 0.37$, $p = 0.127$). **c** Experimenter-rated BBA nervousness, however, was a significant predictor of total NEC freezing duration ($r = 0.50$, $p < 0.05$).

(kappa = 0.77, $p < 0.001$), and between UC-Freeze and the raters' consensus (i.e., modal) classifications (kappa = 0.73, $p < 0.001$), confirming the substantial, above-chance agreement [84] between each pair of raters, and between the average rater and UC-Freeze (Fig. 3c). Finally, although the freezing definition has required the animal to maintain a "tense body posture", UC-Freeze does not account for posture. We hypothesized that this may not be problematic, as UC-Freeze was consistent with human ratings of freezing and humans can be inconsistent when coding "tense body posture". To further test this hypothesis, we had five humans classify "tense body posture" across 80 3-second video clips of UC-Freeze freezing bouts that were selected by an expert to contain "tense body posture" (50%) or not (50%). Results demonstrated very little consistency across human raters, with the average between-rater agreement of kappa = 0.215 (average $R^2 = 0.08$), with only 6/80 videos achieving unanimous agreement of all five raters. Taken together, these analyses validate UC-Freeze as a reliable, sensitive, and specific tool for classifying freezing behavior in rhesus, at a standard comparable to that of human raters.

Validation: UC-Freeze detects established brain-behavior relationships

To further validate UC-Freeze, we first examined the relationship between behavior and well-established metabolic correlates. Specifically, we looked for a relationship between subjects' movement about their enclosures during the NEC, automatically coded by UC-Freeze as *motor activity*, and variation in glucose metabolism in subjects' motor cortices using an a priori ROI. These results demonstrated a significant positive association between motor activity and motor cortex metabolism, as expected ($r = 0.55$, $p < 0.05$; Fig. 3a). These results were corroborated by voxelwise analysis showing a significant relationship between motor activity and metabolism in motor cortex regions ($p < 0.05$, uncorrected; Fig. 3b). These proof-of-principle findings confirm that UC-Freeze can recapitulate a well-established brain-behavior relationship.

Exploring UC-Freeze automated measures

To derive behavioral measures for subsequent correlational analyses, we used UC-Freeze to compute total freezing duration, number of freezing episodes, and mean freezing episode during the NEC context for each animal (Fig. 4a). We observed substantial variability between animals: UC-Freeze identified freezing episodes in all 18 subjects, ranging from 3 episodes in our most infrequent freezer, to 90 in our most prolific freezer. UC-Freeze detected 853 episodes (~3 hours 15 minutes) of total freezing across all animals, which accounted for 10.9% of their total behavior during the NEC. Although we hypothesized that animals would eventually habituate to the presence of the human intruder during the 30-minute NEC context, an analysis of the mean total duration our animals spent freezing during the first and second halves of the NEC suggested that the animals did not habituate (independent-samples t test: $t(34) = 0.27$, $p = 0.79$; Fig. 4b).

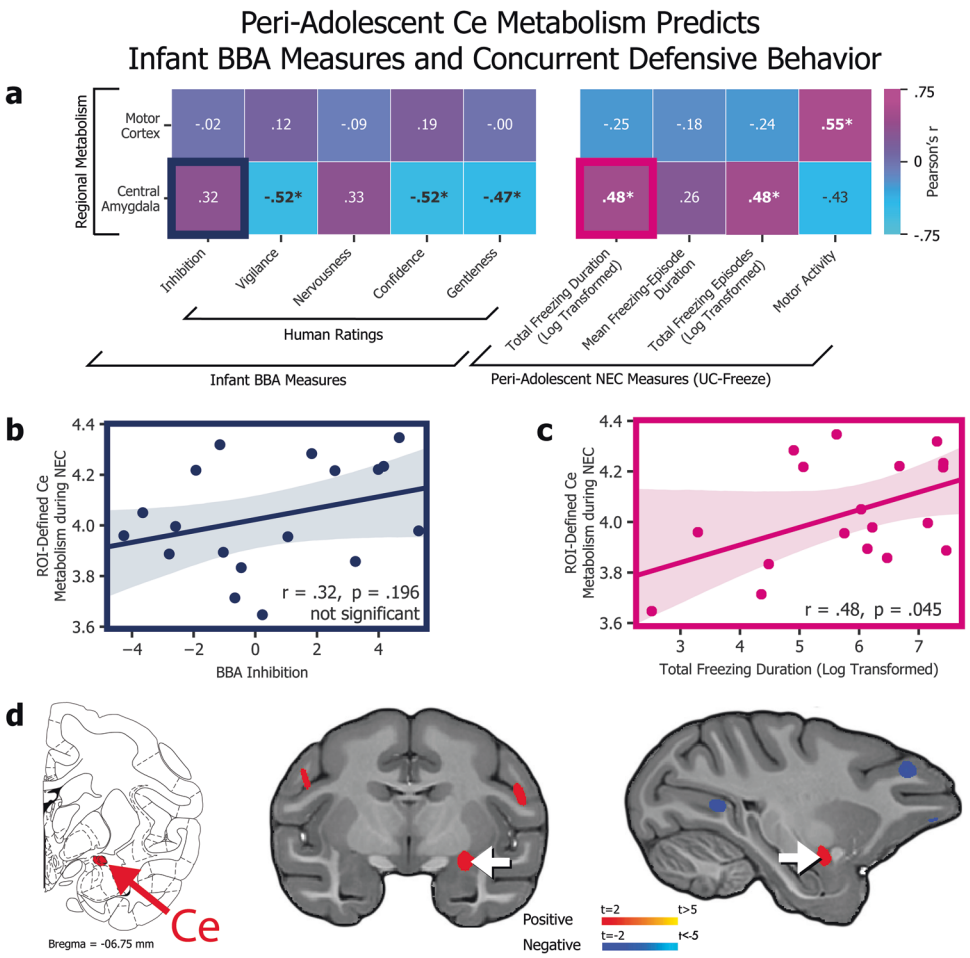

Fig. 5 Peri-adolescent Ce metabolism predicts infant BBA measures and concurrent defensive behavior. a Heatmap of associations between PET-obtained, ROI-defined regional metabolism (*y* axis) and infant BBA measures (*x* axis, left) as well as concurrent NEC behaviors as automatically scored by UC Freeze (*x* axis, right; $*p < 0.05$). **b** The association between Ce ROI metabolism and BBA inhibition was not statistically significant ($r = 0.32$, $p = 0.196$). **c** The association between Ce ROI metabolism and total freezing duration during the NEC, however, was significant ($r = 0.48$, $p < 0.05$). **d** The location of the Ce (shown on the Paxinos et al. atlas, left) corresponds to a voxelwise analysis (middle and right) that revealed a significant correlation between NEC freezing behavior and integrated metabolism in a region of the dorsal amygdala encompassing the Ce ($p < 0.05$, uncorrected). No significant relationships were identified between BST metabolism and infant BBA measures or concurrent defensive behavior (not shown).

UC-Freeze detected large individual differences in animals' split-halves freezing behavior: Some animals froze less during the second half of the NEC, others froze considerably more during the second half, and still others exhibited no notable difference in freezing between the two halves (Fig. 4b).

To capitalize on the ability of UC-Freeze to analyze large datasets, we next examined freezing trends on a per-minute basis by calculating the grand mean average of the animals' probability of freezing during each of the NEC context's 30 1-minute bins. Like our split-halves analysis, individual animal's behavior varied widely, but no overall linear trend in minute-by-minute freezing was observed ($r = -0.25$, $p = 0.18$). These findings are consistent with the view that, on average, our animals' defensive posture did not substantially change during the NEC context.

Infant measures predict peri-adolescent defensive behavior

To test whether infant measures predict variation in defensive behavior in adolescence, we next compared our animals' infant measures to their peri-adolescent freezing and motor activity measured during the NEC (Fig. 5a). We observed no significant relationship between infant-inhibited temperament scores and peri-adolescent total number of freezing episodes ($r = 0.37$, $p = 0.127$), total freezing duration ($r = 0.24$, $p = 0.328$; Fig. 5b), or mean freezing-episode duration ($r = -0.14$, $p = 0.580$) during the NEC context. We further tested the relationships between freezing and BBA experimenter ratings for trait-like *vigilance*, *nervousness*, *confidence*, and *gentleness*. Although the overall measure of inhibited temperament did not significantly predict a greater tendency to freeze during the NEC context, infant nervousness significantly predicted total freezing duration ($r = 0.50$, $p < 0.05$; Fig. 5c) and mean freezing-episode duration ($r = 0.52$, $p < 0.05$). These findings point toward infant nervous temperament as a potential target of future studies aimed at identifying extremely early-life risk factors for the eventual development of anxiety disorders.

Freezing and concurrent FDG

To test whether infant measures predict variation in regional brain metabolism during adolescence, we examined the relationship between EAc metabolism and subjects' defensive behavior as classified by UC-Freeze. Results revealed a significant relationship between animals' integrated Ce metabolism and total time spent freezing, as well as the number of freezing episodes (Fig. 5a). There was no significant relationship between BST metabolism and total freezing duration ($r = 0.32$, $p = 0.20$), nor other UC-Freeze measures of defensive behavior.

ROI-based analyses supported by voxelwise analyses of subjects FDG-PET scans, obtained immediately after exposure to the NEC context (Fig. 1a), revealed a significant relationship between metabolism within an area of the dorsal amygdala encompassing the Ce and subjects' total freezing duration ($r = 0.48$, $p < 0.05$), as well as their total number of freezing episodes ($r = 0.48$, $p < 0.05$). These findings are consistent with previous human [42, 85] and NHP studies [35–37] documenting elevated Ce activation and metabolism, respectively, during threat processing.

DISCUSSION

We developed, validated, and field-tested UC-Freeze, an ML tool for analyzing anxiety-like behavior in rhesus through the semi-automated classification of freezing. Consistent with well-established brain-behavior relationships, UC-Freeze uncovered a significant positive correlation between freezing behavior and increased metabolism in a dorsal amygdala region encompassing the Ce. Because of the increased risk of anxious psychopathology among adolescent females [8], we focused exclusively on a peri-adolescent female cohort. By comparing subjects' infant BI and temperament to their freezing behavior assessed via UC-Freeze, we were able to link infant differences in experimenter-rated BBA nervous temperament to peri-adolescent differences in defensive behavior: Higher nervous temperament ratings by CNPRC staff during infancy predicted more freezing during peri-adolescent exposure to the NEC context.

Large-scale FDG-PET studies of young rhesus at the WNPRC and Harlow Labs have revealed a robust relationship between Ce metabolism and NEC-induced freezing [37, 63]. We replicated this finding at the CNPRC—in animals that have had dramatically different upbringings—by identifying an area of the dorsal amygdala, encompassing the Ce, in which metabolic activity was a significant predictor of NEC-induced freezing. Further, we extended previous work by identifying infant temperament measures that predict peri-adolescent behavior and brain function.

Intriguingly, the Wisconsin researchers have also shown that Ce metabolism is largely attributable to non-heritable influences [37, 63], and can be altered by the overexpression of the plasticity-inducing gene, *NTF3* (neurotrophic factor-3) [86]. In CNPRC animals, we found that Ce regional metabolism was associated with concurrent peri-adolescent freezing, but *not* significantly correlated with infant-inhibited temperament. Other predicted relationships between infant-inhibited temperament and peri-adolescent freezing, as well as BST metabolism, were not statistically significant. While we interpret these findings cautiously in light of our study's limited statistical power, these outcomes hint at the Ce's potential plasticity in response to environmental perturbations. In their large, outdoor, multi-generational social groups, the CNPRC's animals learn from other conspecifics, each with their own idiosyncratic temperament, and experience the formative complexities of social bonds and hierarchies. Raised in this rich social setting, these animals are likely to develop nuanced ways of interacting with others in a variety of contexts, just as humans do. Because of that, the CNPRC's naturalistic conditions provide a unique opportunity to investigate how complex social environments can influence individual differences in BI over time—possibly through Ce plasticity (among other mechanisms).

Together, these observations point to the possibility that Ce metabolism may be particularly relevant to understanding how early-life experience and environment affect the risk of developing anxiety disorders. Future work will be necessary to test this hypothesis and build support for our reported non-significant relationships. Nevertheless, our findings continue to implicate the EAc—and specifically the Ce—as prominently involved in the development of anxious pathology and the expression of defensive behavior. These findings should be considered alongside evidence implicating the EAc in a range of appetitive, consummatory, and addictive behaviors [58, 87–94], as we work toward a more refined understanding that can guide the development of novel interventions [48, 95].

An improved understanding of extremely early-life risk factors for anxiety disorders could lead to premorbid interventions that prevent their onset. In both humans and rhesus, it is challenging to measure behavioral risk factors due to a general lack of motor coordination and the immaturity of threat-response repertoires [56]. Overcoming this challenge could lead to early interventions aimed at blunting organizing effects that contribute to increased risk. Our finding that experimenter-rated nervous temperament in infants predicts peri-adolescent BI in rhesus is consistent with human studies. Such studies have shown that human infants' aversive reactions to negatively valenced stimuli predict BI in childhood [13, 18] and foreshadow the development of anxiety disorders [9, 12, 14, 19]. Our study contributes to an improved understanding of the development of anxiety by directly examining the relationship between infant temperament and peri-adolescent brain function during threat processing. These findings could provide targets for future studies evaluating the longitudinal effects of infant interventions on disordered brain-behavior relationships.

UC-Freeze lowers the bar for other groups to replicate or extend these findings in animals with a diversity of early experiences. More generally, UC-Freeze demonstrates the potential for ML tools to augment existing behavioral neuroscience approaches. A reliance on hand scoring can make behavioral paradigms challenging to scale, since the time required to score each video may be several times greater than the duration of the video itself. Scoring behavior during a 30-minute paradigm administered to hundreds of animals—such as in Fox et al. ($n = 592$) [37]—can impose a significant burden. Nevertheless, the benefit of increased statistical power provided by scale often justifies these efforts. One goal of our study was to provide a proof-of-principle solution to the hand-scoring bottleneck that can arise when behavioral studies are scaled to large cohorts. When video-capture conditions such as lighting and camera position are held constant, UC-Freeze only needs to be manually adjusted in edge cases; for instance, to accurately score an animal that always, or never, freezes. Apart from these edge cases, UC-Freeze operates automatically, trivializing the time commitment required to score behavior, and freeing researchers to engage in other tasks. In addition, UC-Freeze can integrate with modern neuroscientific techniques that enable millisecond-by-millisecond temporal recording and manipulation [96–101]. It is worth noting that UC-Freeze diverges from the classical definition of freezing, by not assessing "tense body posture". Validation analyses confirmed that humans can be inconsistent in applying this definition, and that UC-Freeze was sufficient to identify predicted brain-behavior relationships. That said, it is unclear how this definition will extend to other contexts, beyond NEC, and the importance of classifying posture along with freezing remains an open question for the field. Ultimately, other ML-based approaches that use deep-learning posture estimation and behavioral classification are likely to outperform UC-Freeze, providing *fully* automated and more accurate identification of freezing bouts, as well as enabling a more complete assessment of behavior. Unfortunately, these tools do not yet provide "out of the box" functionality, often requiring significant time investments and technological expertize. Until these tools become validated and widely available, UC-Freeze's semi-automated approach can function as a stopgap to dramatically decrease the amount of effort required to assess freezing.

Although our study was reasonably well-powered by NHP neuroimaging standards, it was unlikely to detect anything less than a large effect as a significant predictor of brain-behavior relationships. Contrary to our predictions, our study did not reveal

a significant correlation between infant-inhibited temperament and peri-adolescent defensive behavior (though the non-significant relationship *was* in the predicted direction; Fig. 5b). However, we refrain from further interpreting this result given the limited statistical power of our study. Previous work has demonstrated infant inhibited temperament to significantly predict lasting inhibited tendencies (i.e., reticence to reach for a food reward) [47], and freezing during NEC has been shown to be moderately stable over time, with the correlation between assessments ~6 months apart being correlated at $\sim r = 0.6$ [35]. A power analysis revealed that, in our $n = 18$ subjects, we had ~80% power to identify a correlation that accounted for ~36% variance (R Studio version 1.0.153's *pwr* package). Nevertheless, the 95% confidence interval for the correlation between infant-inhibited temperament on freezing ranged from −0.26 to 0.64, which is consistent with the possibility that there is a modest relationship between infant-inhibited temperament and adolescent freezing. It will be important to increase statistical power to appropriately assess this relationship. A lack of significant correlation could be important for understanding the instability of BI, differences in the developmental expression of BI, and/or differences between the methodologies for infant and adolescent assessments. We remain interested in the potential relationship between inhibited temperament and peri-adolescent defensive behavior, and are engaged in well-powered studies that explore it thoroughly.

Our findings underscore the importance of examining the developmental timecourse of BI. Peri-adolescent BI reflects both inborn temperament and a multitude of environmental influences that accumulate during maturation. Moving forward, it will be important to scale-up these efforts, investigate sex differences, and integrate these findings with mechanistic studies.

DATA AVAILABILITY

MRI data has been shared with the PRIMatE Data Exchange (PRIME-DE; https://fcon_1000.projects.nitrc.org/indi/indiPRIME.html) Statistical maps have been shared with Neurovault.org (https://identifiers.org/neurovault.collection:16616). All data are available upon request.

CODE AVAILABILITY

UC-Freeze python code is available on github (https://github.com/foxlab-/uc-freeze).

REFERENCES

1. Bandelow B, Michaelis S. Epidemiology of anxiety disorders in the 21st century. Dialogues Clin Neurosci. 2015;17:327–35.
2. Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen H. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. Int J Methods Psychiatr Res. 2012;21:169–84.
3. The US Burden of Disease Collaborators, Mokdad AH, Ballestros K, Echko M, Glenn S, Olsen HE, et al. The state of US health, 1990-2016: burden of diseases, injuries, and risk factors among US states. JAMA. 2018;319:1444.
4. Merikangas KR, Mehta RL, Molnar BE, Walters EE, Swendsen JD, Aguilar-Gaziola S, et al. Comorbidity of substance use disorders with mood and anxiety disorders: results of the internal consortium in psychiatric epidemiology. Addict Behav. 1998;23:893–907.
5. Smith JP, Book SW. Anxiety and substance use disorders: a review. Psychiatr Times. 2008;25:19–23.
6. Smith JP, Randall CL. Anxiety and alcohol use disorders: comorbidity and treatment considerations. Alcohol Res. 2012;34:414–31.
7. Swendsen J, Conway KP, Degenhardt L, Glantz M, Jin R, Merikangas KR, et al. Mental disorders as risk factors for substance use, abuse and dependence: results from the 10-year follow-up of the National Comorbidity Survey. Addiction. 2010;105:1117–28.
8. McLean CP, Asnaani A, Litz BT, Hofmann SG. Gender differences in anxiety disorders: prevalence, course of illness, comorbidity and burden of illness. J Psychiatr Res. 2011;45:1027–35.
9. Chronis-Tuscano A, Degnan KA, Pine DS, Perez-Edgar K, Henderson HA, Diaz Y, et al. Stable early maternal report of behavioral inhibition predicts lifetime social anxiety disorder in adolescence. J Am Acad Child Adolesc Psychiatry. 2009;48:928–35.
10. Clauss JA, Blackford JU. Behavioral inhibition and risk for developing social anxiety disorder: a meta-analytic study. J Am Acad Child Adolesc Psychiatry. 2012;51:1066–75.e1.
11. Clauss JA, Seay AL, VanDerKlok RM, Avery SN, Cao A, Cowan RL, et al. Structural and functional bases of inhibited temperament. Soc Cogn Affect Neurosci. 2014;9:2049–58.
12. Biederman J, Hirshfeld-Becker DR, Rosenbaum JF, Hérot C, Friedman D, Snidman N, et al. Further evidence of association between behavioral inhibition and social anxiety in children. Am J Psychiatry. 2001;158:1673–9.
13. Fox AS, Kalin NH. A translational neuroscience approach to understanding the development of social anxiety disorder and its pathophysiology. Am J Psychiatry. 2014;171:1162–73.
14. Sandstrom A, Uher R, Pavlova B. Prospective association between childhood behavioral inhibition and anxiety: a meta-analysis. Res Child Adolesc Psychopathol. 2020;48:57–66.
15. Kagan J, Reznick JS, Snidman N. The physiology and psychology of behavioral inhibition in children. Child Dev. 1987;58:1459–73.
16. Fox NA, Henderson HA, Marshall PJ, Nichols KE, Ghera MM. Behavioral inhibition: linking biology and behavior within a developmental framework. Annu Rev Psychol. 2005;56:235–62.
17. Fox NA, Snidman N, Haas SA, Degnan KA, Kagan J. The relations between reactivity at 4 months and behavioral inhibition in the second year: replication across three independent samples. Infancy. 2015;20:98–114.
18. Moehler Eva, Kagan Jerome, Oelkers-Ax Rieke, Brunner Romuald, Poustka Luise, Haffner Johann, et al. Infant predictors of behavioural inhibition. Br J Dev Psychol. 2008;26:145–50.
19. Hirshfeld DR, Rosenbaum JF, Biederman J, Bolduc EA, Faraone SV, Snidman N, et al. Stable behavioral inhibition and its association with anxiety disorder. J Am Acad Child Adolesc Psychiatry. 1992;31:103–11.
20. Barbas H. Anatomic basis of cognitive-emotional interactions in the primate prefrontal cortex. Neurosci Biobehav Rev. 1995;19:499–510.
21. Barbas H, Pandya DN. Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. J Comp Neurol. 1989;286:353–75.
22. Barbas H, Zikopoulos B, Timbie C. Sensory pathways and emotional context for action in primate prefrontal cortex. Biol Psychiatry. 2011;69:1133–9.
23. Öngür D, Ferry AT, Price JL. Architectonic subdivision of the human orbital and medial prefrontal cortex. J Comp Neurol. 2003;460:425–49.
24. Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, Mardis ER, et al. Evolutionary and biomedical insights from the rhesus macaque genome. Science. 2007;316:222–34.
25. Preuss TM, Qi H, Kaas JH. Distinctive compartmental organization of human primary visual cortex. Proc Natl Acad Sci. 1999;96:11601–6.
26. Amaral DG. The primate amygdala and the neurobiology of social behavior: implications for understanding social anxiety. Biol Psychiatry. 2002;51:11–7.
27. Liu Z, Li X, Zhang J-T, Cai Y-J, Cheng T-L, Cheng C, et al. Autism-like behaviours and germline transmission in transgenic monkeys overexpressing MeCP2. Nature. 2016;530:98–102.
28. Amaral DG, Schumann CM, Nordahl CW. Neuroanatomy of autism. Trends Neurosci. 2008;31:137–45.
29. Champoux M, Bennett A, Shannon C, Higley JD, Lesch KP, Suomi SJ. Serotonin transporter gene polymorphism, differential early rearing, and behavior in rhesus monkey neonates. Mol Psychiatry. 2002;7:1058–63.
30. Capitanio JP, Miller LA, Schelegle ES, Mendoza SP, Mason WA, Hyde DM. Behavioral inhibition is associated with airway hyperresponsiveness but not atopy in a monkey model of asthma. Psychosom Med. 2011;73:288–94.
31. Capitanio JP. Naturally occurring nonhuman primate models of psychosocial processes. ILAR J. 2017;58:226–34.
32. Chun K, Capitanio JP. Developmental consequences of behavioral inhibition: a model in rhesus monkeys (Macaca mulatta). Dev Sci. 2016;19:1035–48.
33. Golub MS, Hogrefe CE, Widaman KF, Capitanio JP. Iron deficiency anemia and affective response in rhesus monkey infants. Dev Psychobiol. 2009;51:47–59.
34. Kagan J, Reznick JS, Clarke C, Snidman N, Garcia-Coll C. Behavioral inhibition to the unfamiliar. Child Dev. 1984;55:2212–25.
35. Fox AS, Shelton SE, Oakes TR, Davidson RJ, Kalin NH. Trait-like brain activity during adolescence predicts anxious temperament in primates. PLoS ONE. 2008;3:e2570.
36. Fox AS, Oler JA, Shelton SE, Nanda SA, Davidson RJ, Roseboom PH, et al. Central amygdala nucleus (Ce) gene expression linked to increased trait-like Ce metabolism and anxious temperament in young primates. Proc Natl Acad Sci. 2012;109:18108–13.
37. Fox AS, Oler JA, Shackman AJ, Shelton SE, Raveendran M, McKay DR, et al. Intergenerational neural mediators of early-life anxious temperament. Proc Natl Acad Sci. 2015;112:9118–22.

38. Shackman AJ, Fox AS, Oler JA, Shelton SE, Davidson RJ, Kalin NH. Neural mechanisms underlying heterogeneity in the presentation of anxious temperament. Proc Natl Acad Sci. 2013;110:6145–50.
39. Kalin NH, Shelton SE. Defensive behaviors in infant rhesus monkeys: environmental cues and neurochemical regulation. Science. 1989;243:1718–21.
40. Kalin NH, Shelton SE, Rickman M, Davidson RJ. Individual differences in freezing and cortisol in infant and mother rhesus monkeys. Behav Neurosci. 1998;112:251–4.
41. Paulus MP, Stein MB. An insular view of anxiety. Biol Psychiatry. 2006;60:383–7.
42. Etkin A, Wager TD. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. Am J Psychiatry. 2007;164:1476–88.
43. Shin LM, Liberzon I. The neurocircuitry of fear, stress, and anxiety disorders. Neuropsychopharmacology. 2010;35:169–91.
44. Clauss J. Extending the neurocircuitry of behavioural inhibition: a role for the bed nucleus of the stria terminalis in risk for anxiety disorders. Gen Psychiatry. 2019;32:e100137.
45. Blackford JU, Pine DS. Neural substrates of childhood anxiety disorders a review of neuroimaging findings. Child Adolesc Psychiatr Clin N. Am. 2012;21:501–25.
46. Blackford JU, Clauss JA, Avery SN, Cowan RL, Benningfield MM, VanDerKlok RM. Amygdala–cingulate intrinsic connectivity is associated with degree of social inhibition. Biol Psychol. 2014;99:15–25.
47. Fox AS, Harris RA, Rosso LD, Raveendran M, Kamboj S, Kinnally EL, et al. Infant inhibited temperament in primates predicts adult behavior, is heritable, and is associated with anxiety-relevant genetic variation. Mol Psychiatry. 2021;26:6609–18.
48. Fox AS, Shackman AJ. The central extended amygdala in fear and anxiety: closing the gap between mechanistic and neuroimaging research. Neurosci Lett. 2019;693:58–67.
49. Kovner R, Kalin NH. Transcriptional profiling of amygdala neurons implicates PKCδ in primate anxious temperament. Chronic Stress. 2021;5:2470547021989329.
50. Kovner R, Souaiaia T, Fox AS, French DA, Goss CE, Roseboom PH, et al. Transcriptional profiling of primate central nucleus of the amygdala neurons to understand the molecular underpinnings of early-life anxious temperament. Biol Psychiatry. 2020;88:638–48.
51. Kovner R, Fox AS, French DA, Roseboom PH, Oler JA, Fudge JL, et al. Somatostatin gene and protein expression in the non-human primate central extended amygdala. Neuroscience. 2019;400:157–68.
52. Fox AS, Oler JA, Birn RM, Shackman AJ, Alexander AL, Kalin NH. Functional connectivity within the primate extended amygdala is heritable and associated with early-life anxious temperament. J Neurosci. 2018;38:7611–21.
53. Fox AS, Oler JA, Tromp DPM, Fudge JL, Kalin NH. Extending the amygdala in theories of threat processing. Trends Neurosci. 2015;38:319–29.
54. Mobbs D, Hagan CC, Dalgleish T, Silston B, PrÃ©vost C. The ecology of human fear: survival optimization and the nervous system. Front Neurosci 2015;9:55.
55. Mobbs D, Marchant JL, Hassabis D, Seymour B, Tan G, Gray M et al. From threat to fear: the neural organization of defensive fear systems in humans. J Neurosci. 2009;29:12236–43.
56. Kalin NH. The neurobiology of fear. Sci Am. 1993;268:94–101.
57. Blanchard DC, Blanchard RJ Chapter 2.4 Defensive behaviors, fear, and anxiety. In: Handbook of Behavioral Neuroscience. Elsevier, 2008, pp 63–79.
58. Holley D, Fox AS. The central extended amygdala guides survival-relevant tradeoffs: Implications for understanding common psychiatric disorders. Neurosci Biobehav Rev. 2022;142:104879.
59. Shackman AJ, Tromp DPM, Stockbridge MD, Kaplan CM, Tillman RM, Fox AS. Dispositional negativity: an integrative psychological and neurobiological perspective. Psychol Bull. 2016;142:1275–314.
60. Swanson LW, Petrovich GD. What is the amygdala? Trends Neurosci. 1998;21:323–31.
61. Blanchard DC, Griebel G, Pobbe R, Blanchard RJ. Risk assessment as an evolved threat detection and analysis process. Neurosci Biobehav Rev. 2011;35:991–8.
62. Roelofs K. Freeze for action: neurobiological mechanisms in animal and human freezing. Philos Trans R Soc B Biol Sci. 2017;372:20160206.
63. Oler JA, Fox AS, Shelton SE, Rogers J, Dyer TD, Davidson RJ, et al. Amygdalar and hippocampal substrates of anxious temperament differ in their heritability. Nature. 2010;466:864–8.
64. Cohen P, Cohen J, Brook J. An epidemiological study of disorders in late childhood and adolescence - II. Persistence of disorders. J Child Psychol Psychiatry. 1993;34:869–77.
65. Wesselhoeft R, Pedersen CB, Mortensen PB, Mors O, Bilenberg N. Gender–age interaction in incidence rates of childhood emotional disorders. Psychol Med. 2015;45:829–39.
66. Capitanio JP, Mendoza SP, Mason WA, Maninger N. Rearing environment and hypothalamic-pituitary-adrenal regulation in young rhesus monkeys (Macaca mulatta). Dev Psychobiol. 2005;46:318–30.
67. Sclafani V, Del Rosso LA, Seil SK, Calonder LA, Madrid JE, Bone KJ, et al. Early predictors of impaired social functioning in male rhesus macaques (Macaca mulatta). PLoS ONE. 2016;11:e0165401.
68. Rommeck I, Capitanio JP, Strand SC, McCowan B. Early social experience affects behavioral and physiological responsiveness to stressful conditions in infant rhesus macaques (Macaca mulatta). Am J Primatol. 2011;73:692–701.
69. Capitanio JP, Blozis SA, Snarr J, Steward A, McCowan BJ. Do "birds of a feather flock together" or do "opposites attract"? Behavioral responses and temperament predict success in pairings of rhesus monkeys in a laboratory setting. Am J Primatol. 2017;79:e22464.
70. Gottlieb DH, Capitanio JP. Latent variables affecting behavioral response to the human intruder test in infant rhesus macaques (Macaca mulatta). Am J Primatol. 2013;75:314–23.
71. Kalin NH, Sheltona SE. Nonhuman primate models to study anxiety, emotion regulation, and psychopathology. Ann N. Y Acad Sci. 2003;1008:189–200.
72. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of ANTs similarity metric performance in brain image registration. NeuroImage. 2011;54:2033–44.
73. Avants BB, Yushkevich P, Pluta J, Minkoff D, Korczykowski M, Detre J, et al. The optimal template effect in hippocampus studies of diseased populations. NeuroImage. 2010;49:2457–66.
74. Seidlitz J, Sponheim C, Glen D, Ye FQ, Saleem KS, Leopold DA, et al. A population MRI brain template and analysis tools for the macaque. NeuroImage. 2018;170:121–31.
75. Garey LJ. Atlas of the human brain. J Anat. 1997;191:477–8.
76. VanderPlas J, Connolly AJ, Ivezić Ž, Gray A Introduction to astroML: machine learning for astrophysics. In: 2012 Conference on Intelligent Data Understanding. 2012, pp 47–54.
77. Pennington ZT, Dong Z, Feng Y, Vetere LM, Page-Harley L, Shuman T, et al. ezTrack: an open-source video analysis pipeline for the investigation of animal behavior. Sci Rep. 2019;9:19979.
78. Anagnostaras SG, Wood SC, Shuman T, Cai DJ, LeDuc AD, Zurn KR, et al. Automated assessment of pavlovian conditioned freezing and shock reactivity in mice using the video freeze system. Front Behav Neurosci. 2010;4:158.
79. Blázquez-García A, Conde A, Mori U, Lozano JA. A review on outlier/anomaly detection in time series data. ACM Comput Surv. 2022;54:1–33.
80. Seabold S, Perktold J. Statsmodels: econometric and statistical modeling with python. Austin, Texas, 2010, pp 92–96.
81. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O et al. Scikit-learn: machine learning in python. J Mach Learn Python 2011;12:2825−30.
82. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020;17:261–72.
83. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. Permutation inference for the general linear model. NeuroImage. 2014;92:381–97.
84. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med. 2012;22:276–82.
85. Andreatta M, Glotzbach-Schoon E, Mühlberger A, Schulz SM, Wiemer J, Pauli P. Initial and sustained brain responses to contextual conditioned anxiety in humans. Cortex. 2015;63:352–63.
86. Fox AS, Souaiaia T, Oler JA, Kovner R, Kim JM(Hugo), Nguyen J, et al. Dorsal amygdala neurotrophin-3 decreases anxious temperament in primates. Biol Psychiatry. 2019;86:881–9.
87. Warlow SM, Berridge KC. Incentive motivation: 'wanting' roles of central amygdala circuitry. Behav Brain Res. 2021;411:113376.
88. Hardaway JA, Halladay LR, Mazzone CM, Pati D, Bloodgood DW, Kim M, et al. Central amygdala prepronociceptin-expressing neurons mediate palatable food consumption and reward. Neuron. 2019;102:1037–52.e7.
89. Han W, Tellez LA, Rangel MJ, Motta SC, Zhang X, Perez IO, et al. Integrated control of predatory hunting by the central nucleus of the amygdala. Cell. 2017;168:311–24.e18.
90. Cai H, Haubensak W, Anthony TE, Anderson DJ. Central amygdala PKC-δ+ neurons mediate the influence of multiple anorexigenic signals. Nat Neurosci. 2014;17:1240–8.
91. Ponserre M, Fermani F, Gaitanos L, Klein R. Encoding of environmental cues in central amygdala neurons during foraging. J Neurosci. 2022;42:3783–96.
92. Kim J, Zhang X, Muralidhar S, LeBlanc SA, Tonegawa S. Basolateral to central amygdala neural circuits for appetitive behaviors. Neuron. 2017;93:1464–79.e5.
93. Baumgartner HM, Schulkin J, Berridge KC. Activating corticotropin-releasing factor systems in the nucleus accumbens, amygdala, and bed nucleus of stria terminalis: incentive motivation or aversive motivation? Biol Psychiatry. 2021;89:1162–75.
94. Fadok JP, Markovic M, Tovote P, Lüthi A. New perspectives on central amygdala function. Curr Opin Neurobiol. 2018;49:141–7.

95. Campos LJ, Arokiaraj CM, Chuapoco MR, Chen X, Goeden N, Gradinaru V, et al. Advances in AAV technology for delivering genetically encoded cargo to the nonhuman primate nervous system. Curr Res Neurobiol. 2023;4:100086.
96. Chuapoco MR, Flytzanis NC, Goeden N, Christopher Octeau J, Roxas KM, Chan KY et al. Adeno-associated viral vectors for functional intravenous gene transfer throughout the non-human primate brain. Nat Nanotechnol 2023;18:1241–51.
97. Chen X, Ravindra Kumar S, Adams CD, Yang D, Wang T, Wolfe DA, et al. Engineered AAVs for non-invasive gene delivery to rodent and non-human primate nervous systems. Neuron. 2022;110:2242–57.e6.
98. Banko M, Brill E. Scaling to very very large corpora for natural language disambiguation. In: Proceedings of the 39th annual meeting of the association for computational linguistics. association for computational linguistics: Toulouse, France, 2001, pp 26–33.
99. Halevy A, Norvig P, Pereira F. The unreasonable effectiveness of data. IEEE Intell Syst. 2009;24:8–12.
100. Deisseroth K. Optogenetics. Nat Methods. 2011;8:26–9.
101. Resendez SL, Stuber GD. In vivo calcium imaging to illuminate neurocircuit activity dynamics underlying naturalistic behavior. Neuropsychopharmacology. 2015;40:238–9.

ACKNOWLEDGEMENTS

We would like to thank the staff at the California National Primate Research Center (CNPRC) and members of the Fox Lab. This work was supported by NIH grants to JPC (OD010962), ASF (R01MH121735), and the CNPRC (P51OD011107). DH thanks KMM for her support.

AUTHOR CONTRIBUTIONS

Holley: methodology, software, validation, formal analysis, data curation, writing (original draft), visualization. Campos: methodology, validation, formal analysis, investigation, data curation, writing (original draft), visualization, project administration. Drzewiecki: methodology, validation, formal analysis, writing (review and editing). Zhang: formal analysis. Capitanio: conceptualization (BBA program), data curation (BBA), methodology (BBA), resources (BBA data), writing (review and editing). Fox: conceptualization, supervision, validation, formal analysis, methodology, software, visualization, writing (original draft), writing (review and editing), supervision, project administration, funding acquisition.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to A. S. Fox.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.