Investigating the Neurocomputational Architecture of Anxiety with Machine-Learning and Statistical-Modeling Approaches

By

DAN HOLLEY

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Psychology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

_____

Andrew Fox, Chair

_____

Brian Trainor

_____

Erie Boorman

Committee in Charge

2022

Table of Contents

Dedication

For Kara, who knows me better than I know myself, and loves me anyway.

For Roxy, whose companionship saved my life.

For Drew, for relentlessly believing in me, even if I didn't believe in myself.

And for my mother, Brenda, who took me to the Worcester Science Center most weekends.

In remembrance of all those we failed, who left us too soon.

"I spend 10% of my time writing and 90% of my time worrying about why I'm not writing."

~David Foster Wallace

Abstract

Anxiety disorders are among the most common psychiatric diagnoses worldwide and rank prominently among the World Health Organization's leading causes of disability. Existing treatments for anxiety disorders are inconsistently effective and often cause adverse side effects, underscoring the need to develop clinical entry points that lead to new intervention strategies. In the past decade, clinical and basic research efforts have identified uncertain threat anticipation as causal in anxiogenesis and have established sensitivity to uncertain-threat anticipation as a transdiagnostic marker of anxiety disorders. Across a range of temporally "certain-vs-uncertain" threat paradigms in nonhuman animals and human participants alike, uncertain-threat contexts consistently elicit heightened behavioral, physiological, neurobiological, and (in the case of humans) emotional measures of anxiety. However, the research community has no compelling explanation as to *why* individuals become more anxious in uncertain-threat contexts—possibly because the term "uncertainty" is sufficiently imprecise as to allow for many operationalizations. This dissertation builds toward a higher-acuity operationalization of "uncertainty" that enables sophisticated manipulations of its tractable features. *Chapter 1* details our replication and extension of nonhuman-primate research into the neural substrates of threat processing. We used a novel machine-learning approach to uncover a new link between infant temperament and adolescent behavioral and neurobiological response to uncertain threat (N=18). In *Chapter 2*, we argue that the EAc has evolved to function as an arbiter of risk-vs-reward tradeoffs for survival optimization, propose a novel "feature-space" model to explain how distinct pathophysiologies can promote a common clinical phenotype, and discuss the implications of our model in the context of psychiatric disorder. *Chapter 3* recounts our development of a computational model of uncertain-threat anticipation used to decompose "uncertainty" into two tractable features. We held one feature (i.e., discrete threat probability) constant while manipulating the other (i.e., hazard rate) in a statistical threat-learning study (N=42) in which our volunteers made risk-vs-reward tradeoffs during periods of uncertain-threat anticipation. Through this novel approach, we learned that hazard rate causally drives anxious behavioral and emotional responding during uncertain-threat anticipation, irrespective of discrete threat probabilities. Collectively, our work elucidates the neurocomputational architecture of anxiety.

Rhesus Infant Nervous Temperament Predicts Peri-Adolescent Central Amygdala Metabolism & Behavioral Inhibition Measured by a Machine-Learning Approach

Holley D*[1,2], Campos LJ*[1,2], Zhang Y[3], Capitanio JP[1,2], Fox AS[1,2]

Abstract

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.

Author affiliations: [1]University of California, Davis, Department of Psychology; [2]California National Primate Research Center, Davis, California; [3]Columbia University, Department of Statistics; *=contributed equally to this manuscript

Keywords: anxiety disorders, behavioral inhibition, BI, threat processing, amygdala, central extended amygdala, central nucleus, machine learning, statistical modeling

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[11,14–16]. 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 5,000 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 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[10,41–46], 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[45,52,53,58–60]. A range of sensory, evaluative, and contextual inputs converge on the EAc, which projects to downstream effector regions to initiate these defensive responses[13,53,60,61]. 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[62,63]. 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[37,64,65]. 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 (https://github.com/DanHolley/UC-Freeze).

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 & Selection Procedure

Twenty peri-adolescent female rhesus monkeys (*M. mulatta*, *M* [SD]= 2.71 years [.44]) that previously underwent BBA testing during infancy (3- to 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- to 4-month-old inhibited temperament. Because females are at increased risk of developing anxiety and depressive disorders as they transition to adolescence[8,66,67], 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. 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[33,68]. 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[30,69–71].

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 2,000 animals[72]. *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[37,73]. 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 time-course 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 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 3T Siemens Skyra scanner, a dedicated rhesus 8-channel surface coil, with inversion-recovery, fast gradient echo prescription (TI/TR/TE/Flip/FOV/Matrix/Bandwidth:1100ms/2500.0ms/3.65ms/7°/154mm/512×512/240Hz/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[74,75] (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[76]. 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[77] 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[35,37,65]; 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 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[78]. 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).

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 criteria, UC-Freeze automatically classifies that segment as freezing (Fig. 2F). 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.

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]. Results of all relationships tested have been reported in the text and/or in figures 3, 5, and 6. 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 *randomise*[83]. Voxelwise analyses were thresholded at $p<.05$, uncorrected.

Results

Validation: UC-Freeze Detects Established Brain-Behavior Relationships

To 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=.55, p<.05; Fig. 3A). These results were corroborated by voxelwise analysis showing a significant relationship between motor activity and metabolism in motor cortex regions ($p < .05$, uncorrected; Fig. 3B). These proof-of-principle findings confirm that UC-Freeze can recapitulate a well-established brain-behavior relationship.

Validation: Comparing UC-Freeze to Human Raters

To further 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[35,37,65]. 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 interrater reliability (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 (kappa=.77, p<.001), and between UC-Freeze and the raters' consensus (i.e., modal) classifications (kappa=.73, p<.001), confirming the substantial, above-chance agreement[84] between each pair of raters, and between the average rater and UC-Freeze (Fig. 3C). 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.

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 (roughly 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=.79; Fig. 4B). 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=-.25, p=.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=.37, p=.127), total freezing duration (r=.24, p=.328; Fig. 5B), or mean freezing-episode duration (r=-.14, p=.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=.50, p<.05; Fig. 5C) and mean freezing-episode duration (r=.52, p<.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. 6A). There was no significant relationship between BST metabolism and total freezing duration (r=.32, p=.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 ($r=.48$, $p<.05$), as well as their total number of freezing episodes ($r=.48$, $p<.05$). 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, a machine-learning 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[37,65]. 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[37,65], 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.

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., 2015 (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.

Scale can also improve ML accuracy. In an influential study on the effect of increasingly large corpora on the accuracy of competing ML techniques in a confusion-set dis-ambiguation task, Banko and Brill (2001)[95] showed not only that the size of the dataset mattered much more in improving accuracy than the specific approach used, but also that the *worst* approach at a small data volume can emerge as the *best* approach as the size of the data set increases. With this in mind, another goal of our study was to produce large datasets per individual subject in order to maximally leverage the inherent quality of ML approaches to grow increasingly accurate and informative as datasets expand (see also[96]). By decomposing our subjects' 30-fps videos into individual frames, we produced 54,000 similarity scores for each subject. This data-maximalist approach will allow us to develop increasingly granular assessments of subjects' behavior at a temporal resolution that would be impractical and prohibitively time-consuming to achieve by hand. Such granularity will grow more important as NHP researchers continue developing neuroscientific strategies[97,98] that enable millisecond-resolution techniques like optogenetics[99] and fiber photometry[100] in studies of NHP behavior.

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[101]. Contrary to our predictions, our study did not reveal a significant correlation between infant inhibited temperament and peri-adolescent defensive behavior (though it *was* in the predicted direction; Fig. 5B). However, we refrain from further interpreting this effect given the limited statistical power of our study. 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[102]). 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 time-course 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.

Acknowledgements

We would like to thank the staff at the CNPRC. 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. *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.

Conflict of Interest Statement

The authors declare no conflicts of interest.

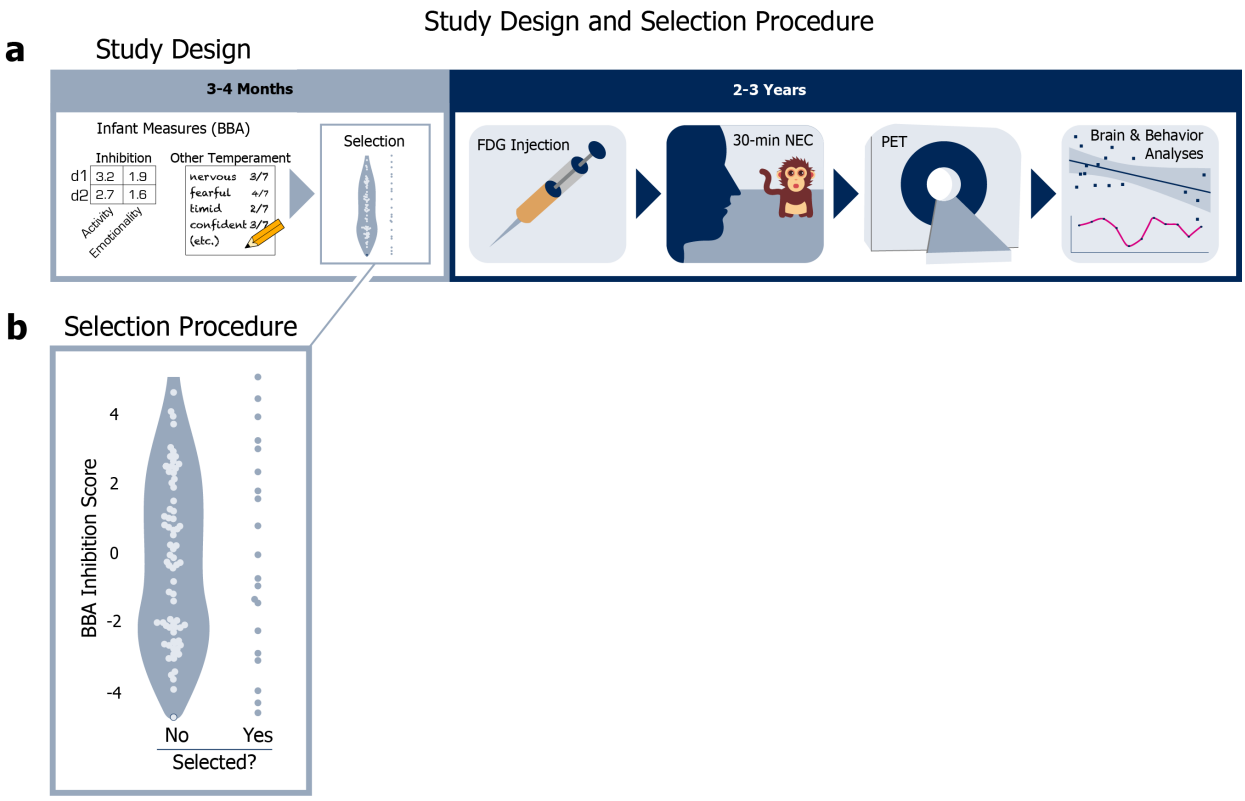

Figure 1. Study design and selection procedure. a) *Study design:* At 3 to 4 months old, all candidate animals were evaluated for a range of infant measures during the BBA. Relevant to our study, the BBA yields objective inhibition scores and other temperament ratings for each animal. At 2 to 3 years 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.

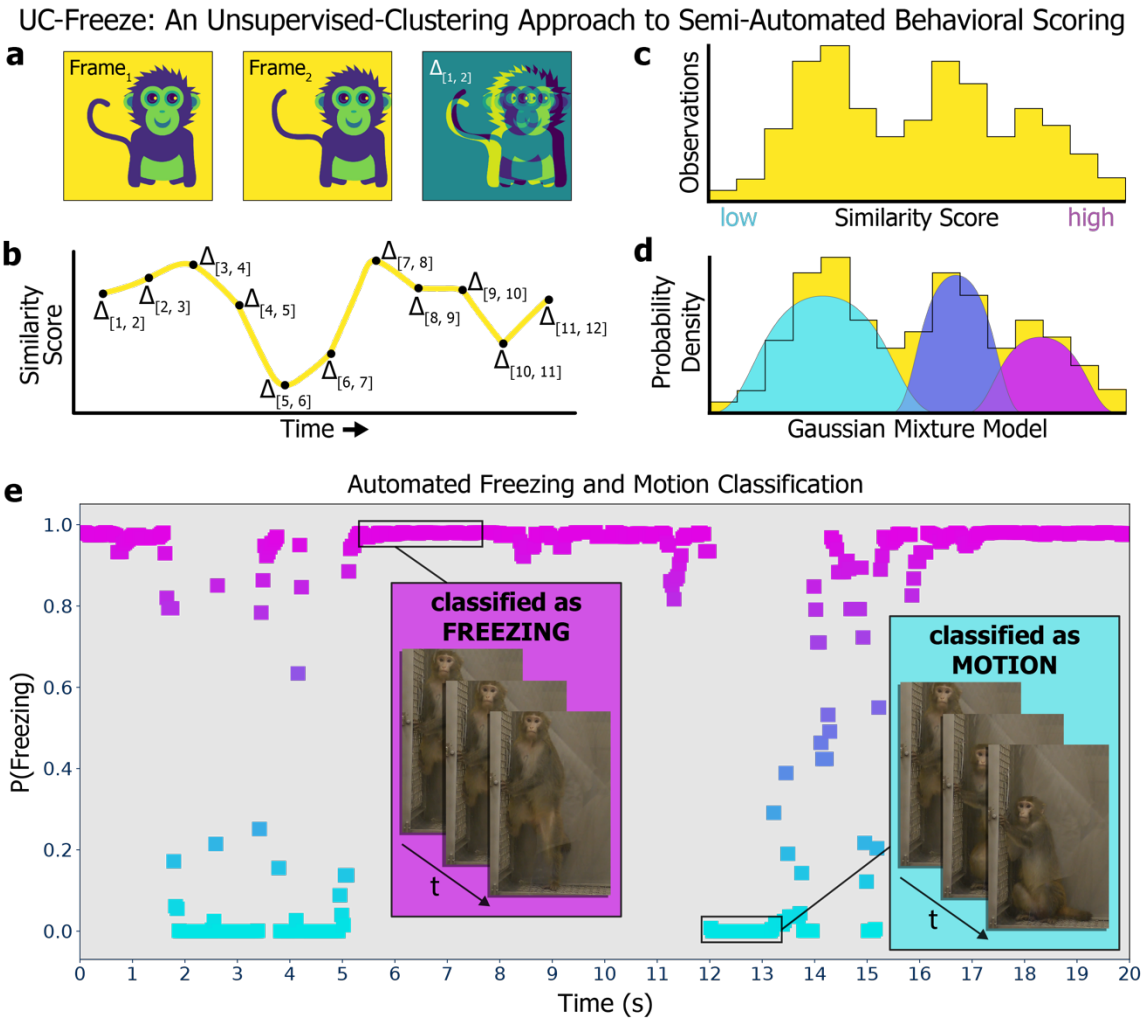

Figure 2. UC-Freeze: An unsupervised-clustering approach to semi-automated behavioral scoring. a) UC-Freeze decomposes 30fps video into individual frames, converts each frame to grayscale, and computes the coefficient of determination value (i.e., r^2, or *similarity score*) for pairs of consecutive 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.

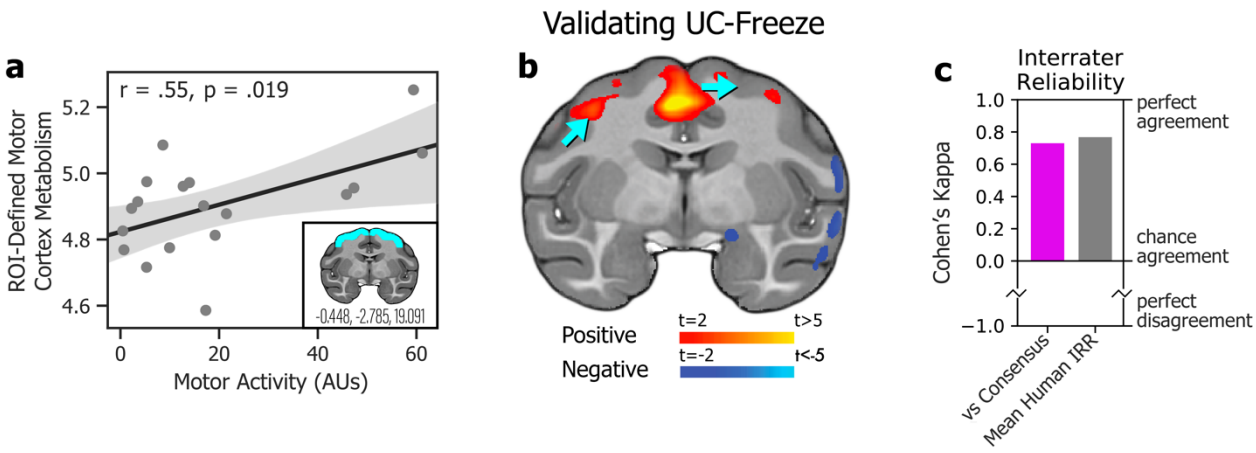

Figure 3. Validating UC-Freeze. a) Subjects' motor activity, as coded by UC-Freeze, significantly predicted integrated motor cortex metabolism in a brain region of interest defined *a priori* by the coordinates x=-0.448, y=-2.785, and z=19.091 (*inset*). **b)** *A posteriori* voxelwise analyses revealed 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=.73, p<.001); and approximated mean human-vs-human IRR (grey; kappa=.77, p<.001), calculated by round-robin comparison. Together, these findings validate UC-Freeze as a reliable tool for scoring freezing in rhesus.

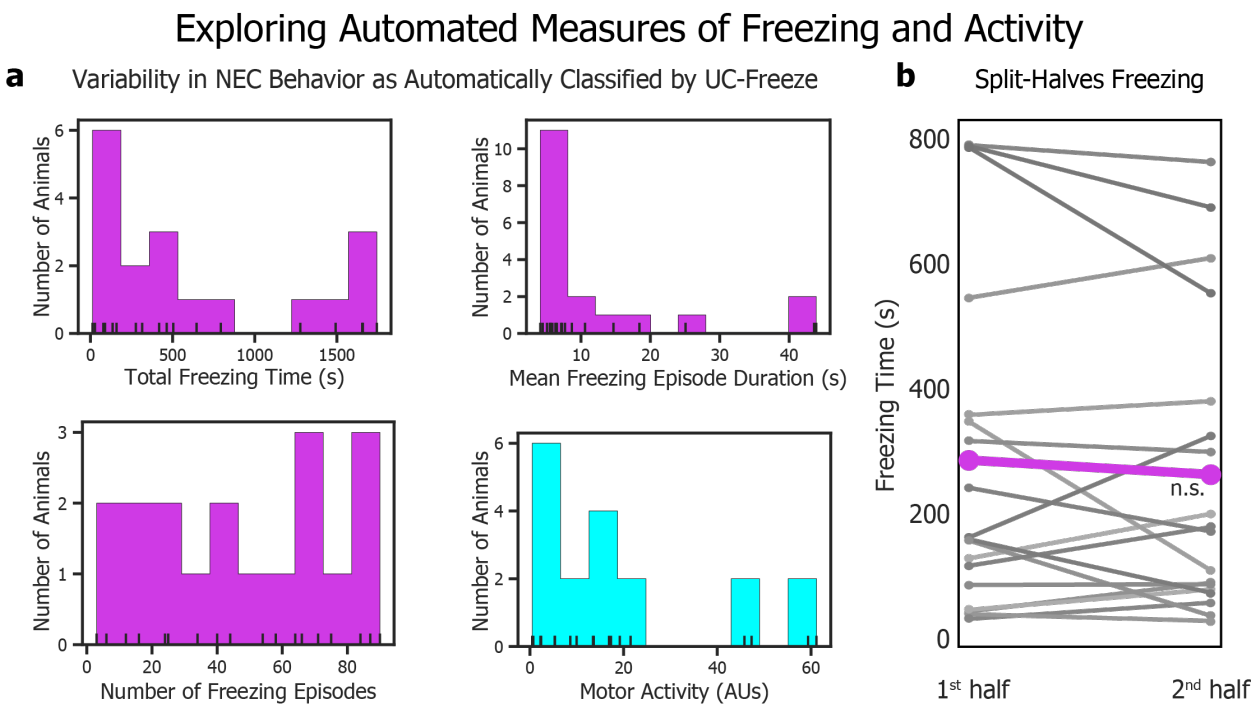

Figure 4. Exploring automated measures of freezing and activity. a) UC-Freeze detected varied total freezing durations, mean freezing-episode durations, total number of freezing episodes, and motor activity in all 18 subjects. **b)** A split-halves analysis revealed no overall trend in subjects' freezing duration in seconds (s) between the first (*M*=299.88s, *SD*=303.25s) and second (*M*=273.12s *SD*=271.65s) halves of the NEC (independent-samples t-test: $t(34)=0.27$, p=.79), indicating that subjects generally did not habituate to the presence of the human intruder during the NEC context.

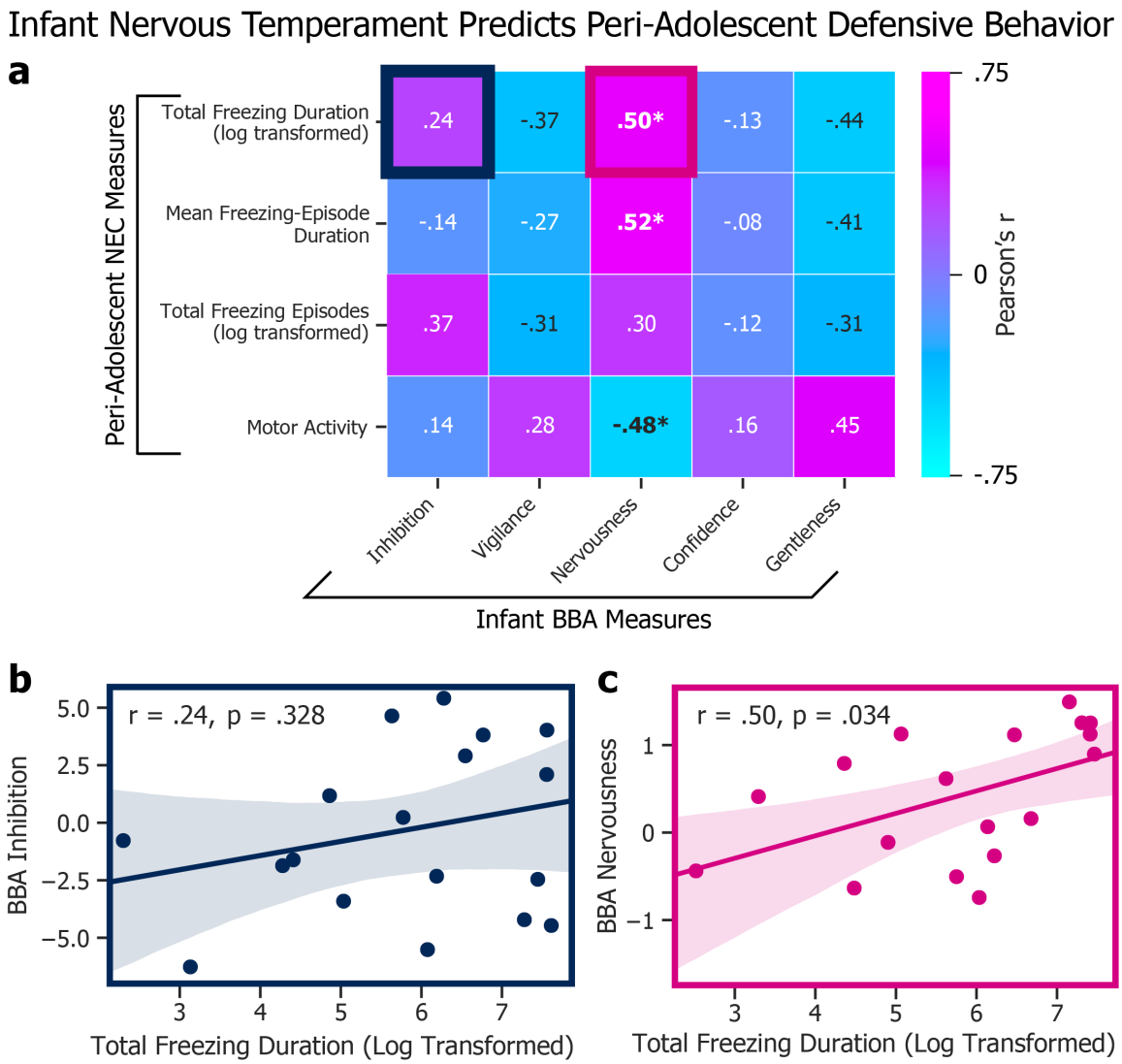

Figure 5. Infant nervous temperament predicts peri-adolescent defensive behavior. a) Heatmap of associations between infant BBA measures and peri-adolescent NEC measures (*=p<.05). **b)** We found no association between BBA inhibition and total NEC freezing duration (r=.37, p=.127). **c)** Experimenter-rated BBA nervousness, however, was a significant predictor of total NEC freezing duration (r=.50, p<.05).

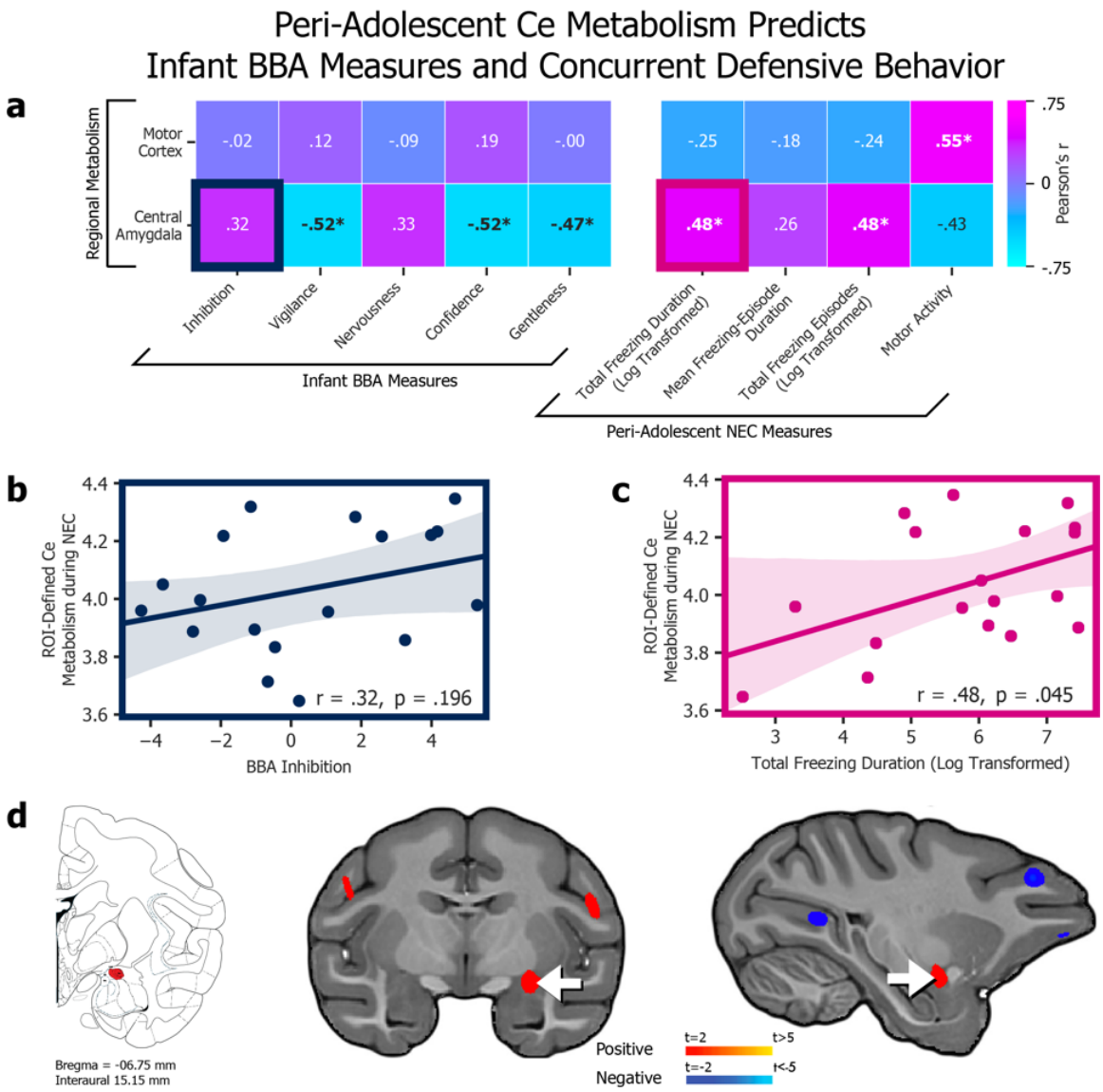

Figure 6. Peri-adolescent Ce metabolism predicts infant BBA measures and concurrent defensive behavior. a) Heatmap of associations between PET-obtained, ROI-defined regional metabolism (*y* axis) and infant BBA measures (*x* axis, *left*) as well as concurrent NEC behaviors as automatically scored by UC Freeze (*x* axis, *right*; *=p < .05). **b)** The association between Ce ROI metabolism and BBA inhibition was not statistically significant (r=.32, p=.196). **c)** The association between Ce ROI metabolism and total freezing duration during the NEC, however, was significant (r=.48, p<.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<.05, uncorrected). No significant relationships were identified between BST metabolism and infant BBA measures or concurrent defensive behavior (*not shown*).

References

1. Bandelow B, Michaelis S. Epidemiology of anxiety disorders in the 21st century. Dialogues Clin Neurosci. 2015 Sep;17(3): 327–35.
2. Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen HU. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. Int J Methods Psychiatr Res. 2012 Sep;21(3):169–84.
3. 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 10;319(14):1444–72.
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 international consortium in psychiatric epidemiology. Addict Behav. 1998 Nov 1;23(6):893–907.
5. Smith JP, Book SW. Anxiety and Substance Use Disorders: A Review. Psychiatr Times. 2008 Oct;25(10): 19–23.
6. 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. Addict Abingdon Engl. 2010 Jun;105(6):1117–28.
7. Smith JP, Randall CL. Anxiety and Alcohol Use Disorders. Alcohol Res Curr Rev. 2012;34(4):414–31
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 Aug;45(8):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 Sep;48(9):928–35.
10. 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 Dec;9(12): 2049–58.
11. Clauss JA, Blackford JU. Behavioral inhibition and risk for developing social anxiety disorder: a meta-analytic study. J Am Acad Child Adolesc Psychiatry. 2012 Oct;51(10):1066-1075.e1.
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 Oct;158(10):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 Nov 1;171(11):1162–73.
14. Sandstrom A, Uher R, Pavlova B. Prospective Association between Childhood Behavioral Inhibition and Anxiety: a Meta-Analysis. J Abnorm Child Psychol. 2020 Jan;48(1): 57–66.
15. Kagan J, Reznick JS, Snidman N. The physiology and psychology of behavioral inhibition in children. Child Dev. 1987 Dec 1;58(6):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 relation between reactivity at 4 months and Behavioral Inhibition in the second year: Replication Across Three Independent Samples. Infancy Off J Int Soc Infant Stud. 2015;20(1):98–114.
18. Moehler E, Kagan J, Oelkers-Ax R, Brunner R, Poustka L, Haffner J, et al. Infant predictors of behavioural inhibition. Br J Dev Psychol. 2008;26(1):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 Jan 1;31(1):103–11.
20. Barbas H. Anatomic basis of cognitive-emotional interactions in the primate prefrontal cortex. Neurosci Biobehav Rev. 1995 Sep 1;19(3):499–510.
21. Barbas H, Pandya DN. Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. J Comp Neurol. 1989;286(3):353–75.
22. Barbas H, Zikopoulos B, Timbie C. Sensory pathways and emotional context for action in primate prefrontal cortex. Biol Psychiatry. 2011 Jun 15;69(12):1133–9.
23. Ongür D, Ferry AT, Price JL. Architectonic subdivision of the human orbital and medial prefrontal cortex. J Comp Neurol. 2003 Jun 2;460(3):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 Apr 13;316(5822):222–34.
25. Preuss TM, Qi H, Kaas JH. Distinctive compartmental organization of human primary visual cortex. Proc Natl Acad Sci U S A. 1999 Sep 28;96(20):11601–6.
26. Amaral DG. The primate amygdala and the neurobiology of social behavior: implications for understanding social anxiety. Biol Psychiatry. 2002 Jan 1;51(1):11–7.
27. Liu Z, Li X, Zhang JT, Cai YJ, Cheng TL, Cheng C, et al. Autism-like behaviours and germline transmission in transgenic monkeys overexpressing MeCP2. Nature [Internet]. 2016 Jan 25 [cited 2016 Feb 2];advance online publication. Available from: http://www.nature.com/nature/journal/vaop/ncurrent/full/nature16533.html#supplementary-information

28. Amaral DG, Schumann CM, Nordahl CW. Neuroanatomy of autism. Trends Neurosci. 2008 Mar;31(3):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(10):1058–63.
30. Capitanio JP, Miller LA, Schelegle ES, Mendoza SP, Mason WA, Hyde DM. Behavioral inhibition is associated with airway hyper-responsiveness but not atopy in a monkey model of asthma. Psychosom Med. 2011 May;73(4):288–94.
31. Capitanio JP. Naturally Occurring Nonhuman Primate Models of Psychosocial Processes. ILAR J. 2017 Dec 1;58(2):226–34.
32. Chun K, Capitanio JP. Developmental consequences of behavioral inhibition: a model in rhesus monkeys (Macaca mulatta). Dev Sci. 2016;19(6):1035–48.
33. Golub MS, Hogrefe CE, Capitanio JP, Widaman KF. Iron Deficiency Anemia and Affective Response in Rhesus Monkey Infants. Dev Psychobiol. 2009 Jan;51(1):47–59.
34. Kagan J, Reznick JS, Clarke C, Snidman N, Garcia-Coll C. Behavioral inhibition to the unfamiliar. Child Dev. 1984;55(6):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 [Internet]. 2008 Jul 2 [cited 2019 Jan 8];3(7). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2430534/
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 U S A. 2012 Oct 30;109(44):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 Jul 21;112(29):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 U S A. 2013 Apr 9;110(15):6145–50.
39. Kalin NH, Shelton SE. Defensive behaviors in infant rhesus monkeys: environmental cues and neurochemical regulation. Science. 1989 Mar 31;243(4899):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 Feb;112(1):251–4.
41. Paulus MP, Stein MB. An insular view of anxiety. Biol Psychiatry. 2006 Aug 15;60(4):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 Oct;164(10):1476–88.
43. Shin LM, Liberzon I. The Neurocircuitry of Fear, Stress, and Anxiety Disorders. Neuropsychopharmacology. 2010 Jan;35(1):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(6): e100137.
45. Blackford JU, Pine DS. Neural substrates of childhood anxiety disorders: a review of neuroimaging findings. Child Adolesc Psychiatr Clin N Am. 2012 Jul;21(3): 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 May;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 May 25;
48. Fox AS, Shackman AJ. The central extended amygdala in fear and anxiety: Closing the gap between mechanistic and neuroimaging research. Neurosci Lett. 2019 Feb 6;693:58–67.
49. Kovner R, Kalin NH. Transcriptional Profiling of Amygdala Neurons Implicates PKCδ in Primate Anxious Temperament. Chronic Stress. 2021 Feb 11;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 May 19;
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 21;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 Off J Soc Neurosci. 2018 Aug 29;38(35): 7611–21.
53. Fox AS, Oler JA, Tromp DPM, Fudge JL, Kalin NH. Extending the amygdala in theories of threat processing. Trends Neurosci. 2015 May;38(5):319–29.
54. Mobbs D, Hagan CC, Dalgleish T, Silston B, Prévost C. The ecology of human fear: survival optimization and the nervous system. Front Neurosci [Internet]. 2015 [cited 2020 Oct 22];9. Available from: https://www.frontiersin.org/articles/10.3389/fnins.2015.00055/full
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 Sep 30;29(39):12236–43.
56. Kalin NH. The Neurobiology of Fear. Sci Am. 2002; 10.

57. Blanchard DC, Blanchard RJ. Chapter 2.4 Defensive behaviors, fear, and anxiety. In: Robert J. Blanchard DCB Guy Griebel and David Nutt, editor. Handbook of Behavioral Neuroscience [Internet]. Elsevier; 2008 [cited 2014 May 30]. P. 63–79. (Handbook of Anxiety and Fear; vol. Volume 17). Available from: http://www.sciencedirect.com/science/article/pii/S1569733907000057
58. Holley D, Fox AS. The Central Extended Amygdala Guides Survival-Relevant Tradeoffs for Action Selection [Internet]. PsyArXiv; 2022 [cited 2022 Jun 17]. Available from: https://psyarxiv.com/z9w3t/
59. Fox AS, Shackman AJ. The central extended amygdala in fear and anxiety: Closing the gap between mechanistic and neuroimaging research. Neurosci Lett. 2017 Nov 30;
60. Shackman AJ, Tromp DPM, Stockbridge MD, Kaplan CM, Tillman RM, Fox AS. Dispositional negativity: An integrative psychological and neurobiological perspective. Psychol Bull. 2016;142(12):1275–314.
61. Swanson LW, Petrovich GD. What is the amygdala? Trends Neurosci. 1998 Aug;21(8):323–31.
62. Blanchard DC, Griebel G, Pobbe R, Blanchard RJ. Risk assessment as an evolved threat detection and analysis process. Neurosci Biobehav Rev. 2011 Mar;35(4):991–8.
63. Roelofs K. Freeze for action: neurobiological mechanisms in animal and human freezing. Philos Trans R Soc B Biol Sci [Internet]. 2017 Apr 19 [cited 2020 Dec 6];372(1718). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5332864/
64. Fox AS. Intergenerational neural mediators of early-life anxious temperament. Neurovault. 2015;
65. 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 Aug;466(7308):864–8.
66. Cohen P, Cohen J, Kasen S, Velez CN, Hartmark C, Johnson J, et al. An epidemiological study of disorders in late childhood and adolescence—I. Age- and gender-specific prevalence. J Child Psychol Psychiatry. 1993 Sep;34(6):851–67.
67. Wesselhoeft R, Pedersen CB, Mortensen PB, Mors O, Bilenberg N. Gender-age interaction in incidence rates of childhood emotional disorders. Psychol Med. 2015 Mar;45(4):829–39.
68. Capitanio JP, Mendoza SP, Mason WA, Maninger N. Rearing environment and hypothalamic-pituitary-adrenal regulation in young rhesus monkeys (Macaca mulatta). Dev Psychobiol. 2005 May;46(4):318–30.
69. 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 [Internet]. 2016 Oct 27 [cited 2019 Jan 22];11(10). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5082922/
70. 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 Jul;73(7):692–701.
71. 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(1):1–11.
72. Gottlieb DH, Capitanio JP. Latent variables affecting behavioral response to the human intruder test in infant rhesus macaques (Macaca mulatta). Am J Primatol. 2013 Apr;75(4):314–23.
73. Kalin NH, Shelton SE. Nonhuman Primate Models to Study Anxiety, Emotion Regulation, and Psychopathology. Ann N Y Acad Sci. 2003;1008(1):189–200.
74. 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 Feb 1;54(3):2033–44.
75. 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 Feb 1;49(3):2457–66.
76. 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 Apr 15;170:121–31.
77. Garey LJ. Atlas of the Human Brain. J Anat. 1997 Oct;191(Pt 3):477–8.
78. Vanderplas J, Connolly A, Ivezić Ž, Gray AG. Introduction to astroML: Machine learning for astrophysics. 2012 Conf Intell Data Underst. 2012;
79. Blázquez-García A, Conde A, Mori U, Lozano JA. A review on outlier/anomaly detection in time series data [Internet]. arXiv; 2020 [cited 2022 Jun 17]. Available from: http://arxiv.org/abs/2002.04236
80. Seabold S, Perktold J. Statsmodels: Econometric and Statistical Modeling with Python. Proc 9th Python Sci Conf. 2010;92–6.
81. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res. 2011 Nov;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 Mar;17(3):261–72.
83. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. Permutation inference for the general linear model. NeuroImage. 2014 May 15;92:381–97.
84. McHugh ML. Interrater reliability: The Kappa statistic. Biochem Medica. 2012 Oct 15;22(3):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 Feb;63:352–63.
86. Fox AS*, Souaiaia T*, Oler JA, Kovner R, Mun J, Nguyen J, et al. Dorsal amygdala neurotrophin-3

decreases anxious temperament in primates. Biol Psychiatry. 2019; 12(15):881-889.
87. Warlow SM, Berridge KC. Incentive motivation: "wanting" roles of central amygdala circuitry. Behav Brain Res. 2021 Aug 6;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 Jun 5;102(5):1037-1052.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 Jan 12;168(1–2):311-324.e18.
90. Cai H, Haubensak W, Anthony T, Anderson DJ. Central amygdala PKC-δ+ neurons mediate the influence of multiple anorexigenic signals. Nat Neurosci. 2014 Sep;17(9):1240–8.
91. Ponserre M, Fermani F, Klein R. Encoding of environmental cues in central amygdala neurons during foraging [Internet]. 2020 Sep [cited 2021 Oct 2] p. 2020.09.28.313056. Available from: https://www.biorxiv.org/content/10.1101/2020.09.28.313056v1
92. Kim J, Zhang X, Muralidhar S, LeBlanc SA, Tonegawa S. Basolateral to Central Amygdala Neural Circuits for Appetitive Behaviors. Neuron. 2017 Mar 22;93(6):1464-1479.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 Jun 15;89(12):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. 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 [Internet]. Toulouse, France: Association for Computational Linguistics; 2001 [cited 2022 Jun 17]. p. 26–33. Available from: https://aclanthology.org/P01-1005
96. Halevy A, Norvig P, Pereira F. The Unreasonable Effectiveness of Data. IEEE Intell Syst [Internet]. 2009 [cited 2022 Jun 17];24(2). Available from: https://ieeexplore.ieee.org/document/4804817
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 [Internet]. 2022 May 27 [cited 2022 Jun 17]; Available from: https://www.sciencedirect.com/science/article/pii/S0896627322004111
98. Chuapoco MR, Flytzanis NC, Goeden N, Octeau JC, Roxas KM, Chan KY, et al. Intravenous gene transfer throughout the brain of infant Old World primates using AAV [Internet]. bioRxiv; 2022 [cited 2022 Jun 17]. p. 2022.01.08.475342. Available from: https://www.biorxiv.org/content/10.1101/2022.01.08.475342v1
99. Deisseroth K. Optogenetics. Nat Methods. 2011 Jan; 8(1): 26–9.
100. Resendez SL, Stuber GD. In vivo Calcium Imaging to Illuminate Neurocircuit Activity Dynamics Underlying Naturalistic Behavior. Neuropsychopharmacology. 2015 Jan;40(1):238–9.
101. Cohen J. A power primer. Psychol Bull. 1992 Jul;112(1): 155–9.
102. Rosario HD, Ford C, Denney B, Weibelzahl S. R package "pwr" [Internet]. 2020 [cited 2022 Jun 17]. Available from: https://github.com/heliosdrm

To appear in *Neuroscience and Behavioral Reviews*, Jan 2023 SI: The Neurobiology of Human Fear and Anxiety

Title:
The central extended amygdala guides survival-relevant tradeoffs: Implications for understanding common psychiatric disorders

Authors:
Dan Holley & Andrew S. Fox

Affiliations:
Department of Psychology and the California National Primate Research Center at the University of California, Davis, Davis, CA, 95616

Abstract
To thrive in challenging environments, individuals must pursue rewards while avoiding threats. Extensive studies in animals and humans have identified the central extended amygdala (EAc)—which includes the central nucleus of the amygdala (Ce) and bed nucleus of the stria terminalis (BST)—as a conserved substrate for defensive behavior. These studies suggest the EAc influences defensive responding and assembles fearful and anxious states. This has led to the proliferation of a view that the EAc is fundamentally a defensive substrate. Yet mechanistic work in animals has implicated the EAc in numerous appetitive and consummatory processes, yielding fresh insights into the microcircuitry of survival- and emotion-relevant response selection. Coupled with the EAc's centrality in a conserved network of brain regions that encode multisensory environmental and interoceptive information, these findings allude to a broader role for the EAc as an arbiter of survival- and emotion-relevant tradeoffs for action selection. Determining how the EAc optimizes these tradeoffs promises to improve our understanding of common psychiatric illnesses like anxiety, depression, alcohol- and substance-use disorders, and anhedonia.

Keywords:
Fear; Anxiety; Defensive Behavior; Reward; Consummatory; Action Selection; Extended Amygdala; Bed Nucleus of the Stria Terminalis; Central Nucleus of the Amygdala; BST; BNST; Ce; CeA

Manuscript:

Optimizing Survival-Relevant Tradeoffs in a Challenging World

The natural world is an unforgiving place, where opportunities to acquire resources, reproduce, and explore must be balanced against ubiquitous threats like predation, starvation, and injury (Blanchard et al., 2011; Blanchard & Blanchard, 2008; Mobbs et al., 2009, 2015). An animal that grazes with reckless abandon might enjoy the short-term benefits of better nutrition, but it's more likely than its vigilant conspecifics to be injured or killed by a predator (Cooper & Blumstein, 2015; Evans & Stempel et al., 2018). Conversely, an animal that tends to forgo its meals and flee at the faintest sign of danger might avoid predators in the short term, but it will eventually suffer malnourishment. Survival-relevant tradeoffs like these pervade the natural world (Fig. 1A, *left*), and the central nervous system evolved to manage them. The human brain also manages *emotion*-relevant tradeoffs, for example the decision of whether to socialize with others or avoid them (Fig. 1A, *right*). While some trepidation in approaching others can be adaptive, an extreme bias toward avoidance can be maladaptive and characteristic of anxiety-related psychopathology (Fox & Kalin, 2014; Shackman et al., 2016). Importantly, the same avoidant behavior could result from any of several biases in the response-selection process (Fig. 1B). How might this selection process be organized in the brain to promote survival in a world of innumerable possibilities? We posit that the brain dynamically integrates sensations, memories, homeostatic signals, preferences, expectations, and other factors into *n*-dimensional feature space where weighted environmental and interoceptive evidence (i.e., $E^*W_k^T$) for survival- and emotion-relevant responses can be represented as values (i.e., $V[R_k]$) and compared (Fig. 1C). The brain must then resolve the feature space to select and trigger adaptive physiology, cognition, and behavior that promote survival and optimize well-being by striking the best balance between risks and rewards.

What constitutes “adaptive” depends on the feature-space inputs, which are unique to individuals at a given moment: Grazing for a few extra seconds as a predator approaches might be adaptive if an animal is especially hungry, if the quality of its food source is high, if its surroundings favor last-second escapes, if nearby conspecifics diffuse the likelihood of being attacked, if the predator is frail or immature, and so on. As the value of each input waxes and wanes, perturbations in the feature space nudge the probability of action selection toward one response or another. Input from a plethora of brain regions shapes the feature space. To avoid “paralysis by analysis” in the management of these high-stakes tradeoffs, specific substrates must integrate across this feature space to rapidly select the most adaptive response. The EAc is well-positioned for this role.

Evidence for an Evolutionarily-Conserved Role for the EAc in Defensive Behavior

In the crucible of natural selection, the primacy of survival has spurred the evolution of defensive adaptations across phylogeny. In mammals, the EAc is a conserved neural substrate that responds to both innate and learned threats. Situated at the center of a distributed network of brain regions that promote fitness in challenging contexts (Fox et al., 2015b; Mobbs et al., 2015), the EAc receives robust direct and indirect input from contextual, sensory, regulatory, and evaluative regions (de Olmos & Heimer, 1999; Swanson & Petrovich, 1998). Two of its major subcomponents—the Ce and the BST—form a functionally coupled circuit (Oler et al., 2012; Oler & Tromp et al., 2017; Tillman et al., 2018; Gorka et al., 2018; Avery et al., 2014) and exhibit similar patterns of gene expression (Bupesh et al., 2011; Fox et al., 2015b; Lein et al., 2007), neurochemistry (Gray & Magnuson, 1992), cellular composition (McDonald, 1982, 1983), and structural connectivity (Fox et al., 2015; Oler & Tromp et al., 2017; Roy et al., 2009). Direct projections from the EAc to effector regions like the periaqueductal gray (PAG) and parvocellular reticular formation (PCRt) trigger selected responses (Han et al., 2017; Tovote et al., 2016). The EAc is therefore well-positioned to synthesize environmental and interoceptive inputs into a

meaningful gestalt, rapidly select optimal defensive responses, and launch those responses into action (Fox et al., 2015b; Mobbs et al., 2015).

Neuroimaging studies of threat-anticipation tasks in humans, and human neuroimaging papers that use the words “fear” and “anxiety,” often report significant activation in the Ce and BST (Fig. 2A, *top*, adapted from Fox & Shackman, 2019, p. 60, Figure 2; Hur et al., 2020; Shackman & Fox, 2016; Somerville et al., 2013; Hur et al., *in press*). In nonhuman primate (NHP) neuroimaging studies, individual differences in rhesus (*Macaca mulatta*) neuroendocrine and behavioral reactivity to potential threats are associated with increased [18F]fludeoxyglucose (FDG) metabolism in the Ce and BST (Fig. 2A, *bottom*; Fox et al., 2008; Fox et al., 2015a; Oler et al., 2010), as well as increased functional connectivity between these regions (Fox et al., 2018; Oler & Tromp et al., 2017). Loss-of-function studies tell a similar story and induce a “threat blind” phenotype in NHPs: rhesus monkeys with gross amygdala lesions, which include the Ce, exhibit more affiliative and sexual behaviors toward intact conspecifics (Emery et al., 2001; Machado & Bachevalier, 2008), are more likely to consume unpalatable foods (Machado et al., 2010), and more readily interact with novel and potentially dangerous objects (Bliss-Moreau et al., 2010, 2011; but see Charbonneau et al., 2021). Similarly, rhesus monkeys with spatially-precise Ce lesions show blunted freezing in response to potential threats and are quicker than intact conspecifics to reach past a snake and retrieve food rewards (Kalin et al., 2004). In rodents, decades of fear-conditioning and threat-related studies have built a foundation for investigating threat processing and have been instrumental to the formulation, testing, and refinement of hypotheses regarding individual responses of the Ce and BST to phasic and sustained threats (Davis et al., 2010; Walker et al., 2009; Walker & Davis, 2008; Marcinkiewcz et al., 2016; Tovote et al., 2016; Perusini & Fanselow, 2015). Despite long-standing evidence of its role in non-defensive processes (e.g., Aggleton, 2000; Baxter & Murray, 2002; Whalen & Phelps, 2009), the

sheer volume of studies implicating the EAc in threat processing could lead one to conclude that it is fundamentally a defensive substrate.

Rodent Studies Uncover the EAc's Diverse Roles in Survival-Related Response Selection

In the past decade, methodological advances have endowed researchers with tools that enable cell type-specific targeting, millisecond-resolution observation, and bidirectional control of neural populations (Chen et al., 2013; Deisseroth, 2011; Resendez & Stuber, 2015; Roth, 2016; Fox & Shackman, 2019). Coupled with well-validated fear-conditioning and threat-related assays, these methods are elucidating the mechanisms that subserve threat processing and have uncovered intermingled populations of EAc neurons that function as substrates for the selection of defensive responses. For example, cell type-specific manipulations within the mouse Ce have identified a competitive inhibitory microcircuit—comprised of intermingled corticotropin releasing hormone-positive (CRH+) and somatostatin-positive (SST+) neurons—that rapidly selects between fleeing and freezing (Fadok et al., 2017; Fig. 3A). Similarly, distinct cell types have been implicated in competitive responses to learned vs. unlearned threats (Isosaka et al., 2015). Other studies have characterized a lateral Ce (CeL) microcircuit that gates conditioned freezing through projections to the medial Ce (CeM; Botta et al., 2015; Ciocchi et al., 2010; Haubensak et al., 2010). This work shows that "CeL_{off}" neurons—which express the anxiety-associated genetic marker protein kinase C-delta (PKCδ+)—form a reciprocal inhibitory microcircuit with intermingled "CeL_{on}"/PKCδ-negative (PKCδ-) neurons. Threat conditioning increases the basal firing rate of the "CeL_{off}"/PKCδ+ population, leading to stronger local inhibition of CeL-CeM projections and increased threat generalization—a transdiagnostic feature of anxiety disorders (Lissek et al., 2010, 2014; Morey et al., 2020). These findings dovetail with work NHPs demonstrating that levels of the transcript encoding for PKCδ in the CeL is associated with individual differences in threat responding (Kovner et al., 2020).

To understand *how* EAc alterations promote pathological anxiety, we need to carefully consider its broader role in arbitrating survival-relevant tradeoffs (i.e., its role in selecting adaptive behavior as a function of $V[R_i]$, or $f(V[R_1], V[R_2]$; see Fig. 2B, C). For instance, researchers have shown that chemogenetic inhibition of CeL PKCδ+ neurons—the same cells that appear to play a mechanistic role in threat generalization—induces risky feeding behavior in mice, as measured by the consumption of bitter tastants that control animals tend to reject (Cai et al., 2014; Ponserre et al., 2020). Other studies have shown that gustatory cortical projections to the Ce encode—and, when manipulated, can even *reverse*—the hedonic value of bitter tastants (Wang et al., 2018). These studies of consummatory behavior are especially interesting in the context of survival-relevant tradeoffs, since aversion to bitterness is an evolved safeguard against the consumption of toxic substances (Bachmanov et al., 2014). Intriguingly, they also hint toward the versatility of the EAc's microcircuitry; that is, the ability of some populations of neurons to bidirectionally control divergent survival behaviors (e.g., eating, threat reactivity) depending on the current context (and experimental probe/assay). Context-dependent repurposing of microcircuits would be an efficient solution to the demands of flexibly responding to the innumerable feature-space perturbations that increase or decrease the adaptiveness of specific emotion-relevant responses. For example, while it may be *generally* maladaptive to graze while predators are nearby, *specific* constraints—such as life-threatening malnutrition—may reshape the feature space so radically that grazing becomes the optimal response. In this case, perhaps "CeL_{on}"/PKCδ- neurons suppress "CeL_{off}"/PKCδ+ neurons to reduce threat responding and promote risky feeding, triggering a "Hail Mary" response as an alternative to imminent death. It is also possible, however, that this appearance of multifunctionality could arise as a product of within-cell-type heterogeneity, and that intermingled groups of ostensibly specific EAc neurons might be further functionally dissociable (e.g., see Zeng, 2022).

An increasing number of studies remind us that the EAc is not a solely defensive substrate. For example, researchers investigating the neural substrates of predatory hunting have begun to dissect the Ce's involvement in prey pursuit and capture. By stimulating the axon terminals of intermingled populations of Ce neurons in mice, Han and colleagues (2017) identified parallel pathways that control appetitive locomotion and consummatory behaviors: a Ce-PAG pathway motivates prey pursuit, while a Ce-PCRt motivates prey consumption. Even in sated animals, activating the Ce-PAG pathway triggers immediate predatory hunting of live or artificial prey, whereas activating the Ce-PCRt pathway triggers immediate biting attacks against these targets, as well as fictive feeding in the absence of prey. Fascinatingly, activating the Ce-PCRt pathway *does not* trigger attacks against other mice, indicating that it is not an indiscriminate "rage circuit," but rather a context-specific circuit for food consumption. In the Kash Lab, researchers investigating the molecular substrates of binge eating discovered a population of prepronociceptin (*Pnoc*)-expressing neurons in the Ce that project food-palatability information to the ventral BST, parabrachial nucleus (PBN), and nucleus of the solitary tract (Hardaway et al., 2019). Activation of Ce *Pnoc* neurons was sufficient to motivate real-time place preference—a widely used index of reward value. Notably, however, the consequences of manipulating these neurons were specific to reward: inhibiting these cells failed to induce anxious behavior in the open field test, elevated plus maze, or other anxiogenic assays. Other work demonstrates that even the Ce's putatively "escape-related" CRH+ neurons can motivate reward seeking in specific contexts. For example, mice will optogenetically self-stimulate "escape-related" CRH+ Ce cells, suggesting an increase in appetitive motivation or hedonic pleasure (Kim et al., 2017). Self-stimulation of these cells has also been shown to increase the amount of effort that rats will expend to obtain sucrose rewards, implying a role in incentive motivation (Baumgartner et al., 2021). Other studies have implicated the EAc in a spectrum of roles ranging from mating behaviors (Wei et al., 2021) and social interaction (Flanigan & Kash, 2020), to binge drinking (Rinker et al., 2017) and nociception (Yu et al., 2021).

The BST, like the Ce, is enriched with distinct neuron populations that mediate several physiological and behavioral features of defensive responding (Kim et al., 2013), making it a priority target for dissecting the mechanisms of anxiety disorders. Moreover, an expanding literature on sex dimorphism in this region hints at its relevance to the profound sex differences in the prevalence of anxiety disorders, which are about twice as common in women compared to men (Lebow & Chen, 2016; Bandelow & Michaelis, 2015). Although less is known about the BST's role in reward and another non-defensive processes, it boasts deep molecular heterogeneity, and its neurons express a range of neuropeptide markers that enable fine-grained modulation of physiological and behavioral survival-related tradeoffs (Gungor & Paré, 2016; Giardino et al., 2018). In mice, for instance, parallel circuits comprised of genetically distinct, lateral hypothalamus (LH)-projecting BST neurons are differentially involved in promoting defensive and appetitive behaviors: one circuit, comprised of CRH+ neurons, promotes avoidance, whereas the other, comprised of cholecystokinin-positive (CCK+) neurons, promotes feeding and mate approach (Giardino et al., 2018). Intriguingly, the latter population may play a key role in addiction (Giardino & Pomrenze, 2021) and appears to interact with estradiol-2 in the presence of cocaine and opioids to reinforce drug-seeking behavior (Ma & Giardino, 2022). This not only highlights the involvement of the BST in non-defensive responding, but also underscores the importance of studying sex as a biological variable in neuroscientific research.

These findings motivate our view that distinct alterations across or within several EAc circuits could give rise to nearly indistinguishable clinical phenotypes, for instance by increasing avoidance (Giardino et al., 2018), dampening incentive motivation (Mahler & Berridge, 2012; Warlow & Berridge, 2021; Baumgartner et al., 2021), shaping hedonic values (Wang et al., 2018), moderating reward-reinforcement signaling (Hardaway et al., 2019), or some combination thereof. Taken together, recent studies of predation and reward demonstrate that the EAc plays a critical

role in both aversive and appetitive survival-related functions—and that the functional "identity" of specific neuron populations is highly context dependent. On balance, these observations render views of the EAc's specificity to threat processing untenable and require us to fundamentally reconsider what the EAc is doing in threatening contexts.

Biological Degeneracy Ensures Partial Redundancy for EAc-Mediated Processes

The EAc does not have a monopoly on selecting between survival-relevant tradeoffs. For instance, in a laboratory paradigm used to induce panic via the inhalation of carbon dioxide (CO_2)-enriched air, even patients with focal bilateral amygdala damage can mount adaptive panic responses (Khalsa et al., 2016). And in freely-behaving mice, a feed-forward excitatory circuit projecting from the dorsomedial superior colliculus (dMSC) to the PAG encodes threat levels and initiates rapid escapes in response to threat stimuli that are parametrically modulated for saliency (Evans & Stempel et al., 2018). Redundancies and "emergency brakes" are to be expected, since evolution favors biological *degeneracy*—that is, "the ability of elements that are structurally different to perform the same function or yield the same output" (Edelman & Gally, 2001, p. 13,763)—over single points of failure. This is consistent with survival as a core determinant of brain evolution across phylogeny. Still, the EAc is uniquely poised to function as an arbiter for survival-relevant tradeoffs. It integrates a wealth of information from myriad regions necessary to encode a survival-relevant feature space (i.e., by computing $f(V[R_1], V[R_2], \ldots V[R_k])$, forms numerous microcircuits capable of rapidly selecting between competing physiological and behavioral responses, and projects to regions that can trigger those responses. Importantly, it is precisely these physiological and behavioral tradeoffs that are shared between survival- and emotion-relevant responses. Therefore, we expect the function of the EAc in survival to be particularly relevant for understanding pathological anxiety and other psychiatric illnesses characterized by prominent alterations in emotion and motivation (e.g., depression, alcohol- and

substance-use disorders, anhedonia). While the EAc is not *required* to mount innate, largely reflexive responses like those we've described here, it seems to be involved in processing both learned (Li, 2019; Fadok et al., 2017; Sanford et al., 2017; Yu, 2017) and unlearned (Isosaka et al., 2015) threats. The Ce exhibits activity-dependent synaptic plasticity (Samson & Paré, 2005), and our work in nonhuman primates suggests that it represents the contributions of learning and experience to the risk of developing anxiety disorders (Holley & Campos et al., 2022; Fox et al., 2015a).

Characterizing Response-Selection Mechanisms in the EAc

We hypothesize that the EAc encodes an *n*-dimensional feature space, where multiple inputs from across the brain converge to form an integrated view that enables adaptive responding to both threats and opportunities. The EAc is hypothesized to play a critical role in normative fear and anxiety (Davis et al., 2010; Fox et al., 2015b; Fox & Shackman, 2019), as well as anxiety-related psychopathology (Avery et al., 2016; Clauss, 2019; Morey et al., 2020; Shackman & Fox, 2021). Although scores of studies document the relationship between alterations in the EAc and differences in threat processing, an expanding mechanistic literature reminds us that the EAc is not threat-specific, and that it guides survival-relevant response selection more broadly. Importantly, lesion studies that find preferential deficits in threat responding do not imply that this region is uninvolved in triggering other responses. The historical tendency to focus on threat processing could reflect experimental biases, or some underlying threat-bias in the EAc's response selection mechanisms. To better understand response selection mechanisms within the EAc must be thoroughly characterized. It is possible, for instance, for dissimilar mechanisms to have the same net effect, thereby promoting a somewhat uniform anxious phenotype via distinct EAc alterations. In fact, we *expect* this to be the case, and to contribute to the challenges in the pharmacological treatment of anxiety disorders (Garakani et al., 2020; Koen & Stein, 2011). For example, a competitive microcircuit that selects between two mutually exclusive behaviors, such

as the Ce CRH+/SST+ microcircuit that selects between fleeing and freezing (Fadok, 2017), could feature any of several alterations that would dispositionally bias an individual toward one behavior over another. A maladaptive bias toward defensive freezing, which is thought to underlie temperamental behavioral inhibition and the risk to develop anxiety-related psychopathology (Fox & Kalin, 2014), could be driven by 1) a preponderance of SST+ neurons (Fig. 3C, *top*), 2) disproportionately strong inhibitory SST+ projections onto CRH+ neurons (Fig. 3C, *middle*), or 3) the tendency of a third population of neurons to inhibit CRH+ neurons (Fig. 3C, *bottom*). Similar outcomes could arise via alterations in the aforementioned "CeL_{on}"/PKCδ- and "CeL_{off}"/PKCδ+ microcircuit. These illustrative mechanisms might respond differently—or not at all—to a common intervention, underscoring a major barrier to the development of one-size-fits-all treatments for psychopathology.

The implications of this within-region cell-type heterogeneity present a challenge for the neuroimaging community. A voxel, the smallest unit of spatial resolution in functional magnetic resonance imaging (fMRI), may represent the activation of hundreds of thousands of neurons. Because of this, blood oxygen level-dependent signal (BOLD) responses collected from intermingled neuron populations that form competitive microcircuits for response selection might look identical in the scanner even when subjects exhibit opposite responses to a given stimulus. But because this heterogeneity is unlikely to be uniform across voxels, it can also lend to the development of hypotheses that move beyond univariate relationships between regional signals and fear/anxiety measures. For example, based on findings from mice, we might hypothesize that the voxels of the basal and lateral regions of the amygdala each contain some mixture of reward- and threat-sensitive cells. With this hypothesis in mind, we might not expect to see differences in activation across these amygdala voxels in a straightforward test of reward vs threat. However, we might expect multivoxel pattern analysis (MVPA; Norman et al., 2006) to reveal *dissociable patterns* of activation that are characteristic of reward or threat processing, because each voxel

has a different mixture of cell types. By parametrically modulating reward or threat information, we may be able to detect changes in patterns—not in any one voxel, but across voxels. Such research could extend MVPA's many contributions (e.g., Chang et al., 2015; Frick et al., 2014; Liu et al., 2015; Norman et al., 2006; Woo et al., 2017) by evaluating hypotheses that posit a conserved organization of reward- and threat-sensitive cells across species. This approach could also be coupled with pharmacological methods: By combining perturbations of the EAc's feature space (i.e., by modulating reward or threat evidence) with drugs believed to target a subset of cell types, we should be able to test hypotheses derived from rodent literature concerning the relationship between specific cell types and the function of the EAc. Such drugs may be useful for these studies, even if they have side-effects or lack clinical efficacy, and include those that target specific serotonin (Sharp & Barnes, 2020), oxytocin (Quintana et al., 2021), and CRH receptors (Zorrilla & Koob, 2010), among others (e.g., neuropeptide Y, cannabinoids, vasopressin, substance P, etc.). These examples illustrate the types of approaches we expect to enable key advances in precision psychiatric diagnostics and treatment in the years to come. Creative study design centered on cross-modality approaches like these are needed to help blunt the enormous public-health burden of anxiety disorders (Beddington et al., 2008) and improve the effectiveness and availability of treatment to the untold millions who suffer their effects (Bandelow & Michaelis, 2015; U.S. Burden of Disease Collaborators, 2018).

Toward an Improved Understanding of Common Psychiatric Disorders

In a seminal review, Rangel, Camerer, and Montague (2008) laid out the processes necessary for action selection to take place in the brain, noting that each process is experimentally tractable: (1) representation of a problem, (2) assignment of values to possible options, (3) selection and implementation of a winning option, (4) evaluation of the outcome, and (5) feedback to enable learning and refinement. These functions are not unique to a single brain region. Here, we have argued that the EAc integrates salient environmental and interoceptive features in an *n*-

dimensional space (akin to steps 1 and 2, above), and guides adaptive responses to challenges and opportunities alike by resolving that feature space to select winning strategies (akin to step 3, above). Although it lies beyond the scope of our mini-review, recent work suggests that the EAc is well-suited to perform steps (4) and (5), for example, via inputs from the PBN (Palmiter, 2018) and ventral tegmental area (Li, 2019), respectively. Moreover, the EAc is differentiated from other brain systems involved in action selection by its direct projections to the effector regions that can induce specific aspects of an emotional response, including species-typical physiological changes (e.g., cardiorespiratory and skin-conductance responses) and behaviors (escape, pursuit, freezing, etc.).

The EAc is uniquely poised to perform survival- and emotion-relevant action selection, and so it is a priority target for understanding psychiatric disorders characterized by prominent alterations in emotion or motivation. However, our expanding knowledge of its neuron populations and their multifunctional, context-dependent involvement in defensive and non-defensive processes should give us pause as we carefully rethink what these findings mean for the study of mental illness. This may require a conceptual reframing of how EAc alterations contribute to pathophysiology. Here, we have proposed approaching survival- and emotion-relevant tradeoffs (Fig. 1) as the outputs of an *n*-dimensional feature space that is encoded and resolved by the EAc (Fig. 2B). This computational approach to understanding survival- and emotion-relevant response selection in the EAc is intended to complement and integrate with other theories of how the brain implements these tradeoffs (e.g., Mobbs et al., 2015; Perusini & Fanselow, 2015; LeDoux & Pine, 2015; etc.).

A major implication of this conceptual reframing is that the same disordered phenotype could arise from alterations in distinct cellular/molecular substrates, which presents challenges for the development of effective treatments. Because the mechanism(s) that the EAc uses to compare

feature vectors for survival- and emotion-relevant decisions are presently unknown (Fig. 2B, *middle*), investigations that parametrically modulate feature-space inputs will be especially valuable in elucidating the mechanisms that select between the EAc's response sets—and, when imbalanced, contribute to maladaptive responding (Fig. 3B). Importantly, there are likely to be *many-to-one* and *one-to-many* relationships between biological dysregulation and psychopathology. As outlined in Figure 3, multiple, distinct biological mechanisms within the EAc could lead to the same output. (Similarly—although not discussed in detail here—a common CeL alteration might differentially bias physiological, cognitive, and behavioral outputs via distinct alterations in downstream mechanisms, for example in regions innervated by CeM outputs). A current challenge for human research is to develop and test hypotheses derived from animal studies to understand the role of the EAc in human psychopathology. Here, NHP studies of the EAc's non-defensive functions (e.g., Parkinson, 2001) will be instrumental in understanding how mechanistic discoveries in rodents relate to the disordered emotion-relevant responses common to clinical populations. A focused effort toward characterizing how the EAc's feature-space inputs are encoded and what the comparison process for response selection entails will enable targeted manipulations of specific cells, genes, and molecules and uncover clinical entry points in the development of new interventions for a range of common psychiatric disorders.

Acknowledgements
This work was supported by NIH Grants R01MH121735, R21MH129851 to ASF, R01MH131264 (Co-I: ASF; PI: A. Shackman) and the California National Primate Research Center (P51OD011107). DH thanks KMM for her support and insights.

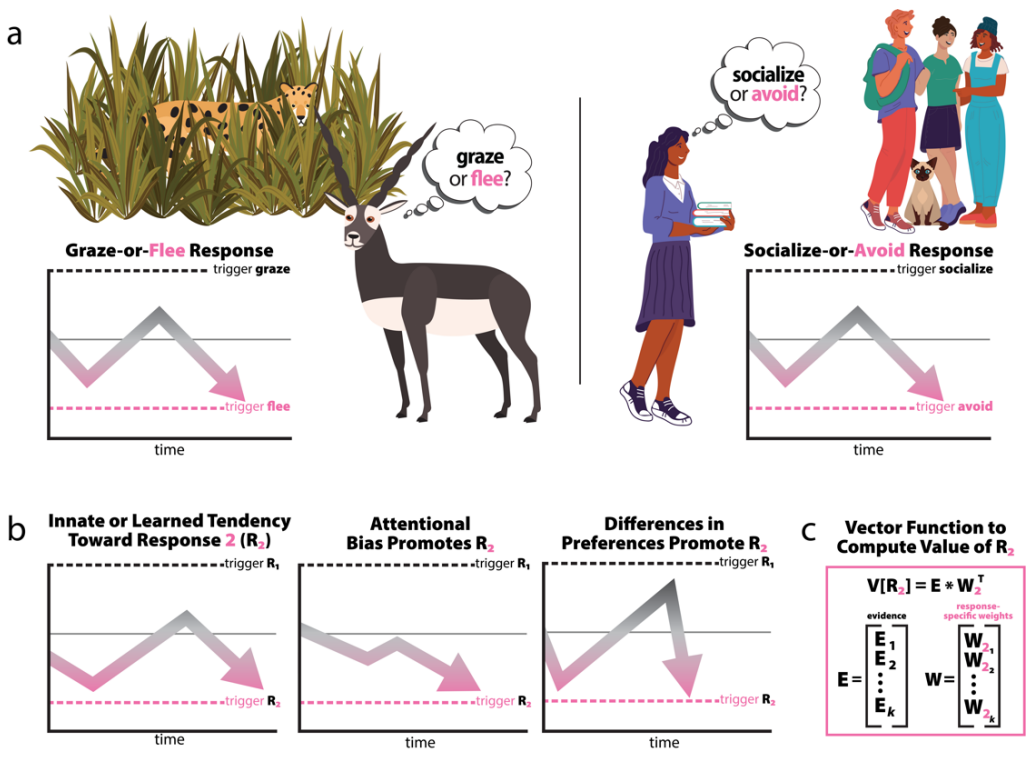

Figure 1. In nature and society, behavior is characterized by risk-vs-reward tradeoffs. **a**) Adaptive responses are selected from competing options. A hungry gazelle detects a predator and must select between grazing and freezing (*left*)—neither of which is inherently maladaptive. This selection process can be defined as a function of the value of each response, i.e., $f(V[R_1], V[R_2])$. While we are agnostic about the specific computations underlying the tradeoffs inherent to response selection, these choices can be conceptualized with simplified drift-diffusion models (DDM; Ratcliff & McKoon, 2008) in which responses are triggered as accumulating evidence surpasses a decision threshold, represented here as dashed lines bisected by a grey line indicating the indifference point. In humans, the systems that underlie these survival-relevant selection processes can select emotion-relevant responses (*right*). **b**) Different underlying processes can trigger the same response. Even with a simplified two-option DDM—which has been useful for characterizing multi-alternative valuation decisions (Krajbich & Rangel, 2011)—different underlying processes can bias individuals toward the same response: an innate or learned tendency toward one response over another (*left*), an attentional bias that leads to disproportionate accumulation of evidence in favor of one response over another (*middle*), or differences in the valuation of evidence between responses (*right*) illustrate sources of bias toward response R_2. **c**) Response selection as a computational process in an *n*-dimensional feature space. The value of any response (e.g., $V[R_2]$) can be conceptualized as the product of all available evidence (e.g., $[E_1, E_2 \ldots E_k]$) times the context-specific weight afforded to each piece of evidence (e.g., $[W_{2_1}, W_{2_2} \ldots W_{2_k}]^T$). In the case of our gazelle, E_1 might represent predator proximity, and W_{2_2} the gazelle's sensitivity to predator proximity in the context of escape decisions. Of note, the weights may comprise a sparse matrix; that is, many pieces of evidence may have no (or little) bearing on a specific response.

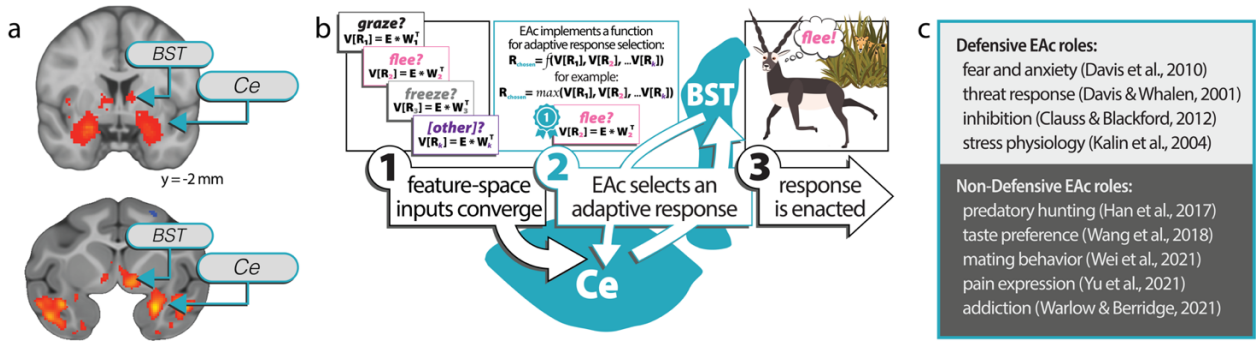

Figure 2. The EAc selects defensive and non-defensive responses. **a**) Studies of humans and rhesus monkeys implicate the EAc in uncertain threat response. As reported in Shackman and Fox, 2016, a *Neurosynth*-enabled (Yarkoni et al., 2011) automated meta-analysis of "fear" and "anxiety" neuroimaging studies in humans reveals Ce and BST activation (*top*), and large-scale (N=592) nonhuman primate neuroimaging studies of response to uncertain threat (Fox et al., 2015a) show that rhesus anxious temperament predicts elevated EAc metabolism during exposure to an uncertain threat represented by an unfamiliar human intruder (*bottom*). **b**) Feature-space model of EAc-implemented function for selecting between graze (R_1), flee (R_2), and freeze (R_3) responses based on the weighted valuations of those responses in each context. In this simplified three-choice model, 1) feature-space inputs encoding salient, weighted environmental and interoceptive evidence converge on the EAc; 2) the EAc represents and resolves the feature space through an unknown selection function (shown here as a placeholder function to represent what is almost certainly a more complicated process; see Krajbich & Rangel, 2011) to guide survival-relevant and emotion-relevant tradeoffs for action selection and adaptive physiology; and 3) instructions to enact the winning response are pushed downstream to effector regions capable of triggering changes in physiology, cognition, and behavior. **c**) An illustrative list of defensive and non-defensive EAc roles highlights the EAc's involvement in diverse response sets. Of note, we use the terms "defensive" and "non-defensive" to be inclusive of physiological, cognitive, and behavioral changes, as well as the phenomenological states that elicit EAc involvement.

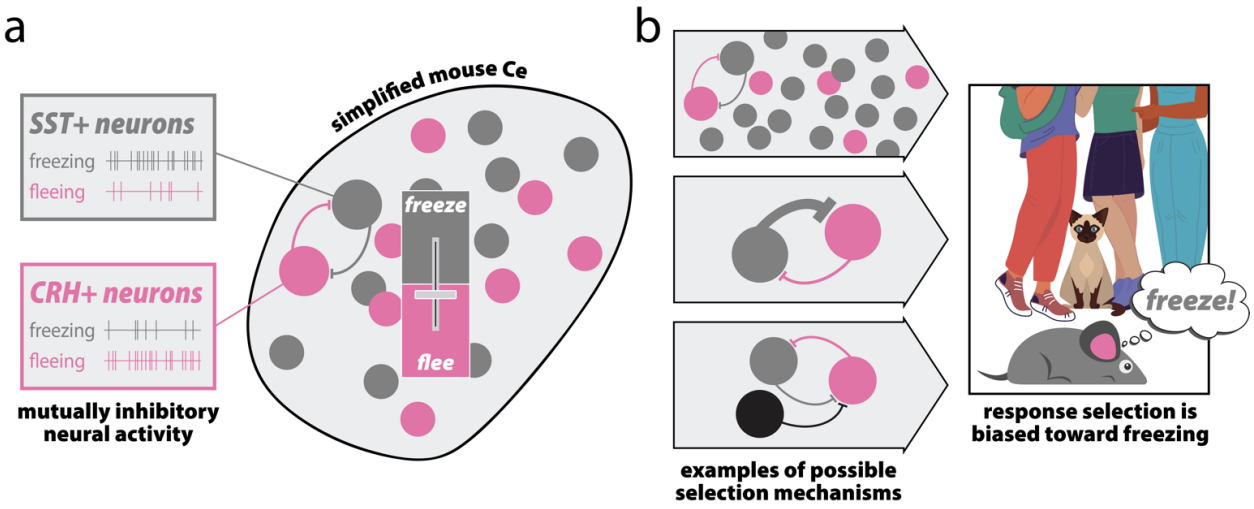

Figure 3. Genetically dissociable microcircuits provide a substrate for response selection through the implementation of a selection function (e.g., $f(V[R_1], V[R_2], ...V[R_k]$; see Fig. 2B). **a**) Mutually inhibitory neural activity in the mouse Ce. A competitive inhibitory microcircuit composed of intermingled, competing populations of SST+ and CRH+ neurons select between freezing and fleeing responses, respectively (adapted from Fadok et al., 2017). The activity of either population generates strong inhibitory postsynaptic currents that suppress the other population, thereby serving as a rapid, winner-take-all mechanism for selecting between active and passive threat response. **b**) Possible mechanisms for response selection. Several distinct mechanisms could dispositionally bias an individual toward passive threat response (i.e., maladaptive freezing), characteristic of behavioral inhibition (Roelofs, 2017; Roelofs, Hagenaars, & Stins, 2010); for example: a preponderance of SST+ neurons (*top*), disproportionately strong SST+ to CRH+ projections (*middle*), or the presence of a third population of neurons that co-inhibits CRH+ neurons (*bottom*). Importantly, while we have highlighted the SST+ and CRH+ microcircuit in the mouse Ce, it is likely that imbalances in other microcircuits—for example, the aforementioned "CeL_{on}"/PKCδ- and "CeL_{off}"/PKCδ+ microcircuit—could drive similar tendencies. We hypothesize similar alterations in other EAc regions, such as the BST.

References:

Aggleton, J. P. (2000). *The amygdala: A functional analysis.* Oxford University Press, Oxford, UK.

Avery, S. N., Clauss, J. A., & Blackford, J. U. (2016). The Human BNST: Functional Role in Anxiety and Addiction. *Neuropsychopharmacology*, *41*(1), 126–141. https://doi.org/10.1038/npp.2015.185

Avery, S. N., Clauss, J. A., Winder, D. G., Woodward, N., Heckers, S., & Blackford, J. U. (2014). BNST neurocircuitry in humans. *NeuroImage, 91*, 311–323. https://doi.org/10.1016/j.neuroimage.2014.01.017

Bachmanov, A. A., Bosak, N. P., Lin, C., Matsumoto, I., Ohmoto, M., Reed, D. R., & Nelson, T. M. (2014). Genetics of Taste Receptors. *Current Pharmaceutical Design*, *20*(16), 2669–2683.

Bandelow, B., Michaelis, S. (2015) Epidemiology of anxiety disorders in the 21st century. *Dialogues in Clinical Neuroscience*, 17(3), 327–35.

Baumgartner, H. M., Schulkin, J., & Berridge, K. C. (2021). Activating Corticotropin-Releasing Factor Systems in the Nucleus Accumbens, Amygdala, and Bed Nucleus of Stria Terminalis: Incentive Motivation or Aversive Motivation? *Biological Psychiatry*, *89*(12), 1162–1175. https://doi.org/10.1016/j.biopsych.2021.01.007

Baxter, M. G., & Murray, E. A. (2002). The amygdala and reward. *Nature Reviews Neuroscience, 3*(7), 563–573. https://doi.org/10.1038/nrn875

Blanchard, D. C., & Blanchard, R. J. (2008). Chapter 2.4 Defensive behaviors, fear, and anxiety. In D. C. B. Robert J. Blanchard Guy Griebel and David Nutt (Ed.), *Handbook of Behavioral Neuroscience: Vol. Volume 17* (pp. 63–79). Elsevier. http://www.sciencedirect.com/science/article/pii/S1569733907000057

Blanchard, D. C., Griebel, G., Pobbe, R., & Blanchard, R. J. (2011). Risk assessment as an evolved threat detection and analysis process. *Neuroscience and Biobehavioral Reviews*, *35*(4), 991–998. https://doi.org/10.1016/j.neubiorev.2010.10.016

Beddington J., Cooper C.L., Field J., Goswami U., Huppert F.A., Jenkins R., Jones H.S., Kirkwood T.B., Sahakian B.J., Thomas S.M. (2008) The mental wealth of nations. *Nature*, 455(7216), 1057-60. doi: 10.1038/4551057a. PMID: 18948946.

Bliss-Moreau, E., Toscano, J. E., Bauman, M., Mason, W. A., & Amaral, D. G. (2010). Neonatal Amygdala or Hippocampus Lesions Influence Responsiveness to Objects. *Developmental Psychobiology*, *52*(5), 487–503. https://doi.org/10.1002/dev.20451

Bliss-Moreau, E., Toscano, J. E., Bauman, M., Mason, W. A., & Amaral, D. G. (2011). Neonatal Amygdala Lesions Alter Responsiveness to Objects in Juvenile Macaques. *Neuroscience*, *178*, 123–132. https://doi.org/10.1016/j.neuroscience.2010.12.038

Botta, P., Demmou, L., Kasugai, Y., Markovic, M., Xu, C., Fadok, J. P., Lu, T., Poe, M. M., Xu, L., Cook, J. M., Rudolph, U., Sah, P., Ferraguti, F., & Lüthi, A. (2015). Regulating anxiety with extrasynaptic inhibition. *Nature Neuroscience*, *18*(10), 1493–1500. https://doi.org/10.1038/nn.4102

Bupesh, M., Abellán, A., & Medina, L. (2011). Genetic and experimental evidence supports the continuum of the central extended amygdala and a mutiple embryonic origin of its principal neurons. *The Journal of Comparative Neurology*, *519*(17), 3507–3531. https://doi.org/10.1002/cne.22719

Cai, H., Haubensak, W., Anthony, T., & Anderson, D. J. (2014). Central amygdala PKC-δ+ neurons mediate the influence of multiple anorexigenic signals. *Nature Neuroscience*, *17*(9), 1240–1248. https://doi.org/10.1038/nn.3767

Chang, L. J., Gianaros, P. J., Manuck, S. B., Krishnan, A., & Wager, T. D. (2015). A Sensitive and Specific Neural Signature for Picture-Induced Negative Affect. *PLoS Biology*, *13*(6), e1002180. https://doi.org/10.1371/journal.pbio.1002180

Charbonneau, J. A., Bennett, J. L., & Bliss-Moreau, E. (2021). Amygdala or hippocampus damage only minimally impacts affective responding to threat. *Behavioral Neuroscience*. https://doi.org/10.1037/bne0000491

Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V., Looger, L. L., Svoboda, K., & Kim, D. S. (2013). Ultra-sensitive fluorescent proteins for imaging neuronal activity. *Nature*, *499*(7458), 295–300. https://doi.org/10.1038/nature12354

Choi, J.-S., & Kim, J. J. (2010). Amygdala regulates risk of predation in rats foraging in a dynamic fear environment. *Proceedings of the National Academy of Sciences*, *107*(50), 21773–21777. https://doi.org/10.1073/pnas.1010079108

Ciocchi, S., Herry, C., Grenier, F., Wolff, S. B. E., Letzkus, J. J., Vlachos, I., Ehrlich, I., Sprengel, R., Deisseroth, K., Stadler, M. B., Müller, C., & Lüthi, A. (2010). Encoding of conditioned fear in central amygdala inhibitory circuits. *Nature*, *468*(7321), 277–282. https://doi.org/10.1038/nature09559

Clauss, J. (2019). Extending the neurocircuitry of behavioural inhibition: A role for the bed nucleus of the stria

terminalis in risk for anxiety disorders. *General Psychiatry*, *32*(6), e100137. https://doi.org/10.1136/gpsych-2019-100137

Cooper, J., William E., & Blumstein, D. T. (Eds.). (2015). *Escaping From Predators: An Integrative View of Escape Decisions*. Cambridge University Press. https://doi.org/10.1017/CBO9781107447189

Davis, M., Walker, D. L., Miles, L., & Grillon, C. (2010). Phasic vs Sustained Fear in Rats and Humans: Role of the Extended Amygdala in Fear vs Anxiety. *Neuropsychopharmacology*, *35*(1), 105–135. https://doi.org/10.1038/npp.2009.109

de Olmos, J. S., & Heimer, L. (1999). The concepts of the ventral striatopallidal system and extended amygdala. *Annals of the New York Academy of Sciences*, *877*, 1–32.

Deisseroth, K. (2011). Optogenetics. *Nature Methods*, *8*(1), 26–29. https://doi.org/10.1038/nmeth.f.324

Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(24), 13763–13768. https://doi.org/10.1073/pnas.231499798

Emery, N. J., Capitanio, J. P., Mason, W. A., Machado, C. J., Mendoza, S. P., & Amaral, D. G. (2001). The effects of bilateral lesions of the amygdala on dyadic social interactions in rhesus monkeys (Macaca mulatta). *Behavioral Neuroscience*, *115*(3), 515–544.

Evans, D. A., Stempel, A. V., Vale, R., Ruehle, S., Lefler, Y., & Branco, T. (2018). A synaptic threshold mechanism for computing escape decisions. *Nature*, *558*(7711), 590–594. https://doi.org/10.1038/s41586-018-0244-6

Fadok, J. P., Krabbe, S., Markovic, M., Courtin, J., Xu, C., Massi, L., Botta, P., Bylund, K., Müller, C., Kovacevic, A., Tovote, P., & Lüthi, A. (2017). A competitive inhibitory circuit for selection of active and passive fear responses. *Nature*, *542*(7639), 96–100. https://doi.org/10.1038/nature21047

Flanigan, M. E., & Kash, T. L. (n.d.). Coordination of social behaviors by the bed nucleus of the stria terminalis. *European Journal of Neuroscience*, *n/a*(n/a). https://doi.org/10.1111/ejn.14991

Fox, A. S., & Kalin, N. H. (2014). A translational neuroscience approach to understanding the development of social anxiety disorder and its pathophysiology. *The American Journal of Psychiatry*, *171*(11), 1162–1173. https://doi.org/10.1176/appi.ajp.2014.14040449

Fox, A. S., Oler, J. A., Birn, R. M., Shackman, A. J., Alexander, A. L., & Kalin, N. H. (2018). Functional Connectivity within the Primate Extended Amygdala Is Heritable and Associated with Early-Life Anxious Temperament. *Journal of Neuroscience*, *38*(35), 7611–7621. https://doi.org/10.1523/JNEUROSCI.0102-18.2018

Fox, A. S., Oler, J. A., Shackman, A. J., Shelton, S. E., Raveendran, M., McKay, D. R., Converse, A. K., Alexander, A., Davidson, R. J., Blangero, J., Rogers, J., & Kalin, N. H. (2015a). Intergenerational neural mediators of early-life anxious temperament. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(29), 9118–9122. https://doi.org/10.1073/pnas.1508593112

Fox, A. S., & Shackman, A. J. (2019). The central extended amygdala in fear and anxiety: Closing the gap between mechanistic and neuroimaging research. *Neuroscience Letters, 693*, 58–67. https://doi.org/10.1016/j.neulet.2017.11.056

Fox, A. S., Oler, J. A., Tromp, D. P. M., Fudge, J. L., & Kalin, N. H. (2015b). Extending the amygdala in theories of threat processing. *Trends in Neurosciences*, *38*(5), 319–329. https://doi.org/10.1016/j.tins.2015.03.002

Fox, A. S., Shelton, S. E., Oakes, T. R., Davidson, R. J., & Kalin, N. H. (2008). Trait-Like Brain Activity during Adolescence Predicts Anxious Temperament in Primates. *PLoS ONE*, *3*(7). https://doi.org/10.1371/journal.pone.0002570

Frick, A., Gingnell, M., Marquand, A. F., Howner, K., Fischer, H., Kristiansson, M., Williams, S. C. R., Fredrikson, M., & Furmark, T. (2014). Classifying social anxiety disorder using multivoxel pattern analyses of brain function and structure. *Behavioural Brain Research*, *259*, 330–335. https://doi.org/10.1016/j.bbr.2013.11.003

Garakani, A., Murrough, J. W., Freire, R. C., Thom, R. P., Larkin, K., Buono, F. D., & Iosifescu, D. V. (2020). Pharmacotherapy of Anxiety Disorders: Current and Emerging Treatment Options. *Frontiers in Psychiatry*, *11*. https://www.frontiersin.org/article/10.3389/fpsyt.2020.595584

Giardino, W. J., Eban-Rothschild, A., Christoffel, D. J., Li, S. B., Malenka, R. C., & de Lecea, L. (2018). Parallel circuits from the bed nuclei of stria terminalis to the lateral hypothalamus drive opposing emotional states. *Nature Neuroscience*, *21*(8), 1084–1095. https://doi.org/10.1038/s41593-018-0198-x

Giardino, W. J., & Pomrenze, M. B. (2021). Extended amygdala neuropeptide circuitry of emotional arousal: Waking up on the wrong side of the bed nuclei of stria terminalis. *Frontiers in Behavioral Neuroscience*, *15*, 613025. https://doi.org/10.3389/fnbeh.2021.613025

Gorka, A. X., Torrisi, S., Shackman, A. J., Grillon, C., & Ernst, M. (2018). Intrinsic functional connectivity of the central

nucleus of the amygdala and bed nucleus of the stria terminalis. *NeuroImage, 168*, 392–402. https://doi.org/10.1016/j.neuroimage.2017.03.007

Gray, T. S., & Magnuson, D. J. (1992). Peptide immunoreactive neurons in the amygdala and the bed nucleus of the stria terminalis project to the midbrain central gray in the rat. *Peptides*, *13*(3), 451–460. https://doi.org/10.1016/0196-9781(92)90074-d

Grillon, C. (2008). Models and mechanisms of anxiety: Evidence from startle studies. *Psychopharmacology*, *199*(3), 421–437. https://doi.org/10.1007/s00213-007-1019-1

Gungor, N. Z., & Paré, D. (2016). Functional heterogeneity in the bed nucleus of the stria terminalis. *The Journal of Neuroscience*, *36*(31), 8038–8049. https://doi.org/10.1523/JNEUROSCI.0856-16.2016

Han, W., Tellez, L. A., Rangel, M. J., Motta, S. C., Zhang, X., Perez, I. O., Canteras, N. S., Shammah-Lagnado, S. J., van den Pol, A. N., & de Araujo, I. E. (2017). Integrated Control of Predatory Hunting by the Central Nucleus of the Amygdala. *Cell*, *168*(1–2), 311-324.e18. https://doi.org/10.1016/j.cell.2016.12.027

Hardaway, J. A., Halladay, L. R., Mazzone, C. M., Pati, D., Bloodgood, D. W., Kim, M., Jensen, J., DiBerto, J. F., Boyt, K. M., Shiddapur, A., Erfani, A., Hon, O. J., Neira, S., Stanhope, C. M., Sugam, J. A., Saddoris, M. P., Tipton, G., McElligott, Z., Jhou, T. C., … Kash, T. L. (2019). Central Amygdala Prepronociceptin-Expressing Neurons Mediate Palatable Food Consumption and Reward. *Neuron*, *102*(5), 1037-1052.e7. https://doi.org/10.1016/j.neuron.2019.03.037

Haubensak, W., Kunwar, P. S., Cai, H., Ciocchi, S., Wall, N. R., Ponnusamy, R., Biag, J., Dong, H.-W., Deisseroth, K., Callaway, E. M., Fanselow, M. S., Lüthi, A., & Anderson, D. J. (2010). Genetic dissection of an amygdala microcircuit that gates conditioned fear. *Nature*, *468*(7321), 270–276. https://doi.org/10.1038/nature09553

Holley, D., Campos, L. J., Zhang, Y., Capitanio, J. P., & Fox, A. S. (2022). Rhesus nervous temperament predicts peri-adolescent central amygdala metabolism and behavioral inhibition measured by a machine-learning approach. *BioRxiv* https://doi.org/10.1101/2022.07.26.501512

Hur, J., Kuhn, M., Grogans, S. E., Anderson, A. S., Islam, S., Kim, H. C., Tillman, R. M., Fox, A. S., Smith, J. F., DeYoung, K. A., Shackman, A. J. (*in press*). Anxiety-related fronto-cortical activity is associated with dampened stressor reactivity in the real world. *Psychological Science*. [coordinate table on bioRxiv; maps on NeuroVault]

Hur, J., Smith, J. F., DeYoung, K. A., Anderson, A. S., Kuang, J., Kim, H. C., Tillman, R. M., Kuhn, M., Fox, A. S., & Shackman, A. J. (2020). Anxiety and the Neurobiology of Temporally Uncertain Threat Anticipation. *The Journal of Neuroscience*, *40*(41), 7949–7964. https://doi.org/10.1523/JNEUROSCI.0704-20.2020

Isosaka, T., Matsuo, T., Yamaguchi, T., Funabiki, K., Nakanishi, S., Kobayakawa, R., & Kobayakawa, K. (2015). Htr2a-Expressing Cells in the Central Amygdala Control the Hierarchy between Innate and Learned Fear. *Cell*, *163*(5), 1153–1164. https://doi.org/10.1016/j.cell.2015.10.047

Kalin, N. H., Shelton, S. E., & Davidson, R. J. (2004). The role of the central nucleus of the amygdala in mediating fear and anxiety in the primate. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *24*(24), 5506–5515. https://doi.org/10.1523/JNEUROSCI.0292-04.2004

Khalsa, S. S., Feinstein, J. S., Li, W., Feusner, J. D., Adolphs, R., & Hurlemann, R. (2016). Panic Anxiety in Humans with Bilateral Amygdala Lesions: Pharmacological Induction via Cardiorespiratory Interoceptive Pathways. *The Journal of Neuroscience*, *36*(12), 3559–3566. https://doi.org/10.1523/JNEUROSCI.4109-15.2016

Kim, J., Zhang, X., Muralidhar, S., LeBlanc, S. A., & Tonegawa, S. (2017). Basolateral to central amygdala neural circuits for appetitive behaviors. *Neuron*, *93*(6), 1464-1479.e5. https://doi.org/10.1016/j.neuron.2017.02.034

Kim, S. Y., Adhikari, A., Lee, S. Y., Marshel, J. H., Kim, C. K., Mallory, C. S., Lo, M., Pak, S., Mattis, J., Lim, B. K., Malenka, R. C., Warden, M. R., Neve, R., Tye, K. M., & Deisseroth, K. (2013). Diverging neural pathways assemble a behavioural state from separable features in anxiety. *Nature*, *496*(7444), 219–223. https://doi.org/10.1038/nature12018

Koen, N., & Stein, D. J. (2011). Pharmacotherapy of anxiety disorders: A critical review. *Dialogues in Clinical Neuroscience*, *13*(4), 423–437.

Kovner, R., Souaiaia, T., Fox, A. S., French, D. A., Goss, C. E., Roseboom, P. H., Oler, J. A., Riedel, M. K., Fekete, E. M., Fudge, J. L., Knowles, J. A., & Kalin, N. H. (2020). Transcriptional Profiling of Primate Central Nucleus of the Amygdala Neurons to Understand the Molecular Underpinnings of Early-Life Anxious Temperament. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2020.05.009

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(33), 13852–13857. https://doi.org/10.1073/pnas.1101328108

LeDoux, J. E., & Pine, D. S. (2016). Using neuroscience to help understand fear and anxiety: A two-system framework. *The American Journal of Psychiatry, 173*(11), 1083–1093. https://doi.org/10.1176/appi.ajp.2016.16030353

Lebow, M. A., & Chen, A. (2016). Overshadowed by the amygdala: The bed nucleus of the stria terminalis emerges as key to psychiatric disorders. *Molecular Psychiatry, 21*(4), 450–463. https://doi.org/10.1038/mp.2016.1

Lein, E. S., Hawrylycz, M. J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A. F., Boguski, M. S., Brockway, K. S., Byrnes, E. J., Chen, L., Chen, L., Chen, T.-M., Chin, M. C., Chong, J., Crook, B. E., Czaplinska, A., Dang, C. N., Datta, S., … Jones, A. R. (2007). Genome-wide atlas of gene expression in the adult mouse brain. *Nature*, *445*(7124), 168–176. https://doi.org/10.1038/nature05453

Li, B. (2019). Central amygdala cells for learning and expressing aversive emotional memories. *Current Opinion in Behavioral Sciences*, *26*, 40–45. https://doi.org/10.1016/j.cobeha.2018.09.012

Lissek, S., Kaczkurkin, A. N., Rabin, S., Geraci, M., Pine, D. S., & Grillon, C. (2014). Generalized Anxiety Disorder is Associated with Overgeneralization of Classically Conditioned-Fear. *Biological Psychiatry*, *75*(11), 909–915. https://doi.org/10.1016/j.biopsych.2013.07.025

Lissek, S., Rabin, S., Heller, R. E., Lukenbaugh, D., Geraci, M., Pine, D. S., & Grillon, C. (2010). Overgeneralization of Conditioned Fear as a Pathogenic Marker of Panic Disorder. *The American Journal of Psychiatry*, *167*(1), 47–55. https://doi.org/10.1176/appi.ajp.2009.09030410

Liu, F., Guo, W., Fouche, J.-P., Wang, Y., Wang, W., Ding, J., Zeng, L., Qiu, C., Gong, Q., Zhang, W., & Chen, H. (2015). Multivariate classification of social anxiety disorder using whole brain functional connectivity. *Brain Structure & Function*, *220*(1), 101–115. https://doi.org/10.1007/s00429-013-0641-4

Ma, Y., & Giardino, W. J. (2022). Neural circuit mechanisms of the cholecystokinin (CCK) neuropeptide system in addiction. *Addiction Neuroscience, 3,* 1-6. https://doi.org/10.1016/j.addicn.2022.100024

Machado, C. J., & Bachevalier, J. (2008). Behavioral and hormonal reactivity to threat: Effects of selective amygdala, hippocampal or orbital frontal lesions in monkeys. *Psychoneuroendocrinology*, *33*(7), 926–941. https://doi.org/10.1016/j.psyneuen.2008.04.012

Machado, C. J., Emery, N. J., Mason, W. A., & Amaral, D. G. (2010). Selective changes in foraging behavior following bilateral neurotoxic amygdala lesions in rhesus monkeys. *Behavioral Neuroscience*, *124*(6), 761–772. https://doi.org/10.1037/a0021560

Mahler, S. V., & Berridge, K. C. (2012). What and when to "want"? Amygdala-based focusing of incentive salience upon sugar and sex. *Psychopharmacology, 221*(3), 407–426. https://doi.org/10.1007/s00213-011-2588-6

Marcinkiewcz, C. A., Mazzone, C. M., D'Agostino, G., Halladay, L. R., Hardaway, J. A., DiBerto, J. F., Navarro, M., Burnham, N., Cristiano, C., Dorrier, C. E., Tipton, G. J., Ramakrishnan, C., Kozicz, T., Deisseroth, K., Thiele, T. E., McElligott, Z. A., Holmes, A., Heisler, L. K., & Kash, T. L. (2016). Serotonin engages an anxiety and fear-promoting circuit in the extended amygdala. *Nature*, *537*(7618), 97–101. https://doi.org/10.1038/nature19318

McDonald, A. J. (1982). Cytoarchitecture of the central amygdaloid nucleus of the rat. *The Journal of Comparative Neurology*, *208*(4), 401–418. https://doi.org/10.1002/cne.902080409

McDonald, A. J. (1983). Neurons of the bed nucleus of the stria terminalis: A golgi study in the rat. *Brain Research Bulletin*, *10*(1), 111–120.

Mobbs, D., Hagan, C. C., Dalgleish, T., Silston, B., & Prévost, C. (2015). The ecology of human fear: Survival optimization and the nervous system. *Frontiers in Neuroscience*, *9*. https://doi.org/10.3389/fnins.2015.00055

Mobbs, D., Marchant, J. L., Hassabis, D., Seymour, B., Tan, G., Gray, M., Petrovic, P., Dolan, R. J., & Frith, C. D. (2009). From Threat to Fear: The Neural Organization of Defensive Fear Systems in Humans. *The Journal of Neuroscience*, *29*(39), 12236–12243. https://doi.org/10.1523/JNEUROSCI.2378-09.2009

Morey, R. A., Haswell, C. C., Stjepanović, D., Dunsmoor, J. E., & LaBar, K. S. (2020). Neural correlates of conceptual-level fear generalization in posttraumatic stress disorder. *Neuropsychopharmacology*, *45*(8), 1380–1389. https://doi.org/10.1038/s41386-020-0661-8

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*(9), 424–430. https://doi.org/10.1016/j.tics.2006.07.005

Oler, J. A., Birn, R. M., Patriat, R., Fox, A. S., Shelton, S. E., Burghy, C. A., Stodola, D. E., Essex, M. J., Davidson, R. J., & Kalin, N. H. (2012). Evidence for coordinated functional activity within the extended amygdala of non-human and human primates. *NeuroImage*, *61*(4), 1059–1066. https://doi.org/10.1016/j.neuroimage.2012.03.045

Oler, J. A., Fox, A. S., Shelton, S. E., Rogers, J., Dyer, T. D., Davidson, R. J., Shelledy, W., Oakes, T. R., Blangero,

J., & Kalin, N. H. (2010). Amygdalar and hippocampal substrates of anxious temperament differ in their heritability. *Nature*, *466*(7308), 864–868. https://doi.org/10.1038/nature09282

Oler, J. A., Tromp, D. P. M., Fox, A. S., Kovner, R., Davidson, R. J., Alexander, A. L., McFarlin, D. R., Birn, R. M., E Berg, B., deCampo, D. M., Kalin, N. H., & Fudge, J. L. (2017). Connectivity between the central nucleus of the amygdala and the bed nucleus of the stria terminalis in the non-human primate: Neuronal tract tracing and developmental neuroimaging studies. *Brain Structure & Function*, *222*(1), 21–39. https://doi.org/10.1007/s00429-016-1198-9

Palmiter, R. D. (2018). The parabrachial nucleus: CGRP neurons function as a general alarm. *Trends in Neurosciences*, *41*(5), 280–293. https://doi.org/10.1016/j.tins.2018.03.007

Parkinson, J. A., Crofts, H. S., McGuigan, M., Tomic, D. L., Everitt, B. J., & Roberts, A. C. (2001). The role of the primate amygdala in conditioned reinforcement. *The Journal of Neuroscience, 21*(19), 7770–7780. https://doi.org/10.1523/JNEUROSCI.21-19-07770.2001

Perusini, J. N., & Fanselow, M. S. (2015). Neurobehavioral perspectives on the distinction between fear and anxiety. *Learning & Memory, 22*(9), 417–425. https://doi.org/10.1101/lm.039180.115

Ponserre, M., Fermani, F., & Klein, R. (2020). *Encoding of environmental cues in central amygdala neurons during foraging* (p. 2020.09.28.313056). https://doi.org/10.1101/2020.09.28.313056

Quintana, D. S., Lischke, A., Grace, S., Scheele, D., Ma, Y., & Becker, B. (2021). Advances in the field of intranasal oxytocin research: Lessons learned and future directions for clinical research. *Molecular Psychiatry, 26*(1), 80–91. https://doi.org/10.1038/s41380-020-00864-7

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews. Neuroscience*, *9*(7), 545–556. https://doi.org/10.1038/nrn2357

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Resendez, S. L., & Stuber, G. D. (2015). In vivo Calcium Imaging to Illuminate Neurocircuit Activity Dynamics Underlying Naturalistic Behavior. *Neuropsychopharmacology*, *40*(1), 238–239. https://doi.org/10.1038/npp.2014.206

Rinker, J. A., Marshall, S. A., Mazzone, C. M., Lowery-Gionta, E. G., Gulati, V., Pleil, K. E., Kash, T. L., Navarro, M., & Thiele, T. E. (2017). Extended Amygdala to Ventral Tegmental Area Corticotropin-Releasing Factor Circuit Controls Binge Ethanol Intake. *Biological Psychiatry*, *81*(11), 930–940. https://doi.org/10.1016/j.biopsych.2016.02.029

Roelofs K. (2017). Freeze for action: neurobiological mechanisms in animal and human freezing. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, *372*(1718), 20160206. https://doi.org/10.1098/rstb.2016.0206

Roelofs, K., Hagenaars, M. A., & Stins, J. (2010). Facing freeze: social threat induces bodily freeze in humans. *Psychological Science*, *21*(11), 1575–1581. https://doi.org/10.1177/0956797610384746

Roth, B. L. (2016). DREADDs for Neuroscientists. *Neuron*, *89*(4), 683–694. https://doi.org/10.1016/j.neuron.2016.01.040

Roy, A. K., Shehzad, Z., Margulies, D. S., Kelly, A. M. C., Uddin, L. Q., Gotimer, K., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2009). Functional connectivity of the human amygdala using resting state fMRI. *NeuroImage*, *45*(2), 614–626. https://doi.org/10.1016/j.neuroimage.2008.11.030

Samson, R. D., & Paré, D. (2005). Activity-dependent synaptic plasticity in the central nucleus of the amygdala. *The Journal of Neuroscience, 25*(7), 1847–1855. https://doi.org/10.1523/JNEUROSCI.3713-04.2005

Sanford, C. A., Soden, M. E., Baird, M. A., Miller, S. M., Schulkin, J., Palmiter, R. D., Clark, M., & Zweifel, L. S. (2017). A central amygdala CRF circuit facilitates learning about weak threats. *Neuron, 93*(1), 164–178. https://doi.org/10.1016/j.neuron.2016.11.034

Shackman, A. J., & Fox, A. S. (2021). Two decades of anxiety neuroimaging research: New insights and a look to the future. *The American Journal of Psychiatry, 178*(2), 106–109. https://doi.org/10.1176/appi.ajp.2020.20121733Shackman, A. J., & Fox, A. S. (2016). Contributions of the Central Extended Amygdala to Fear and Anxiety. *The Journal of Neuroscience*, *36*(31), 8050–8063. https://doi.org/10.1523/JNEUROSCI.0982-16.2016

Shackman, A. J., Tromp, D. P. M., Stockbridge, M. D., Kaplan, C. M., Tillman, R. M., & Fox, A. S. (2016). Dispositional negativity: An integrative psychological and neurobiological perspective. *Psychological Bulletin*, *142*(12), 1275–1314. https://doi.org/10.1037/bul0000073

Sharp, T., & Barnes, N. M. (2020). Central 5-HT receptors and their function; present and future.

Neuropharmacology, 177, 108155. https://doi.org/10.1016/j.neuropharm.2020.108155
Somerville, L. H., Wagner, D. D., Wig, G. S., Moran, J. M., Whalen, P. J., & Kelley, W. M. (2013). Interactions between transient and sustained neural signals support the generation and regulation of anxious emotion. *Cerebral Cortex (New York, N.Y.: 1991)*, *23*(1), 49–60. https://doi.org/10.1093/cercor/bhr373
Swanson, L. W., & Petrovich, G. D. (1998). What is the amygdala? *Trends in Neurosciences*, *21*(8), 323–331. https://doi.org/10.1016/s0166-2236(98)01265-x
Tillman, R. M., Stockbridge, M. D., Nacewicz, B. M., Torrisi, S., Fox, A. S., Smith, J. F., & Shackman, A. J. (2018). Intrinsic functional connectivity of the central extended amygdala. *Human Brain Mapping*, *39*(3), 1291–1312. https://doi.org/10.1002/hbm.23917
Tovote, P., Esposito, M. S., Botta, P., Chaudun, F., Fadok, J. P., Markovic, M., Wolff, S. B. E., Ramakrishnan, C., Fenno, L., Deisseroth, K., Herry, C., Arber, S., & Lüthi, A. (2016). Midbrain circuits for defensive behaviour. *Nature*, *534*(7606), 206–212. https://doi.org/10.1038/nature17996
U.S. Burden of Disease Collaborators: Mokdad, A.H., Ballestros, K., Echko, M., Glenn, S., Olsen, H.E., et al. (2018). The state of US Health, 1990-2016: Burden of diseases, injuries, and risk factors among U.S. states. JAMA, 319(14):1444–1472
Walf, A. A., & Frye, C. A. (2007). The use of the elevated plus maze as an assay of anxiety-related behavior in rodents. *Nature Protocols, 2*(2), 322–328. https://doi.org/10.1038/nprot.2007.44
Walker, D. L., & Davis, M. (2008). Role of the extended amygdala in short-duration versus sustained fear: A tribute to Dr. Lennart Heimer. *Brain Structure & Function*, *213*(1–2), 29–42. https://doi.org/10.1007/s00429-008-0183-3
Walker, D. L., Miles, L. A., & Davis, M. (2009). Selective Participation of the Bed Nucleus of the Stria Terminalis and CRF in Sustained Anxiety-Like versus Phasic Fear-Like Responses. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *33*(8), 1291–1308. https://doi.org/10.1016/j.pnpbp.2009.06.022
Wang, L., Gillis-Smith, S., Peng, Y., Zhang, J., Chen, X., Daniel Salzman, C., Ryba, N. J. P., & Zuker, C. S. (2018). The coding of valence and identity in the mammalian taste system. *Nature*, *558*(7708), 127–131. https://doi.org/10.1038/s41586-018-0165-4
Warlow, S. M., & Berridge, K. C. (2021). Incentive motivation: 'wanting' roles of central amygdala circuitry. *Behavioural Brain Research*, *411*, 113376. https://doi.org/10.1016/j.bbr.2021.113376
Wei, D., Talwar, V., & Lin, D. (2021). Neural circuits of social behaviors: Innate yet flexible. *Neuron*. https://doi.org/10.1016/j.neuron.2021.02.012
Whalen, P. J., & Phelps, E. A. (2009). *The human amygdala*. Guilford Press, New York City, NY.
Woo, C.-W., Chang, L. J., Lindquist, M. A., & Wager, T. D. (2017). Building better biomarkers: Brain models in translational neuroimaging. *Nature Neuroscience*, *20*(3), 365–377. https://doi.org/10.1038/nn.4478
Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, *8*(8), 665–670. https://doi.org/10.1038/nmeth.1635
Yu, K., Ahrens, S., Zhang, X., Schiff, H., Ramakrishnan, C., Fenno, L., Deisseroth, K., Zhao, F., Luo, M. H., Gong, L., He, M., Zhou, P., Paninski, L., & Li, B. (2017). The central amygdala controls learning in the lateral amygdala. *Nature Neuroscience, 20*(12), 1680–1685. https://doi.org/10.1038/s41593-017-0009-9
Yu, W., Pati, D., Pina, M. M., Schmidt, K. T., Boyt, K. M., Hunker, A. C., Zweifel, L. S., McElligott, Z. A., & Kash, T. L. (2021). Periaqueductal gray/dorsal raphe dopamine neurons contribute to sex differences in pain-related behaviors. *Neuron*, *109*(8), 1365-1380.e5. https://doi.org/10.1016/j.neuron.2021.03.001
Zeng, H. (2022). What is a cell type and how to define it? *Cell, 185*(18), 2739-2755, https://doi.org/10.1016/j.cell.2022.06.031
Zorrilla, E. P., & Koob, G. F. (2010). Progress in corticotropin-releasing factor-1 antagonist development. *Drug Discovery Today, 15*(9-10), 371–383. https://doi.org/10.1016/j.drudis.2010.02.011

Reimagining “uncertainty” in the study of threat anticipation: A statistical-learning approach

Holley, D, & Fox, AS

Abstract

Anxiety disorders are among the most common psychiatric conditions and impair the well-being of individuals, organizations, and society as a whole (Bandelow & Michaelis, 2015; Kessler et al., 2012; Beddington et al., 2008). Refining our understanding of the processes that make us anxious and contribute to these disorders promises to inform new treatment and prevention approaches that blunt their deleterious effects. In the past decades, sensitivity to uncertainty has emerged as a transdiagnostic marker of anxiety disorders (Grupe & Nitschke, 2013). In laboratory studies, uncertain-threat conditions are consistently found to be more anxiogenic than certain-threat conditions (Schmitz & Grillion, 2012; Somerville et al., 2013; Hur et al., 2020). However, we currently have no explanation as to *why* individuals become more anxious in uncertain-threat contexts. One way to approach this challenge is to parameterize components of uncertainty and test their effects on behavioral and emotional outcomes. Here, we developed and validated a computational model of uncertain-threat anticipation that isolates the *hazard rate*—that is, the evolving probability that a threat will occur, taking past information into account—from experimenter-defined momentary threat probabilities [i.e., P(shock)] at specific timepoints. By manipulating the hazard rate while holding P(shock) constant in a statistical threat-learning task in which participants balanced the risk of receiving aversive shocks against the pursuit of cash rewards (N=42), we showed that the hazard rate drives threat avoidance and subjective anxiety on a momentary basis, irrespective of P(shock). These findings suggest that individuals compute, track, and leverage specific environmental threat statistics to optimize risk-vs-reward tradeoffs. Our results shed new light on the computational architecture of threat processing and set the stage for model-based studies into the effects of statistical manipulations on anxiety and its neurobiological substrates.

Author affiliations: Department of Psychology and the California National Primate Research Center at the University of California, Davis, Davis, CA, 95616

Keywords: anxiety, uncertainty, threat processing, computational modeling, statistical learning

Introduction

Anxiety disorders affect an estimated 1 in 3 people during their lifetime and are a leading cause of disability worldwide (Bandelow & Michaelis, 2015; World Health Organization, 2017). Anxiety is characterized by a behavioral tendency to avoid potential threats, and a concomitant emotional experience (Davis, Walker, Miles, & Grillion, 2010). When it is adaptive, anxiety can focus our attention and prime our bodies and minds to meet challenges and seize opportunities. However, when anxiety is extreme, prolonged, or contextually inappropriate, it can lead to disorders that impair our ability to function and thrive (Fox & Kalin, 2014; Shackman et al., 2016; Meacham & Bergstrom, 2012). These disorders rarely occur in isolation and are commonly comorbid with depression (Gorman, 1996; Garber & Weersing, 2010; Wu & Fang, 2014) and substance use disorders (Swendsen et al., 2010). They predict a range of adverse life outcomes, such as suicidal ideation and suicide attempts (Nepon, Belik, Bolton, & Sareen, 2010; Sareen et al., 2005), and are substantially more prevalent in women than in men (McLean, Asnaani, Litz, & Hoffman, 2011). Because existing treatments are inconsistently effective and often cause adverse side effects (Bystritsky, 2006; Griebel & Holmes, 2013), building an improved understanding of the pathophysiology of anxiety disorders is imperative and has the potential to guide new treatments and prevention strategies. Despite this need, key aspects of anxiety are poorly understood.

Decomposing “Uncertainty” into a Set of Tractable Parameters

Heightened sensitivity to uncertain-threat anticipation is a transdiagnostic marker of anxiety that cuts across phenotypes (Grupe & Nitschke, 2013; Yassa, Hazlett, Stark, & Hoehn-Saric, 2012; Krain et al., 2008; Simmons, Matthews, Paulus, & Stein, 2008), paradigms (Shackman et al., 2016), and theoretical perspectives (Hur et al., 2020; Grupe & Nitschke, 2013; Blanchard, Griebel,

Pobbe, & Blanchard, 2011; Moors, Ellsworth, Scherer, & Frijda, 2013; Barrett, Quigley, & Hamilton, 2016; Anderson, Carleton, Diefenbach, & Han, 2019; Mobbs et al., 2009; Mobbs et al., 2015; Moscarrello & Penzo, 2022). However, “uncertainty” in the context of psychological and neuroscientific research is often vaguely defined, leading to an unreliable literature and presenting barriers to experimental replication, extension, and cross-species translation. Furthermore, although uncertain-threat anticipation is widely accepted as intrinsically more anxiogenic than certain threat anticipation, virtually nothing is known about *why* this is the case. This may be because the term “uncertainty” is sufficiently imprecise as to allow for multiple interpretations. To elucidate the relationships between environmental stimuli and anxiogenesis, we must first identify precise, tractable “uncertainty”-associated parameters.

Computational psychiatry is ideally suited to advance our understanding of uncertain-threat anticipation through model-based approaches that parameterize “uncertainty,” thereby enabling characterizations of aberrant processes that contribute to disorder (Montague, Dolan, Friston, & Dayan, 2012). Statistical characterizations of the environment have led to radical advances in our understanding of language (Kuhl, 2004), vision (Krajbich & Rangel, 2011), and behavioral economics (Krajbich, Lu, Camerer, & Rangel, 2012; Rangel, Camerer, & Montague, 2008), highlighting the utility of computational approaches in the study of psychological phenomena. By contrast, clinical studies of anxiety disorders often focus on contexts that elicit disordered responses (e.g., *social* anxiety disorder), or features of the disordered response (e.g., *panic* attack). Computational psychiatry approaches aim to integrate with the extant literature and augment future studies by bringing a higher degree of statistical sophistication to these efforts, thereby helping move the field beyond coarse, categorical variables (i.e., “certain” vs “uncertain”) and toward the parametric manipulation of specific anxiogenic components. To extend previous work centered on comparisons between “certain” and “uncertain” conditions (e.g., Somerville et al., 2013; Hur et al., 2020), we identified features of the statistical environment that were

confounded in certain-vs-uncertain paradigms, and developed a task that enabled us to dissociate and manipulate those features in novel “uncertain A”-vs-“uncertain B” comparisons (Fig. 1).

A Computational Model of Uncertain-Threat Anticipation

Using a theory-based approach, we reasoned that anxiety may increase during uncertain anticipation because people update their expectations as time passes. We posit that tracking the probability dynamics of potential threats is necessary to optimize adaptive responses in real-time, because doing so allows individuals to appropriately balance risk-vs-reward tradeoffs. Notably, the probability dynamics over time are dramatically different in certain and uncertain contexts. In the context of *certain*-threat anticipation, the computation governing these estimates would be straightforward: If an individual knows that a countdown will end with a shock at time 0, then P(shock) at timepoints 3, 2, 1, and 0 would equal 0, 0, 0, and 1, respectively (Fig. 1A, B). The computation becomes more complex—and, for our purposes, more interesting—during a period of *uncertain*-threat anticipation, when the individual knows that a shock will occur *sometime during* the countdown, but not necessarily at 0. In this case, the individual must draw from experience to continuously estimate the evolution of P(shock) over time. For example, in the above countdown, at time 3 the shock could occur at one of four potential times, whereas at time 1, the shock could only occur at one of two times. Assuming uniform-distributed P(shock) values of .25 at each of four timepoints during a period of uncertain-threat anticipation, the *hazard rate* (Goel, Khanna, & Kishore, 2010)—that is, P(shock), given that shock has not occurred yet—for timepoints 3, 2, 1, and 0 would equal .25, .33, .5, and 1, respectively (see *Methods*). These statistical dynamics motivate our interest in exploring the dissociable contributions of P(shock) values (in the above example, [.25, .25, .25, .25]) and hazard-rate values (i.e., [.25, .33, .5, 1]) to anxious behavioral and emotional responses in uncertain-threat contexts.

Our Approach

Uncertain-threat anticipation elicits characteristic avoidant behaviors and concomitant emotional responses (Davis et al., 2010; Shackman et al., 2016). To test the individual contributions of P(shock) and hazard-rate values to anxious behaviors and feelings, we designed a statistical threat-learning task that allowed us to hold momentary P(shock) values constant while manipulating the hazard rate (Fig. 1C, D; detailed in *Methods*) as participants made decisions about whether to persist (and earn greater rewards) or escape (and avoid shocks) in distinct uncertain-threat environments.

Shock Workups and Training Phase

In our study, N=42 participants first underwent a shock workup to determine a shock voltage that was "uncomfortable but not unbearable." Next, they underwent a roughly 30-minute training phase during which two innocuous shapes were paired, at random, with two statistically distinct uncertain-threat environments that were matched for P(shock) but which differed in hazard rate (Fig. 1D). In the training phase, participants saw each shape 50 times, in a predetermined and pseudo-randomized trial order. Each trial ended in the administration of an aversive electrical shock at one of six timepoints: 5, 10, 15, 20, 25, or 30 seconds after the shape appeared. Importantly, participants did not have access to any timekeeping devices to track shock timing. In the *early-threat* environment (indicated in our figures by the pink circle), shock timing was drawn from a statistical distribution such that, at the aforementioned six timepoints, $P(shock)_{EARLY}$=[.35, .13, .13, .13, .13, .13]; in the *late-threat* environment (indicated in our figures by the gray square), shock timing was defined by $P(shock)_{LATE}$=[.13, .13, .13, .13, .13, .35] (Fig. 1D, *top*). Critically, between 5 and 25 seconds, P(shock) values in both the *early-* and *late-threat* environments were held constant, but the hazard rates differed (Fig. 1D), allowing us to dissociate these tractable features of uncertain-threat anticipation.

Post-Training Questionnaire

Our goal in the testing phase was to evaluate the effects of statistical threat-learning on avoidance behaviors and concomitant emotional responses in uncertain-threat environments. To ensure learning occurred during training, before beginning the testing phase we asked our participants to estimate the statistics of the early- and late-threat environments (N=42, counterbalanced), via an online questionnaire. Participants estimated a mean shock-delivery time of 9.46s (SD=4.62s) in the early-threat environment and a mean shock-delivery time of 20.40s (SD=6.99s) in the late-threat environment. Ground-truth means were 14.75s and 20.25s, respectively. Interestingly, the subjects significantly underestimated the average time of shock-delivery in the early-threat environment only (independent-samples *t*-test = 7.15, p<.001), which is aligned with recent findings reporting similar time-contraction effects during induced-anxiety tasks (Sarigiannidis et al., 2020). We then asked our participants (N=33) to estimate the number of shocks that occurred in each of the six epochs for both environments. Of note, participants' estimates were unbounded. The estimates revealed a right-skewed early-threat distribution and a left-skewed late-threat distribution (Fig. 2D) that roughly approximated the mirror-inverse symmetry of our ground-truth P(shock) distribution (Fig. 1D, *top*). By the 16-20s epoch, participants reliably expected there to be fewer shocks administered in the early-threat environment (independent-samples *t*-test = 3.61, p<.001). Collectively, our post-training data indicated that our participants had learned about the statistics of the threat environments.

Testing Phase

Following the questionnaire, participants underwent a testing phase that allowed us to evaluate the impacts of statistical threat-learning on risk-vs-reward tradeoffs in distinct uncertain-threat environments. During the testing phase, volunteers earned $0.01 for every second that a shape was on-screen. At the 5-second mark of each trial, if a participant had not yet been shocked they

gained volitional control over “escape” decisions; that is, the participant could press the spacebar at any time to end the current trial and advance to the next trial, thereby foregoing future gains in that trial in favor of avoiding being shocked. Critically, measuring the timing of escape decisions afforded us an objective measure of anxious behavior with strong ecological footing (Cooper, William, & Blumstein, 2015; Blanchard et al., 2011), making it amenable to a variety of cross-species translational approaches (Evans & Stempel, 2018; Choi & Kim, 2010). Before beginning the testing phase, volunteers were informed that “the shapes would be the same ones that they learned about in the training phase,” and were instructed to “use what they learned about the shapes to decide whether and when to escape.” Participants accrued cash rewards on a second-by-second basis until the end of the trial, regardless of whether it ended with shock administration or escape (Fig. 2E).

Results

Higher Hazard Rates Causes More Avoidance Behavior

In the testing phase, participants’ behavioral responses reflected our computational model of uncertain-threat anticipation and supported the hypothesis that differences in hazard rate drive differences in escape behavior irrespective of P(shock). When P(shock) was held constant, differences in hazard rate caused significant differences in avoidance behavior: Participants were much more likely to forego potential future gains and escape from the early-threat environment, where the hazard-rate is always higher than that of the late-threat environment during the P(shock)-matched epochs (Figs. 1D, 3A; independent-samples *t*-tests, N=42: [6-10s: t=8.09, p<.001], [11-15s: t=7.80, p<.001], [16-20s: t=4.71, p<.001], [21-25s: t=8.30, p<.001]). Critically, participants were more likely to escape during the 16-20s and 21-25s epochs, *even though they reliably reported receiving fewer shocks in the early-threat environment during this period*, underscoring the potential for anxious anticipation to elicit maladaptive avoidance. Moreover, the

escape trends of each environment during the P(shock)-matched epochs revealed a pattern of avoidance behavior that conspicuously resembled each environment's hazard rate (Figs. 1D, 3A). These effects reversed when the hazard rate and P(shock) values converged on 1 during the final 26-30s epoch. Because of this convergence, our model made no predictions about volunteers' behavioral or emotional responses for this epoch. We hesitate to interpret this effect since, in this epoch, hazard rate and P(shock) are confounded.

To further examine the differences in behavioral responses elicited between the early- and late-threat environments, we used Kaplan-Meier survival analysis (KMSA) to compare the time course of escape behavior while accounting for trials that were "censored" by shocks (Fig. 3B). KMSA provides a versatile means of analyzing a range of time-to-event data (e.g., Goel et al., 2010; Lim, 2020). Importantly, because participants could not escape from any trial until after the first 5-second epoch had passed, their escape decisions began at the outset of a 20-second period during which P(shock) values in both early- and late-threat environments were held constant, but during which the early-threat environment's hazard rate was always higher. We hypothesized that participants would behave as though the early-threat environment was more threatening, and be more motivated to avoid the potential threat (i.e., by escaping), *even though the initial spike in momentary threat probability in the early-threat environment had already passed by the time they were permitted to escape*. As predicted, KMSA confirmed that when participants gained volitional control over their escape decisions (see Fig. 2E and *Methods*), their survival rates differed dramatically between environments, with significantly lower survival at all timepoints of the early-threat environment (log-rank test, N=21, χ^2=259.30, p<.005). If P(shock) *per se* drove anxiety and guided escape behavior, the survival curves should not have significantly differed across the four epochs during which those values were matched between environments (Fig. 1D, *top*). These findings were further supported by logistic regression analyses examining the main effect of hazard rate on escape decisions across conditions. These analyses revealed a significant effect

of hazard rate ($z=20.67$, $p<.001$), which survived after controlling for environment ($z=21.02$, $p<.001$) and was significant in each environment alone (z-statistics > 5.03, p values < .001). Together, these findings indicate that participants' avoidance behavior was driven by changes in hazard rate, irrespective of P(shock).

Higher Hazard Rates Elicit Greater State Anxiety

Finally, we sought to test the critical hypothesis that the early-threat environment would be perceived as more anxiogenic, despite both environments being uncertain and all trials ending with a shock (if a participant chose not to escape). In a final, 30-second inescapable trial for each environment (counterbalanced), administered at the end of the testing phase (see *Methods*), participants reported significantly greater anxiety following shock administration in the early-threat environment (Wilcoxon rank-sum test, N=21, $U=328.5$, $p<.005$). And in a two-response, forced-choice question posed at the conclusion of the study—"*Overall, which shape made you more anxious*?"—90.4% of participants identified the shape representing the early-threat environment as more anxiogenic (binomial test, N=21, $p<.001$). Collectively, our findings indicate that the hazard rate is sufficient to drive changes in the emotional experience of anxiety—even when P(shock) is held constant.

Discussion

"Uncertainty" is Amenable to Computational Modeling

Among anxiety researchers, "uncertainty" is accepted as intrinsically more anxiogenic than certainty (Grupe & Nitschke, 2013)—yet previous research has provided no compelling explanation for *why* this is the case. If uncertainty is indeed important in anxiogenesis and anxious pathology, then a concrete understanding of *why* it makes us more anxious is imperative. One approach to this challenge is to build computational models that decompose "uncertainty" into

tractable parts. We have begun that effort by developing and validating a computational model of uncertain-threat anticipation. We used our model to evaluate the differences in anxious behavioral and emotional responses to innocuous visual stimuli that had been paired with aversive electrical shocks in a statistical threat-learning paradigm. By decomposing "uncertainty" into two precise components—that is, P(shock) and hazard rate—and holding the former constant while parametrically manipulating the latter, we uncovered previously undocumented relationships between environmental threat statistics, avoidant behavior, and anxious phenomenology. Fascinatingly, our results indicate that participants (1) track a relatively accurate approximation of the hazard rate as the probability dynamics of a temporally uncertain threat evolve over time, (2) that they use that information to optimize risk-vs-reward tradeoffs in uncertain-threat environments, and (3) that they feel more anxious in uncertain-threat environments characterized by higher hazard rates. Our results can complement and integrate with existing theories of how the brain implements survival- and emotion-relevant tradeoffs across a range of contexts (e.g., Mobbs et al., 2015; Perusini & Fanselow, 2015; LeDoux & Pine, 2015; Cooper et al., 2015; Choi & Kim, 2010). Moreover, our findings can guide computationally sophisticated, model-based research into the neural substrates of survival optimization.

Identifying the Neural Substrates of Hazard Computations and Survival Optimization

Our findings set the stage for studies using functional neuroimaging (e.g., fMRI) to probe the neural substrates that compute, track, and leverage threat environments' hazard rates to optimize survival. Such studies could incorporate hazard-rate regressors in voxelwise mass-univariate analyses to identify regions where variation in blood oxygen level-dependent (BOLD) response covaries with the hazard rate. We suspect that the central extended amygdala (EAc)—which includes the central nucleus of the amygdala (Ce) and the bed nucleus of the stria terminalis (BST), and which is strongly implicated in threat responding (Clauss, 2019; Fox et al., 2015a; Fox et al., 2015b; Hur et al., 2020; Avery, Clauss, & Blackford, 2016; Somerville et al., 2013;

Shackman et al., 2016; Holley & Campos et al., 2022)—may emerge as one substrate where the hazard rate is integrated into a computation for selecting optimal survival- and emotion-relevant responses to a variety of defensive and non-defensive contexts (Holley & Fox, 2022). Although not designed for this purpose, our post-hoc analyses on a previously published dataset derived from a certain-vs-uncertain threat-anticipation task support this hypothesis: We found that a distributed network of brain regions implicated in threat processing and including the BST tracked hazard rate on a moment-to-moment basis, above and beyond the effect of uncertainty as it is frequently modeled (Fig. 4; Hur et al., 2020; unpublished data). Functional neuroimaging studies designed to test the effect of hazard rate could help resolve inconsistencies in the extant literature; namely, the tendency of some studies (e.g., Somerville et al., 2013; Somerville, Whalen, & Kelley, 2010), but not others (e.g., Hur et al., 2020; Fox & Shackman, 2019), to reveal greater sustained activation in BST compared to Ce during uncertain-threat anticipation. Insofar as behavioral and emotional responses are driven by activity in the brain, our findings indicate that accounting for the hazard rate will be vital to the interpretation of this literature, and that it must be considered in future research. Importantly, because the BST is sexually dimorphic (Allen & Gorski, 1990; Chung, De Vries, & Swaab, 2002), and because anxiety disorders are significantly more prevalent in women than in men (McLean et al., 2011), a focused effort to identify sex differences in hazard-rate sensitivity could provide fresh insights into to the relationships between BST functional dimorphism and the risk of psychopathology.

Clinical Implications of Increased Computational Sophistication in Uncertainty Studies

As functional-neuroimaging studies in humans elucidate the relationships between hazard rate and BOLD response, the synergistic efforts of rodent researchers will lead to the discovery of specific mechanisms in the brain that compute, track, and leverage the hazard rate to guide behavior. This would open the door to studies that manipulate specific circuits, cells, and molecules to modulate anxious behavior by effectively changing how the hazard rate is weighted

in computations that select behavior, potentially uncovering new clinical entry points (Holley & Fox, 2022; Moscarello & Penzo, 2022). Critically, we expect that alterations in these mechanisms are likely to account for the avoidant behavior and pervasive worry characteristic of anxiety disorders (Davis et al., 2010; Grupe & Nitschke, 2013). Building a high-acuity understanding of the aberrant computations that drive avoidant behavior in uncertain-threat contexts could be critical to advancing precision psychiatric approaches (Fernandes et al., 2017; Friston, Redish, & Gordon, 2017). This will grow increasingly important as we refine our understanding of how different pathophysiologies might give rise to a common disordered phenotype, thereby ensuring that "one-size-fits-all" treatments will benefit only a fraction of patients (Holley & Fox, 2022). A focused effort to understand the neurocomputational architecture of threat processing, on the other hand, would aid in the development of new diagnostic and treatment approaches that target aberrant, anxiogenic computations at their source. Our work is a step in this direction.

In closing, we have shown that "uncertainty" is unequivocally *not* categorical. Rather, tractable components of uncertainty—here, P(shock) and hazard rate—can be parametrically modulated to drive significant differences in emotional and behavioral response to conditioned threat stimuli. There are likely to be numerous other components of uncertainty that, when identified, will be amenable to manipulations like ours. This should be considered both in study design and literature review. By reimagining uncertainty and exploring its computational architecture and aberrations, we stand to learn not only *why* uncertainty makes us so anxious, but *how* we can restore psychosocial functioning to the countless millions who suffer under anxiety's yoke.

Methods

Participants

A total of 53 adults volunteered for our study. All volunteers were recruited via UC Davis' SONA paid research participation system. Our inclusion criteria were age (18-40), vision (20-20 or corrected to 20-20), and English language fluency. Our exclusion criteria were ongoing psychiatric treatment, history of psychiatric disorder or neurological disorder/injury, ongoing illicit drug use, and pregnancy or possibility of pregnancy. All activities were approved by, and conducted in strict accordance with, the policies of the UC Davis Institutional Review Board under authorization 1716796-1. Because our study featured aversive electrical shocks, in addition to informed consent all participants were repeatedly reminded that they could withdraw at any time. Data from initial participants (5F/4M, mean age = 23.11 years, SD=3.37 years) were used for protocol refinement and are not included in our analyses. Of the 44 participants who volunteered thereafter, our team involuntarily withdrew 2 participants whose shock tolerance exceeded our paradigm's maximum shock level, bringing our total cohort to N=42 (31F/11M, mean age = 21.68 years, SD=3.12 years).

Shock Equipment

To administer shock stimuli, we used the STMEPM-MRI System (Biopac), which consists of a constant-voltage Stimulator Module (STM100C; range: ±10 V), Stimulus Isolation Adapter (STMISOC), Isolated Power Supply (IPS100C), and MRI-compatible Filter/Cable Set (MECMRI-STMISO). The Stimulator Module provides safe, fully programmable, real-time computer control over the electrical stimulus train (i.e., pulse duration, repetition, onset, and amplitude). The STMEPM-MRI System is intrinsically safe—despite possible errors in user stimulation setup or programming—under all operating conditions. Specifically, the strongest possible pulse under open-circuit conditions (160 mJ at 500 ohms, or 200 V) is well below the levels detailed in IEC 60601-2-10:2015 (max allowed: 300 mJ at 500 ohms, or 500 V), the harmonized, international regulatory standard relating to the safety of nerve and muscle stimulators. Stimuli were delivered to the musculature of the hand, between thumb and forefinger, via disposable electrodes. The timing and maximum number of shocks were strictly controlled by our predetermined experimental

schedule. The duration of every shock administered during the study was 0.1s. Volunteers were fully informed of these aspects of shock delivery, without deception, during the informed consent process and again before data collection. During the informed consent process and prior to data collection, our team repeatedly stressed that participants could stop at any time.

Shock Workups

We used a progressive workup method to determine an appropriate shock value for each participant. Participants were instructed to help us identify a shock level that was “uncomfortable but not unbearable” and were reminded that they “should prefer to not receive the shock.” To find this level, we began with a device output of 20 V and worked up as necessary in increments of 20 V toward a maximum of 200 V, administering two shocks at each intensity and obtaining positive consent from the participant at each step before increasing the value. Once we arrived at the appropriate shock level, confirmed by each participant, we entered that value for the training and testing phases of our study. Participants were informed that their shock level could not be changed midway through the experiment, but that they were free to withdraw at any time if the shocks became too intense or, conversely, were insufficient to motivate engagement in the task.

‘Escape from the Shocks’ Paradigm

Our paradigm consists of a training phase and testing phase. The training phase lasts roughly 30 minutes and is intended to teach participants about the temporal threat statistics of two “environments”—one in which shocks tend to occur early (*early*), and another in which the shocks tend to occur late (*late*), as shown in Figures 1B and 1C. Each environment was represented by one of two distinct shapes, which were randomly assigned for each participant (Fig. 2A). During training, participants sat in front of a computer screen and were connected to the shock equipment. Watches, mobile phones, and other distracting and/or timekeeping devices were silenced, collected, and secured. Participants saw 50 presentations each of the *early* and *late* shapes in an order that was predetermined by a randomly-selected trial structures drawn from six available structures (each created using the approach detailed below in *Block Generation*). Each shape appeared on screen following a 1-second intertrial interval (ITI, indicated by a + symbol in our

figures), remained on screen for 5 to 30 seconds, and then disappeared with the administration of a shock, followed by the next ITI (Fig 2C). All stimulus feedback was administered at 5-second intervals to facilitate modeling in future fMRI studies. During training, participants had no volitional control over whether they would be shocked (i.e., they could not “escape”); this was visually represented by a lock icon in the upper-left corner of the screen. Importantly, participants were not given any specific information about the timing distributions of shocks or how those distributions related to the shapes, although they were instructed to pay close attention and told that “depending on the shape, the timing of the shocks might be different” (akin to instructions provided to participants in a reward-focused statistical-learning task conducted by McGuire and Kable, 2015). We measured attention to infer learning by asking a one-back memory question—i.e., “Which shape did you just see?”—after every third shape. Participants were informed that failing to answer at least 75% of these questions correctly would result in withdrawal from the study, although no participants were withdrawn under this criterion. After the training phase, participants estimated the average shock timing of each of the shapes and the number of shocks delivered by each shape during each of the six 5-second bins (Fig 2D). Following a break, the participants completed the testing phase.

During the testing block, participants once again saw 50 presentations of each of the *early* and *late* shapes, with presentation order and shock timing once again predetermined by a second trial structure, randomly drawn from the five available structures that were not chosen for the training phase. An example of a testing-phase trial is shown in Figure 2E. In the testing phase, participants earned $0.01 per second while shapes were on screen. If a trial continued beyond 5 seconds (i.e., if a shock had not yet been delivered), the aforementioned lock icon in the upper left corner of the screen would disappear, and pressing the computer's spacebar would then bypass shock administration and advance the program to the next ITI. This was fully explained to participants, who were told, “If you are not shocked within the first 5 seconds, the lock icon will disappear, and then you can decide whether and when to escape” by pressing the spacebar. Participants were told that they would still earn rewards even if they received a shock on a given trial and were instructed to use what they learned about the shapes during training to strike whatever balance between risks and rewards they found appropriate. Participants were again asked to perform a one-back attention task after every third trial. (Again, all participants scored above our threshold 75%

accuracy.) At the end of the training phase, a cross-section of N=21 participants performed a final, inescapable trial for each of the two shapes. Each shape (counterbalanced) was presented for the maximum time (30s) and terminated with the delivery of an inescapable shock. Participants rated their anxiety from 1 (none) to 5 (extreme) immediately following each shock. These same participants (N=21) also answered a two-option, forced-choice question, “Overall, which shape made you more anxious?” (counterbalanced).

Block Generation

To create blocks of trials that recapitulated the statistical dynamics of our model, we first built a “perfect” exemplar block in which the probabilities for *early* and *late* threat stimuli adhered exactly to our model at all timepoints. As an unbiased approach to adding random noise to our environments, we generated 10,000 block structures consisting of 50 *early* and 50 *late* stimuli with timing randomly generated but bounded by the discrete probabilities of our exemplar’s bins (Fig. 1D, *top*). We then narrowed these 10,000 randomly generated blocks down to 21 candidate blocks that were at least 99% correlated to the statistics of our exemplar. Since each subject would experience two block schedules, we incorporated a process to ensure diversity between the order of stimuli in the blocks. To do this, we repeatedly computed a pairwise correlation matrix for the candidate blocks, beginning with all 21 candidates and systematically eliminating the most highly correlated candidates until 6 “most dissimilar” candidates remained. All simulations and analyses were conducted in Python 3.8.3 using the pandas and NumPy libraries. An example of one of the 6 block schedules used in our study is shown in Figure 2B.

Statistical Analyses

All statistical analyses were performed in Python v3.8.3. Parametric (i.e., independent-samples *t*-tests) and nonparametric (i.e., Wilcoxon rank-sum tests) hypothesis tests were performed using associated modules from the scipy.stats library (i.e., ttest_ind and mannwhitneyu, respectively). Kaplan-Meier survival analyses were performed using Python’s lifelines library (v0.27.1; Davidson-Pilon, 2019), and differences between the curves were evaluated via a log-rank test using lifelines’ statistics.logrank_test function. The following logistic regression analyses were performed using Pythong’s statsmodels library v0.13.2: smf.logic(“escape

~ hazard_vector", df); smf.logit("escape ~ hazard_vector+threat_env", df); smf.logit("escape ~ hazard_vector ", df[df['threat_env']=='late_threat']); and smf.logit("escape ~ hazard_vector ", df[df['threat_env']=='early_threat']).

Threat-Probability Dynamics of Time-to-Event Analyses

Here, we investigate the impact of the evolving probability of an outcome during a period of anticipation. These values—i.e., the *hazard rate*, [h]—can be computed for any point in the anticipation period as:

$$h(t) = \frac{p(t)}{1 - P(t)}$$

Where, h(t) is the hazard rate, p(t) is the probability of the event occurring at this specific moment in time, and P(t) is the probability that the event will occur in a given time interval, i.e. up to time t. This cumulative probability, P(t), is defined as the sum of the perceived probability, p(t) from time 0 to t:

$$P(t) = \sum_{i=0}^{t} p(i)$$

During anticipation of an event where the exact time of the outcome is known, P(t) will equal zero until the delivery of the outcome. During anticipation of an event where the distribution of the negative event is uncertain, P(t) will increase over the interval in which p(t) is greater than 0. Thus, the denominator of h(t) decreases over time, resulting in the same probability [p(t)] being associated with increased hazard rate [h(t)] until the anticipation period ends. The cumulative hazard rate throughout a time interval is:

$$H(t) = \sum_{i=0}^{t} h(i)$$

Notably, the H(t) is unconstrained, meaning that it could increase without being bound by some maximum value. This means that, unlike the actual probability [p(t)] or the cumulative probability [P], the cumulative hazard rate [H] can exceed 100% by the time that an outcome occurs (i.e., t=T). We hypothesized that participants would use the hazard rate, as opposed to the experimenter-defined momentary probability of threat, or *P(shock)*, to guide behavior, and that environments with a higher hazard rate would elicit more anxiety than those with a lower hazard rate.

Codebase and Systems

The codebase for our study's shock-workup, stimulus-presentation, and shock-delivery software can be found here: https://github.com/DanHolley/Escape-the-Shocks-study

We built our study interface using Python 3.8.3 on a 2017 MacBook Pro with a 2.8 GHz Quad-Core Intel i7 CPU and a Radeon Pro 555 2 GB GPU, running MacOS 11.2.3. We relied on a variety of libraries to support device drivers, stimulus presentation, and human-computer interaction. Please refer to the codebase for our requirements, while keeping in mind that requirements may differ across systems.

Acknowledgements

This work was supported by NIH Grants R01MH121735, R21MH129851 to ASF, and the California National Primate Research Center (P51OD011107). DH would like to thank KMM for her insights and support. ASF and DH thank collaborators JH, JFS, AJS and colleagues for their generous assistance with proof-of-principle tests of our model in their MTC fMRI dataset.

Conflict of Interest Statement

The authors report no conflicts of interest.

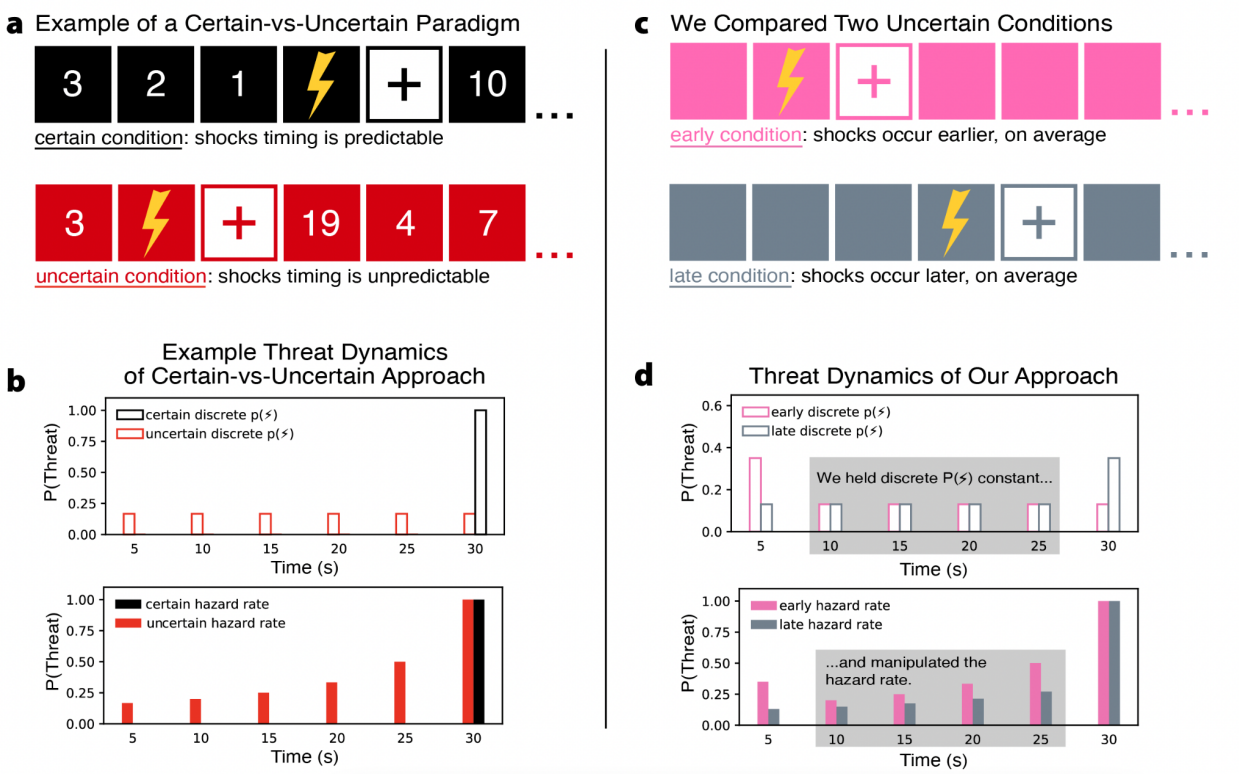

Figure 1. **A computational approach to the study of uncertain-threat anticipation. a)** studies of threat anticipation often compare a certain condition and an uncertain condition. **b)** In certain-vs-uncertain comparisons P(threat) and hazard rate values are confounded, prohibiting causal investigation about the contributions of either to the behavioral expression or emotional experience of anxiety. In the certain-threat condition, note that both P(threat) and hazard-rate values remain at zero until the final epoch, when they peak at 1. **c)** We addressed this confound by comparing two uncertain conditions. **d)** In these conditions, we held P(shock) constant across several time bins while manipulating the hazard rate, thereby allowing us to dissociate the effects of hazard rate from the effects of P(shock) values.

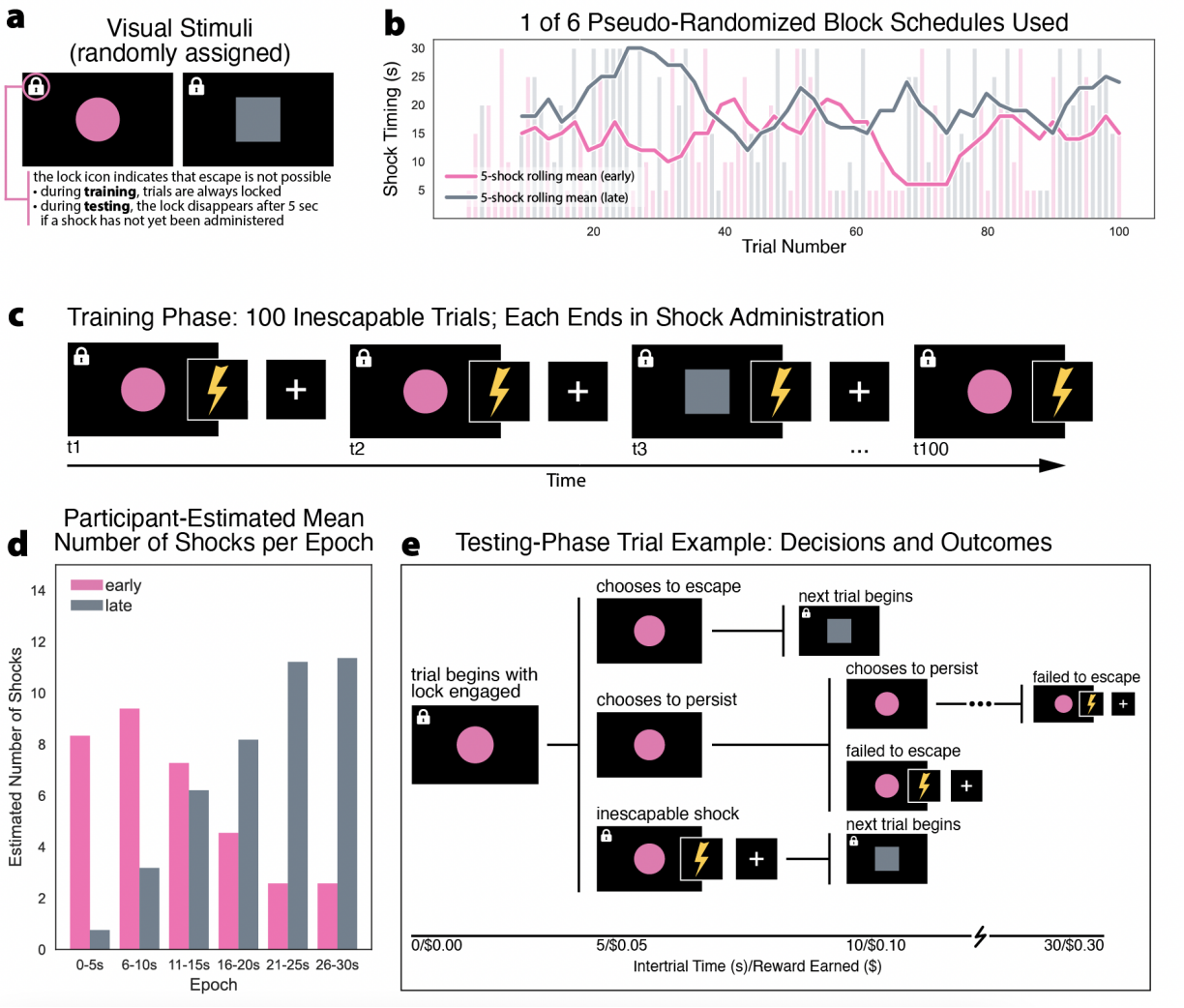

Figure 2. **'Escape from the Shocks' paradigm. a)** We paired randomly assigned, innocuous visual stimuli with shocks during periods of uncertain-threat anticipation. Shock timing was dictated by the statistical dynamics of early- and late-threat environments, shown in Figure 1D. **b)** An example of a block schedule used in our study. At the outset of the study, each participant was assigned two schedules, randomly drawn without replacement from six pseudo-randomized block schedules that recapitulated our early- and late-threat environments' P(shock) values and hazard rates. Bold lines show the 5-trial rolling mean of shock-administration times. **c)** In the training phase, participants are exposed to each shape 50 times in a predetermined, pseudo-randomized order, with all trials ending in shock administration at a pre-set, individually titrated intensity. **d)** Participants' post-training estimates of environmental statistics, such as the mean number of shocks administered per epoch, indicate that learning occurs over the course of the training phase. **e)** An illustrative example of a testing-phase trial, with possible outcomes depicted based on shock timing and the participant's risk-vs-reward decisions.

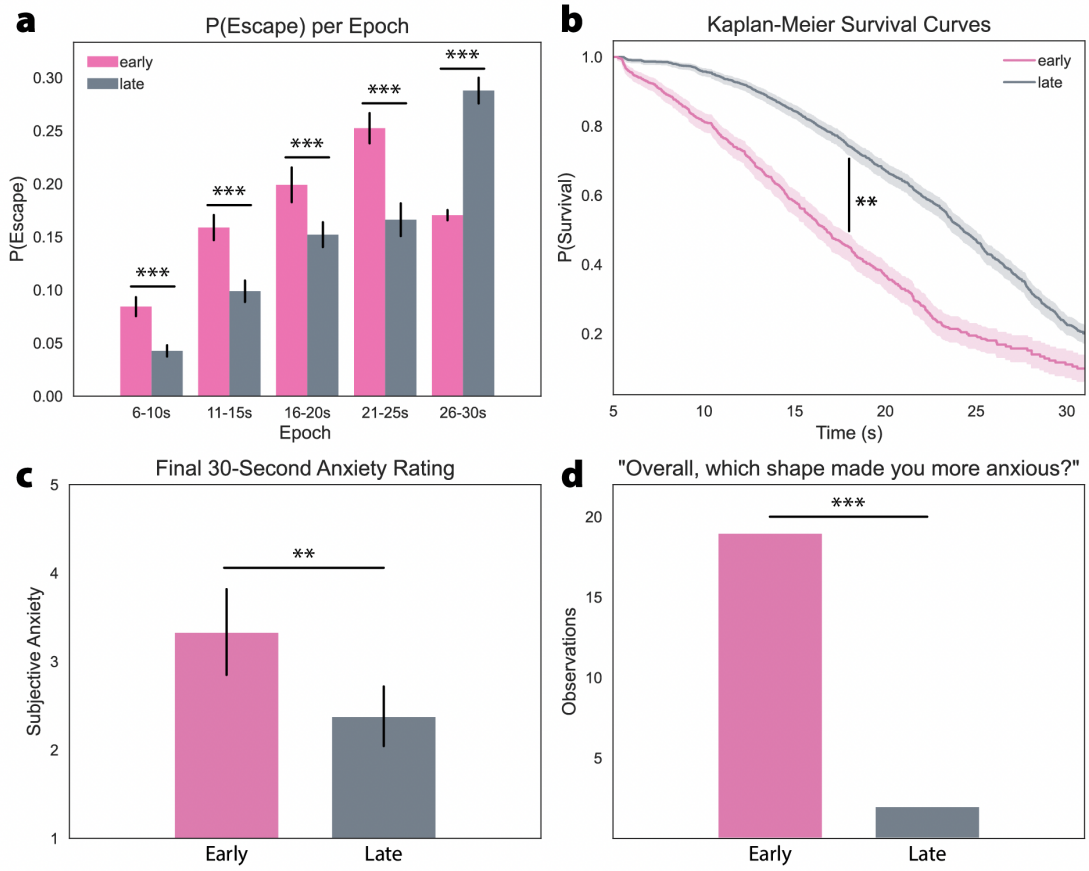

Figure 3. **Higher threat hazard rates cause significantly more anxiety. a)** In epochs during which P(shock) values were matched but hazard rates varied (see Fig. 1D), participants were significantly more likely to escape from the early-threat environment, where the hazard rate was higher at all P(shock)-matched timepoints (independent-samples *t*-tests, N=42: [6-10s: t=8.09, p<.001], [11-15s: t=7.80, p<.001], [16-20s: t=4.71, p<.001], [21-25s: t=8.30, p<.001]). Escapes were locked in 0-5s epoch. Although data is shown our model makes no predictions for escape behavior during the final epoch, during which P(shock) values and hazard rates converge on 1 and are therefore confounded. Although the escape trend reversed and was significant in this epoch (independent-samples *t*-test, t=18.22, p<.001), we refrain from interpreting these effects. **b)** Kaplan-Meier (KM) survival analysis revealed a significant difference between escape activity at all timepoints along the early- and late-threat environments' respective KM curves (log-rank test, N=42, χ^2=259.3, p<.005). **c)** In a final, 30-second, inescapable trial for each environment (counterbalanced), participants reported significantly more anxiety in the early-threat environment (Wilcoxon rank-sum test, N=21, U=328.5, p<.005) **d)** In a two-response, forced-choice test to determine which shape caused more anxiety overall (counterbalanced), participants showed a significant bias toward the shape paired with the early-threat environment (binomial test, N=21, p < .001). Legend: ** = p<.005, *** = p<.001

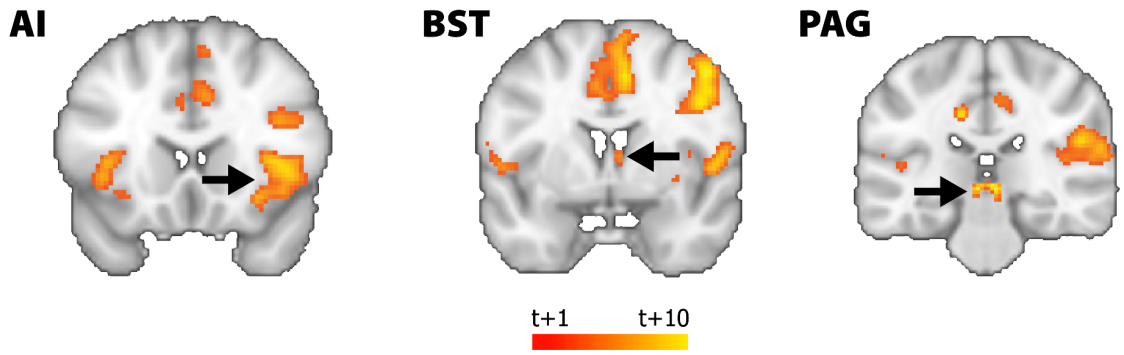

Figure 4. A distributed network of threat-sensitive brain regions tracks the hazard rate on a momentary basis. As a proof-of-principle test of our model in a functional-neuroimaging context, our collaborators at University of Maryland (see *Acknowledgements*) performed voxelwise mass-univariate analyses to test a hazard-rate regressor derived from our model against data from their certain-vs-uncertain threat-anticipation paradigm, the Maryland Threat Countdown (MCT; Hur et al., 2020; N=220, thresholded at t>10). Although the MTC was not designed to test our hypotheses, our regressor identified a main effect of hazard rate, collapsed across conditions, and captured BOLD variation across a distributed network of regions implicated in threat processing (e.g., see Mobbs et al., 2015): The anterior insula (AI; *left*), BST (*middle*), and periaqueductal gray (PAG; *right*) tracked the hazard rate on a moment-to-moment basis.

References

Allen, L. S., & Gorski, R. A. (1990). Sex difference in the bed nucleus of the stria terminalis of the human brain. *The Journal of Comparative Neurology*, *302*(4), 697–706. https://doi.org/10.1002/cne.903020402

Anderson, E. C., Carleton, R. N., Diefenbach, M., & Han, P. (2019). The relationship between uncertainty and affect. *Frontiers in Psychology*, *10*, 2504. https://doi.org/10.3389/fpsyg.2019.02504

Avery, S. N., Clauss, J. A., & Blackford, J. U. (2016). The human BNST: Functional role in anxiety and addiction. *Neuropsychopharmacology*, *41*(1), 126–141. https://doi.org/10.1038/npp.2015.185

Bandelow, B., & Michaelis, S. (2015). Epidemiology of anxiety disorders in the 21st century. *Dialogues in Clinical Neuroscience*, *17*(3), 327–335. https://doi.org/10.31887/DCNS.2015.17.3/bbandelow

Barrett, L. F., Quigley, K. S., & Hamilton, P. (2016). An active inference theory of allostasis and interoception in depression. *Philosophical Transactions of the Royal Society of London — Series B, Biological Sciences*, *371*(1708), 20160011. https://doi.org/10.1098/rstb.2016.0011

Beddington, J., Cooper, C. L., Field, J., Goswami, U., Huppert, F. A., Jenkins, R., Jones, H. S., Kirkwood, T. B., Sahakian, B. J., & Thomas, S. M. (2008). The mental wealth of nations. *Nature*, *455*(7216), 1057–1060. https://doi.org/10.1038/4551057a

Blanchard, D. C., Griebel, G., Pobbe, R., & Blanchard, R. J. (2011). Risk assessment as an evolved threat detection and analysis process. *Neuroscience and Biobehavioral Reviews*, *35*(4), 991–998. https://doi.org/10.1016/j.neubiorev.2010.10.016

Bystritsky A. (2006). Treatment-resistant anxiety disorders. *Molecular Psychiatry*, *11*(9), 805–814. https://doi.org/10.1038/sj.mp.4001852

Choi, J. S., & Kim, J. J. (2010). Amygdala regulates risk of predation in rats foraging in a dynamic fear environment. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(50), 21773–21777. https://doi.org/10.1073/pnas.1010079108

Chung, W. C., De Vries, G. J., & Swaab, D. F. (2002). Sexual differentiation of the bed nucleus of the stria terminalis in humans may extend into adulthood. *The Journal of Neuroscience*, *22*(3), 1027–1033. https://doi.org/10.1523/JNEUROSCI.22-03-01027.2002

Clauss, J. (2019). Extending the neurocircuitry of behavioural inhibition: A role for the bed nucleus of the stria terminalis in risk for anxiety disorders. *General Psychiatry*, *32*(6), e100137. https://doi.org/10.1136/gpsych-2019-100137

Cooper, J., William E., & Blumstein, D. T. (Eds.). (2015). *Escaping from predators: An integrative view of escape decisions*. Cambridge University Press, Cambridge, UK. https://doi.org/10.1017/CBO9781107447189

Davidson-Pilon, C. (2019). Lifelines: Survival analysis in Python. *Journal of Open Source Software, 4(*40), 1317. https://doi.org/10.21105/joss.01317

Davis, M., Walker, D. L., Miles, L., & Grillon, C. (2010). Phasic vs sustained fear in rats and humans: Role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacology, 35*(1), 105–135. https://doi.org/10.1038/npp.2009.109

Evans, D. A., Stempel, A. V., Vale, R., Ruehle, S., Lefler, Y., & Branco, T. (2018). A synaptic threshold mechanism for computing escape decisions. *Nature*, *558*(7711), 590–594. https://doi.org/10.1038/s41586-018-0244-6

Fernandes, B. S., Williams, L. M., Steiner, J., Leboyer, M., Carvalho, A. F., & Berk, M. (2017). The new field of 'precision psychiatry'. *BMC Medicine*, *15*(1), 80. https://doi.org/10.1186/s12916-017-0849-x

Fox, A. S., & Kalin, N. H. (2014). A translational neuroscience approach to understanding the development of social anxiety disorder and its pathophysiology. *The American Journal of Psychiatry*, *171*(11), 1162–1173. https://doi.org/10.1176/appi.ajp.2014.14040449

Fox, A. S., Oler, J. A., Shackman, A. J., Shelton, S. E., Raveendran, M., McKay, D. R., Converse, A. K., Alexander, A., Davidson, R. J., Blangero, J., Rogers, J., & Kalin, N. H. (2015a). Intergenerational neural mediators of early-life anxious temperament. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(29), 9118–9122. https://doi.org/10.1073/pnas.1508593112

Fox, A. S., Oler, J. A., Tromp, D. P. M., Fudge, J. L., & Kalin, N. H. (2015b). Extending the amygdala in theories of threat processing. *Trends in Neurosciences*, *38*(5), 319–329. https://doi.org/10.1016/j.tins.2015.03.002

Fox, A. S., & Shackman, A. J. (2019). The central extended amygdala in fear and anxiety: Closing the gap between mechanistic and neuroimaging research. *Neuroscience Letters, 693*, 58–67. https://doi.org/10.1016/j.neulet.2017.11.056

Friston, K. J., Redish, A. D., & Gordon, J. A. (2017). Computational nosology and precision psychiatry. *Computational Psychiatry*, *1*, 2–23. https://doi.org/10.1162/CPSY_a_00001

Garber, J., & Weersing, V. R. (2010). Comorbidity of anxiety and depression in youth: Implications for treatment and prevention. *Clinical Psychology*, *17*(4), 293–306. https://doi.org/10.1111/j.1468-2850.2010.01221.x

Goel, M. K., Khanna, P., & Kishore, J. (2010). Understanding survival analysis: Kaplan-Meier estimate. *International Journal of Ayurveda Research, 1*(4), 274–278. https://doi.org/10.4103/0974-7788.76794

Gorman, J. M. (1996). Comorbid depression and anxiety spectrum disorders. *Depression and Anxiety, 4*, 160-168.

Griebel, G., & Holmes, A. (2013). 50 years of hurdles and hope in anxiolytic drug discovery. *Nature Reviews Drug Discovery*, *12*(9), 667–687. https://doi.org/10.1038/nrd4075

Grupe, D. W., & Nitschke, J. B. (2013). Uncertainty and anticipation in anxiety: An integrated neurobiological and psychological perspective. *Nature Reviews Neuroscience*, *14*(7), 488–501. https://doi.org/10.1038/nrn3524

Holley, D., Campos, L. J., Zhang, Y., Capitanio, J. P., & Fox, A. S. (2022). Rhesus nervous temperament predicts peri-adolescent central amygdala metabolism and behavioral inhibition measured by a machine-learning approach. *BioRxiv.* preprint doi: https://doi.org/10.1101/2022.07.26.501512

Holley, D., & Fox, A. S. (2022, in press). The central extended amygdala guides survival- and emotion-relevant tradeoffs: Implications for understanding common psychiatric disorders. To appear in *Neuroscience and Behavioral Reviews, special edition,* January 2023. preprint doi: 10.31234/osf.io/z9w3t

Hur, J., Smith, J. F., DeYoung, K. A., Anderson, A. S., Kuang, J., Kim, H. C., Tillman, R. M., Kuhn, M., Fox, A. S., & Shackman, A. J. (2020). Anxiety and the neurobiology of temporally uncertain threat anticipation. *The Journal of Neuroscience*, *40*(41), 7949–7964. https://doi.org/10.1523/JNEUROSCI.0704-20.2020

Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research*, *21*(3), 169–184. https://doi.org/10.1002/mpr.1359

Krajbich, I., Lu, D., Camerer, C., & Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology*, *3*, 193. https://doi.org/10.3389/fpsyg.2012.00193

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(33), 13852–13857. https://doi.org/10.1073/pnas.1101328108

Kuhl P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831–843. https://doi.org/10.1038/nrn1533

LeDoux, J. E., & Pine, D. S. (2016). Using neuroscience to help understand fear and anxiety: A two-system framework. *The American Journal of Psychiatry, 173*(11), 1083–1093. https://doi.org/10.1176/appi.ajp.2016.16030353

Lim., T. (2020). Applying survival analysis for customer retention: A U.S. regional mobile service operator. *2020 Zooming Innovation in Consumer Technologies Conference (ZINC)*, 338-342, doi: 10.1109/ZINC50678.2020.9161811

McGuire, J. T., & Kable, J. W. (2015). Medial prefrontal cortical activity reflects dynamic re-evaluation during voluntary persistence. *Nature Neuroscience, 18*(5), 760–766. https://doi.org/10.1038/nn.3994

McLean, C. P., Asnaani, A., Litz, B. T., & Hofmann, S. G. (2011). Gender differences in anxiety disorders: Prevalence, course of illness, comorbidity and burden of illness. *Journal of Psychiatric Research*, *45*(8), 1027–1035. https://doi.org/10.1016/j.jpsychires.2011.03.006

Meacham, F., & T Bergstrom, C. (2016). Adaptive behavior can produce maladaptive anxiety due to individual differences in experience. *Evolution, Medicine, and Public Health*, *2016*(1), 270–285. https://doi.org/10.1093/emph/eow024

Mobbs, D., Hagan, C. C., Dalgleish, T., Silston, B., & Prévost, C. (2015). The ecology of human fear: Survival optimization and the nervous system. *Frontiers in Neuroscience*, *9*. https://doi.org/10.3389/fnins.2015.00055

Mobbs, D., Marchant, J. L., Hassabis, D., Seymour, B., Tan, G., Gray, M., Petrovic, P., Dolan, R. J., & Frith, C. D. (2009). From threat to fear: The neural organization of defensive fear systems in humans. *The Journal of Neuroscience*, *29*(39), 12236–12243. https://doi.org/10.1523/JNEUROSCI.2378-09.2009

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, *16*(1), 72–80. https://doi.org/10.1016/j.tics.2011.11.018

Moors, A., Ellsworth, P. C., Scherer, K. R., & Frijda, N. H. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review, 5*(2), 119-124. https://doi.org/10.1177/1754073912468165

Moscarello, J. M., & Penzo, M. A. (2022). The central nucleus of the amygdala and the construction of defensive modes across the threat-imminence continuum. *Nature Neuroscience*, *25*(8), 999–1008. https://doi.org/10.1038/s41593-022-01130-5

Nepon, J., Belik, S. L., Bolton, J., & Sareen, J. (2010). The relationship between anxiety disorders and suicide attempts: Findings from the National Epidemiologic Survey on Alcohol and Related Conditions. *Depression and Anxiety*, *27*(9), 791–798. https://doi.org/10.1002/da.20674

Perusini, J. N., & Fanselow, M. S. (2015). Neurobehavioral perspectives on the distinction between fear and anxiety. *Learning & Memory, 22*(9), 417–425. https://doi.org/10.1101/lm.039180.115

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–556. https://doi.org/10.1038/nrn2357

Sareen, J., Cox, B. J., Afifi, T. O., et al. (2005). Anxiety disorders and risk for suicidal ideation and suicide attempts: A population-based longitudinal study of adults. *JAMMA Psychiatry, 62*(11):1249–1257. doi:10.1001/archpsyc.62.11.1249

Sarigiannidis, I., Grillon, C., Ernst, M., Roiser, J. P., & Robinson, O. J. (2020). Anxiety makes time pass quicker while fear has no effect. *Cognition*, *197*, 104116. https://doi.org/10.1016/j.cognition.2019.104116

Shackman, A. J., Tromp, D. P. M., Stockbridge, M. D., Kaplan, C. M., Tillman, R. M., & Fox, A. S. (2016). Dispositional negativity: An integrative psychological and neurobiological perspective. *Psychological Bulletin*, *142*(12), 1275–1314. https://doi.org/10.1037/bul0000073

Somerville, L. H., Wagner, D. D., Wig, G. S., Moran, J. M., Whalen, P. J., & Kelley, W. M. (2013). Interactions between transient and sustained neural signals support the generation and regulation of anxious emotion. *Cerebral Cortex*, *23*(1), 49–60. https://doi.org/10.1093/cercor/bhr373

Somerville, L. H., Whalen, P. J., & Kelley, W. M. (2010). Human bed nucleus of the stria terminalis indexes hypervigilant threat monitoring. *Biological Psychiatry, 68*, 416-424.

Swendsen, J., Conway, K. P., Degenhardt, L., Glantz, M., Jin, R., Merikangas, K. R., Sampson, N., & Kessler, R. C. (2010). Mental disorders as risk factors for substance use, abuse and dependence: Results from the 10-year follow-up of the National Comorbidity Survey. *Addiction, 105*(6), 1117–1128. https://doi.org/10.1111/j.1360-0443.2010.02902.x

World Health Organization. (2017). *Depression and other common mental health disorders.* World Health Organization Document Production Services, Geneva, Switzerland.

Wu, Z., & Fang, Y. (2014). Comorbidity of depressive and anxiety disorders: Challenges in diagnosis and assessment. *Shanghai Archives of Psychiatry*, *26*(4), 227–231. https://doi.org/10.3969/j.issn.1002-0829.2014.04.006