Psychopharmacology (2019) 236:99–110
https://doi.org/10.1007/s00213-018-5005-6

ORIGINAL INVESTIGATION

Predicting extinction phenotype to optimize fear reduction

M. H. Monfils[1,2,3] · H. J. Lee[1,2] · N. E. Keller[2] · R. F. Roquet[1] · S. Quevedo[1] · L. Agee[1] · R. Cofresi[2] · J. Shumake[1,3]

Received: 6 April 2018 / Accepted: 16 August 2018 / Published online: 14 September 2018

Abstract

Fear conditioning is widely employed to study dysregulations of the fear system. The repeated presentation of a conditioned stimulus in the absence of a reinforcer leads to a decrease in fear responding—a phenomenon known as extinction. From a translational perspective, identifying whether an individual might respond well to extinction prior to intervention could prove important to treatment outcomes. Here, we test the hypothesis that CO_2 reactivity predicts extinction phenotype in rats, and that variability in CO_2 reactivity as well as extinction long-term memory (LTM) significantly predicts orexin activity in the lateral hypothalamus (LH). Our results validate a rat model of CO_2 reactivity and show that subcomponents of behavioral reactivity following acute CO_2 exposure explain a significant portion of the variance in extinction LTM. Furthermore, we show evidence that variability in CO_2 reactivity is also significantly predictive of orexin activity in the LH, and that orexin activity, in turn, significantly accounts for LTM variance. Our findings open the possibility that we may be able to use CO_2 reactivity as a screening tool to determine if individuals are good candidates for an extinction/exposure-based approach.

Keywords Extinction phenotype · Fear conditioning · Fear extinction · CO_2 · Individual differences

In fear conditioning, an initially neutral conditional stimulus comes to elicit fear expression after its pairing to an unconditional, aversive stimulus. The subsequent repeated presentation of the conditioned stimulus in the absence of a reinforcer leads to a progressive decrease in fear responding—a phenomenon and paradigm known as extinction. Exposure therapy, a therapeutic approach employed in clinical settings, shares characteristics and mechanisms with extinction, and mounting evidence suggests that there are considerable differences in individual responding to both exposure therapy and extinction (Shumake et al. 2014, 2018; Bush et al. 2007; Galatzer-Levy et al. 2013; Schwartze et al. 2017). From a translational perspective, identifying whether an individual might respond well to extinction prior to intervention could prove important to treatment outcomes. Effectively, if we could determine, prior to treatment, that an individual is not a good candidate for extinction/exposure therapy, they could be reasonably assigned to another treatment strategy.

In a recent study, Sharko et al. (2017) found that differences in orexin activity in the hypothalamus significantly account for individual differences in extinction phenotype in rats. Individual differences in extinction as well as CO_2 exposure have been respectively found to activate orexin neurons in the lateral hypothalamus (Johnson et al. 2011). Orexin from the lateral hypothalamus (LH) modulates amygdala threat (fear) learning (Sears et al. 2013), and orexin receptor antagonism has been found to facilitate extinction from context and cued fear conditioning (Flores et al. 2014). Furthermore, antagonism of orexin receptors increases the recruitment of basolateral amygdala (BLA) neurons that project to the infralimbic cortex during extinction (Flores et al. 2017). Those very same neurons (the IL projecting BLA neurons) are the ones found to be active during extinction (Senn et al. 2014), supporting the notion that individual differences in orexin activation in the hypothalamus could account for individual differences in extinction (Sharko et al. 2017). In humans, adults with anxiety disorders display heightened emotional reactivity to a single inhalation of 35% CO_2;

This article belongs to a Special Issue on Psychopharmacology of Extinction.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00213-018-5005-6) contains supplementary material, which is available to authorized users.

✉ M. H. Monfils
marie.monfils@gmail.com

[1] Department of Psychology, University of Texas at Austin, Austin, TX, USA

[2] Institute for Neuroscience, Austin, TX, USA

[3] Institute for Mental Health Research, Austin, TX, USA

however, data investigating prospective linkages between emotional reactivity to CO_2 and susceptibility are limited (Telch et al. 2012). We propose that CO_2-reactivity might prove an important tool to identify extinction phenotype. Consistent with this idea, CO_2 reactivity predicts the later development of PTSD symptom severity (Telch et al. 2012), and individuals with PTSD show deficits in extinction (Pitman et al. 2012), and dysregulation of HPA axis (Yehuda 2009; Michopoulos et al. 2017).

Here, we test the hypothesis that CO_2 reactivity predicts extinction phenotype in rats, that variability in CO_2 reactivity significantly predicts orexin activity in the LH, and that, in turn, orexin activity in the LH predicts extinction long-term memory (LTM).

Methods

Procedures were conducted in compliance with the National Institutes of Health Guide for the Care and Use of Experimental Animals and were approved by the Institutional Animal Care and Use Committee of the University of Texas at Austin.

Husbandry

Throughout all experimental procedures, subjects were housed in pairs in temperature and humidity-controlled transparent polyethylene cages and were maintained on a 12-h/12-h light/dark cycle with food and water available ad libitum. Subjects consisted of 56 male Sprague-Dawley rats (approximately 80 days of age), obtained from Envigo (Houston, TX, USA).

Apparatus

All experimental manipulations (fear conditioning, extinction, long-term memory) were administered in the same context (operant conditioning chambers; Coulbourn Instruments, Whitehall, PA). Each chamber was equipped with a stainless-steel rod flooring connected to a shock generator (Model H10-11R-TC-SF; Coulbourn Instruments, Whitehall, PA) and individually enclosed in a sound-insulated box (Isolation Cubicle, Model H10-24T; Coulbourn Instruments, Whitehall, PA). Chambers were illuminated with a red light. Behavior was recorded by infrared digital cameras (Panasonic, model wvBP334, Osaka, Japan) mounted on the ceiling of each unit. Stimulus presentation was automated using FreezeFrame2 software (Coulbourn Instruments, Whitehall, PA). Equipment was cleaned with Windex (SC Johnson, Racine, WI) between each session.

Experimental timeline

Rats were first screened for reactivity to CO_2 ($n = 34$) or Normoxic Air ($n = 22$). Then, at least 5 days later, they were fear conditioned using 3 tone shock pairings. The next day, they received an extinction session (19 CSs). The day after extinction, they received a long-term memory test (LTM). At least 4 days after the LTM test, a subset of rats were tested in the elevated plus maze, and another subset were tested in the light-dark box. At least 6 days later, all rats received a CO_2 challenge and were sacrificed 1 h later for later immunohistochemistry processing (see Fig. 1).

CO_2 screening

Flow cages (12″ width × 12″ height × 24″ length) were custom built using plexi-glass. Gas flow was regulated using a two-stage regulator (Praxair, Inc., Danbury, CT, USA). An infusion hose was placed to allow air to enter the chamber. Infusion of hypercarbic gas blended with normoxic air (25% CO_2) began 0.5 min after placement of the rat in the chamber and continued for 2 min (induction phase; shown in Fig. 2a). After 2 min, the gas flow was held constant at 25% (25% CO_2 Hold phase). After 2 additional minutes, the gas flow was terminated and the cage was flushed to allow rapid equilibration with atmospheric air (Flush-out phase). The rat was left in the chamber for an additional 4 min and then was transferred to its original home cage. Tests of CO_2 levels across the different stages of CO_2 flow are shown in Fig. 2. Additionally, as a control condition, a subset of rats were exposed to only normoxic air in the plexi-glass chambers for a similar 9-min duration.

CO_2 behavioral analysis

We developed a detailed scoring system to describe and quantify the behaviors observed in response to a CO_2 challenge in rats. Each behavior was quantified at baseline (30 s), during CO_2 induction (2 min), during the CO_2 hold period (2 min), and during the flush-out period (4 min). For the analyses, the flush-out period is divided into 2, so that all the periods of interest are quantified for the same duration. Each session was videotaped from 2 different angles. Behaviors from each angle were scored and averaged for offline analyses. The following behaviors were quantified: Ambulation (A) (time spent moving around. Any displacement of paws). Grooming (G) (time spent grooming), Labored breathing (L) (deep and long breaths, usually noticeable from movements of the torso). Rearing (R) (number of times rat stands on rear legs). For coding purposes, induction was referred to as phase 1, 25% hold as phase 2, and the first and second half of the flush-out period as phases 3 and 4. As such, Ambulation recorded during the first half of the flush-out phase will be reported as "A3".

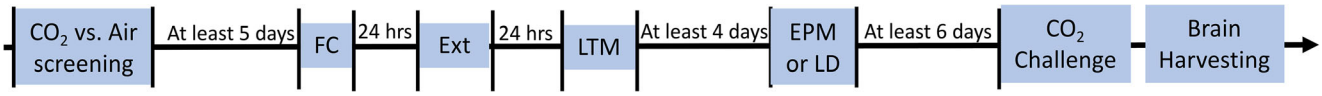

Fig. 1 Experimental timeline. Rats were first screened for reactivity to CO_2 ($n = 34$) or Normoxic Air ($n = 22$). Then, at least 5 days later, they were fear conditioned using 3 tone shock pairings. The next day, they received an extinction session (19 CSs). The day after extinction, they received a long-term memory test (LTM). At least 4 days after the LTM test, a subset of rats were tested in the elevated plus maze, and another subset were tested in the light-dark box. At least 6 days later, all rats received a CO_2 challenge and were sacrificed 1 h later for later immunohistochemistry processing

Fear conditioning

Subjects were placed in the conditioning chambers, allowed to habituate for 3 min, and then fear conditioned with 3, 20-s, 5 kHz, 80 dB tones (CS). Each co-terminated with a 0.5-s, 0.7 mA footshock (US). The interval between each CS was 120 s. After conditioning, each subject remained in the chamber for 3 min and was then returned to its home cage.

Extinction

The day after conditioning, rats were returned to the conditioning chambers and allowed to acclimate for 3 min. Subjects received 19 unreinforced CSs. The interval between CSs was variable, with a mean of 180 s. Upon completion of extinction, subjects remained in the conditioning chambers for 3 min.

Long-term memory test

The day after extinction, the rats were brought back to the conditioning chambers, allowed to acclimate for 3 min, then received 4 unreinforced CS presentations.

Behavioral scoring: Freezing

Freezing was defined as the absence of all movement aside from breathing and ear twitching, not including sleeping or resting. Behavior was scored manually from videos by an experimenter blind to experimental conditions. The total amount of CS-induced freezing was expressed as a percentage of total time spent freezing during each 20-s CS. We operationalized post-acquisition conditioned fear as mean freezing over the first 2 tone-alone trials 24 h after acquisition (i.e., the first 2 trials of extinction), end of extinction conditioned fear as mean freezing over the last 2 trials of extinction, and extinction long-term memory as freezing over the first 2 trials of the LTM test, to minimize the likelihood of re-extinction, and provide a more accurate estimate of long-term memory post extinction.

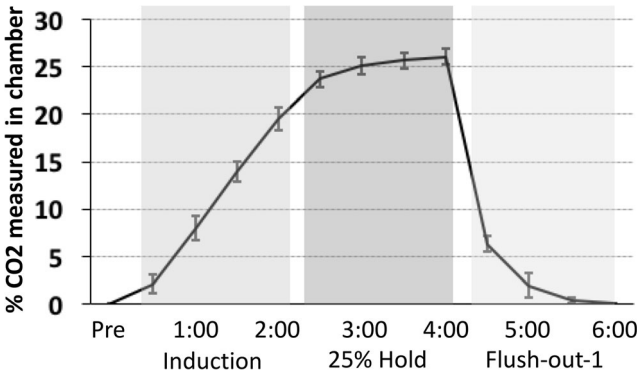

Fig. 2 Measurements of CO_2 concentration in chamber during "induction", "hold", and "flush-out" phases. Mean (± SEM) percent CO_2 concentrations in the CO_2 chamber during the induction, 25% hold, and flush-out periods, as measured by a device located directly inside the chamber. Note that after 2 min and 30 s of CO_2 infusion, the levels inside the chamber reach the desired concentration of 25%

Elevated plus maze

Rats were individually placed in the middle of the elevated plus maze (a maze that possesses 2 open arms, and 2 closed arms). They were allowed to move freely throughout the maze for a 5-min period. A camera mounted on the ceiling recorded the rats' behavior for later offline analysis. We analyzed time (in seconds) and percent time spent in the open arms and percent time spent in the closed arms.

Light-dark test

Rats were tested for 2 days in exploratory chambers that automatically measured horizontal (ambulatory) and vertical (rearing) activity under a light–dark condition. This test allowed us to measure activity in response to a relatively threatening environment (the light) versus a relatively safe environment (the dark box). We quantified the traditional metrics (latency to exit dark box and time spent in light box) of the light–dark test (Crawley and Goodwin 1980) in addition to the total activity metrics, which consisted of absolute horizontal and vertical activity (ambulatory distance and rearing counts, respectively) and time-normalized versions of the same metrics (velocity and rearing duration, respectively). Velocity reflects the vigor of movement (distance in cm covered per second of time), and rearing duration reflects the mean length of a single rear (seconds spent rearing per rearing count).

CO2 challenge and brain harvesting

At the end of the experiment, all rats received a CO_2 challenge (as described above under CO_2 screening). One hour after this CO_2 challenge, rats received a lethal dose of sodium

pentobarbital and were intra-cardially perfused with phosphate buffered saline followed by 4% buffered paraformaldehyde. Next, the brains were extracted, stored in paraformaldehyde overnight, and then stored in cryoprotectant (30% phosphate-buffered sucrose) until sectioning.

Orexin-cFos immunohistochemistry

Methodological details can be found in our previous publications (Lee et al. 2005, 2010, Jones et al. 2013; Jones and Monfils 2016; Cunha et al. 2010). Briefly, for each rat, the entire brain was sectioned at a thickness of 30 μm and 4 sets of serial floating sections stored in phosphate buffered saline at − 4 °C. Tissue to be analyzed was rinsed, blocked, and incubated in rabbit anti-cFos antibody (1:1000 dilution; Immunostar) for 72 h at 4 degrees, and mouse anti-orexin antibody (R&D Systems) for 24 h. Tissue was then incubated overnight at − 4 °C with direct conjugate fluorescent secondary antibodies (Anti-mouse 568 Alexa Fluor; Anti-rabbit 488 Alexa Fluor, Abcam), the sections mounted on slides, and coverslipped with fluorescence mounting medium (Prolong® Gold Antifade Mountant, Invitrogen).

Orexin-cFos imaging and quantification

Cells were imaged and counted according to our previously described procedures (Lee et al. 2010; Jones et al. 2013). Briefly, sections from one series were visualized on a fluorescence microscope (Zeiss) with an objective magnification of × 20 through an eyepiece with a magnification of 12.5X. The lateral hypothalamus region identified (defined according to the Paxinos & Watson brain atlas, and observed between − 2.8 and − 3.3 mm from Bregma), and a minimum of 3 sections from similar anterior-posterior planes per brain were photographed bilaterally from an experimenter blind to behavioral outcome. Imaging of each excited fluorofore was performed separately, and care was taken to ensure that there was no bleed through between emissions at the different excitation wavelengths. Cells from the lateral hypothalamus that expressed cFos, orexin A, or both were quantified offline by a different observer who was also blind to behavioral outcome. Since the brains were cut in 4 series, the sections from which cells were counted were not adjacent from one another (the sections that we sampled from were approximately 120 μm apart). All the cells from the sampled region that expressed cFos and/or orexin were counted using ImageJ software (NIH) for Mac. Care was taken to match sections for each brain and the regions were sampled from both the left and right hemispheres.

Statistical analyses

All statistical analyses were performed using R (R Core Team 2016) using the following packages: beset (Shumake 2018) and car (Fox and Weisberg 2011). Fear acquisition, extinction, and long-term memory were compared between CO2-exposed and Air-exposed rats using separate repeated measures ANOVA for each behavior.

To determine which combinations of CO_2-reactivity behaviors accounted for the most important portion of the variance in the extinction long-term memory freezing, we used a modified version of the "best subset" approach to linear regression. Linear regression models are generally easy to fit and interpret, and can provide a good degree of accuracy, since they take into account the whole range of values, and information is thus not lost (which can often occur when variables are categorized into subgroups). It is often the case, however, that not all the variables included into a linear regression model contribute their fair share to the variance explained. In such cases, it is best to weigh the contribution of different components vs. their cost to the analysis (in other words, make sure the cost/benefit analysis makes them worth being included, since including irrelevant variables leads to unnecessary complexity). The best subset approach fits a different linear model for every possible combination of predictor variables. We then used resampling (k-fold cross-validation where k = 10) to estimate how well each model would predict new samples in terms of mean squared error. Each model was repeatedly refit to random subsamples of the data and tested for how well it could predict the remainder of the data that it was not fit to. This results in a sampling distribution of prediction errors. The "best" model was then chosen as the simplest model (the one with the fewest predictors) that was within one standard error of the model that was, on average, best at predicting new data. A detailed explanation of how the beset package implements the best-subsets approach is freely available (Shumake 2018). In addition, since cross-validation estimates of prediction error tend to be overly optimistic if they are used to select the best model (a concept referred to as selection bias), we also ran a nested cross-validation (Cawley and Talbot 2010; Taylor and Tibshirani 2015). This procedure nests the cross-validation used to select the best model within a cross-validation used to estimate test error (prediction error on a new sample), such that test error is evaluated on holdout examples neither used to fit the models nor to select the best models, thus providing a fairer estimate of the model's generalizability.

Results

We analyzed the following behaviors observed in response to a CO_2 challenge in rats: Ambulation (A), Grooming (G), Labored Breathing (L), and Rearing (R) across 4 different phases: CO_2 induction (phase 1), CO_2 hold (phase 2), and flush-out period (phases 3 and 4). As such, Ambulation recorded during the first half of the Flush-out phase, for

example, is reported as "A3". Three R^2 statistics will be reported for each best model: the full-sample R^2 (the unadjusted R^2 obtained by fitting the model to the full data set), the CV-selection R^2 (the R^2 obtained by predicting random holdout examples and used to select the best model), and the CV-test R^2 (the R^2 obtained by predicting holdout examples that were not used for model fitting or model selection).

CO2 reactivity is a significant predictor of post-extinction long-term memory

We operationally defined post-extinction LTM as the mean freezing of the first 2 trials of LTM. For our initial data analysis on post-extinction long-term memory (LTM), we used the best-subset approach to estimate our best model that is within 1 standard error, with the lowest cross-validation error, and with the fewest number of predictors. With those parameters, the model suggests that the best predictive effect of CO_2 reactivity is from predictor A3. A3 alone yields an full-sample R^2 of 0.19, with a CV-selection R^2 of .08, and a CV-test R^2 of − .19. Thus, the CO_2 reactivity predictor A3 accounts for 19% of the variance in post extinction LTM in the observed sample, and is expected to account for between − 19 and 8% of the variance in post extinction LTM in new samples (cross-validation analysis). We also analyzed how sensitive the selection outcome was to the randomized assignments of observations to cross-validation folds and found that the A3 model was selected 52 times out of 100. If we allow the model to run unconstrained (that is, if we remove the 1 standard error rule, and allow more predictors), the best model generated includes R3, G3, and L2, and the full sample yields an R^2 of 0.4, with a CV-selection R^2 of 0.27, and a CV-test R^2 of − .173. If we query the best subset analysis for the best generated model for 2 predictors, the best 2-predictor solution was A3 + G3 (full-sample $R^2 = 0.31$, CV-selection $R^2 = 0.18$, CV-test $R^2 =$ − .267). We also examined whether the best-subset analysis could generate the best model for the end of extinction, and found that it did not (full sample $R^2 = 0$). When testing the best subset analysis on the rats that received normoxic air rather than CO2, the full sample R^2 was 0. When specifically testing our best predictor (A3) for the CO_2 sample on the rats that were exposed to normoxic air, the full sample R^2 was .01, suggesting that CO_2 reactivity, and not simply baseline ambulation, specifically explained variance in post-extinction LTM. We separately ran our analysis on the mean of all 4 LTM trials, since it could provide a more stable estimate of post-extinction LTM (albeit at the risk of re-engaging extinction mechanisms) (see Supplementary Materials).

In looking at the R^2 and the standard error of the estimate for the training sample and the cross-validation as a function of the number of predictors, it is apparent that while adding more predictors explains a larger portion of the variance in the training sample, it yields diminishing returns for the cross-validation (Fig. 3 and 4). In fact, not only are there diminishing returns, more than 4 predictors was actively detrimental to finding a reproducible solution. Figure 3 also

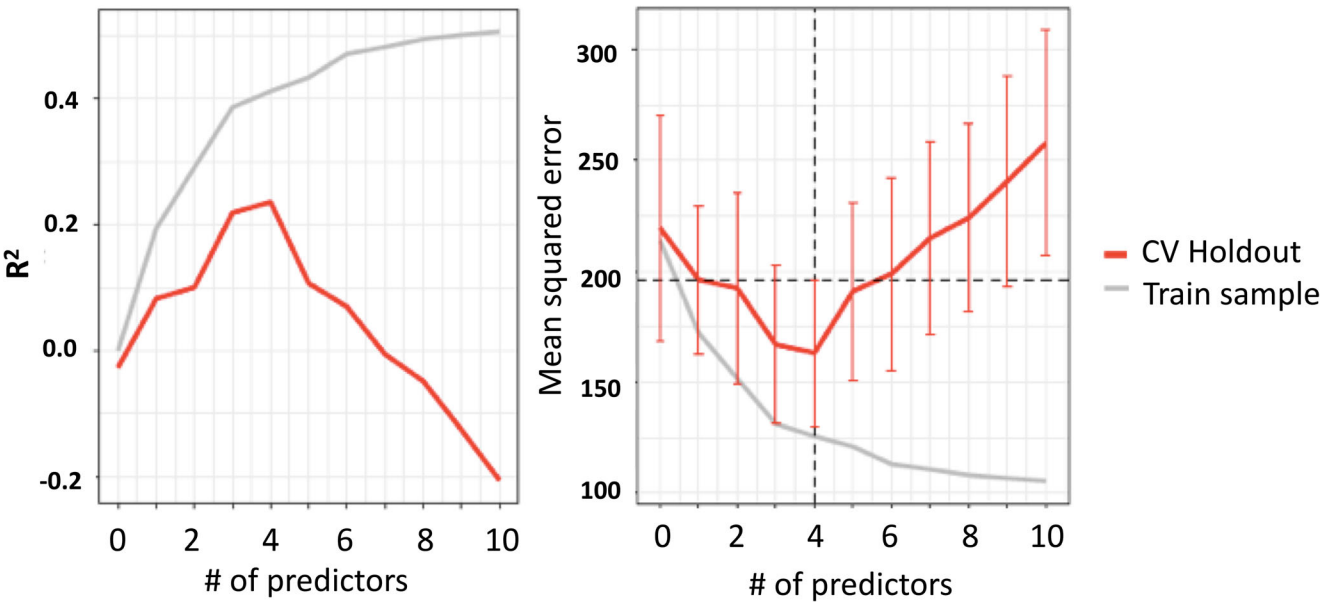

Fig. 3 R^2 and mean squared error for the training sample and the cross-validation as a function of the number of predictors. It is apparent that while adding more predictors explains a larger portion of the variance in the training sample, it yields diminishing returns for the cross-validation. In fact, not only are there diminishing returns, more than 4 predictors is actively detrimental to finding a reproducible solution. Our data also suggest that there is severe overfitting when too many predictors are included—the predictive R^2 is negative, which indicates that the model is predicting noise, and that the predictions are actually worse than simply predicting the mean for everyone. Similarly, as the mean square error continually decreases for the training sample, the error decreases for up to 4 predictors in the CV Holdout, and then increases

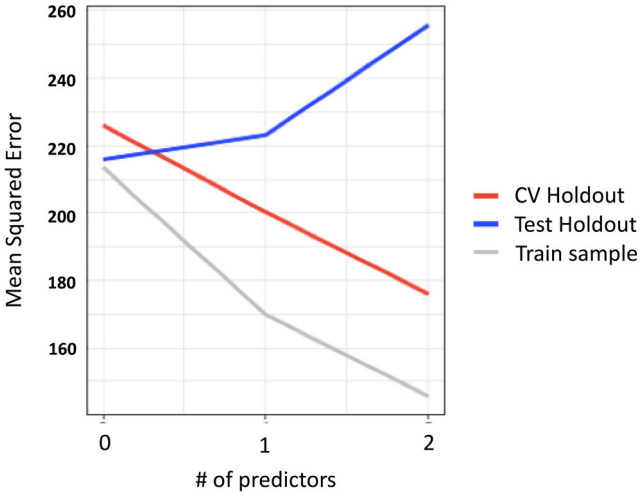

Fig. 4 Mean squared error for the training sample and the inner and outer cross-validation as a function of the number of predictors. It is apparent that as the number of predictors increase from 1 to 2, the standard error of estimate decreases for our training sample and the inner cross-validation (CV Holdout); however, under the same circumstances, the standard error of estimate increases for the outer-cross validation (Test Holdout)

conveys that there was severe overfitting when too many predictors were included—the predictive R^2 becomes negative, which indicates that the model is predicting noise, and that the predictions are actually worse than simply predicting the mean for everyone. A negative predictive R^2 was also observed for even the single-predictor model for the test cross-validation, which provides a meta-estimate of how well models selected using this procedure on this sample can predict new observations. This statistic principally reflects that the current sample size may be too small to both fit and select a model that will reliably generalize to future samples.

Two of the predictors' (R3 and A3; Rearing and ambulation during the flush-Out phase of CO2 challenge) relationship with LTM are shown in Fig. 5, with R3 individually being the weakest and A3 the strongest of the 4 best predictors. The differential predictive effect of R3 and A3 may reflect an adaptive behavioral response: increased movement and exploratory movement might maximize the opportunity to flee when a threat is no longer present see Fig. 6. That the same 4 behaviors (A3, R3, G3, and L2) were found in various combinations across the best predictive scenarios suggests that some might be tapping into similar predictive concepts, and as such, might also be highly exchangeable. To examine the possibility that some of these behaviors might be conceptually related, we ran a principal component analysis (PCA) on all the predictor variables. Our results revealed that the first 2 principal components accounted for ~ 50% of the overall variance, and the first four principal components accounted for nearly 80% of the variance. A correlation matrix for A3, R3, G3, and L2 with the first 4 principal components is found in Table 1. In Fig. 7, we show the relationship between the vectors for all the predictor variables, which conveys that A3 was predominantly associated with principal component 2, G3 was predominantly associated with principle component 1, and R3 and L2 were associated with a combination of the two.

Figure 8 shows the effect in good and poor extinguishers (lower and upper 40% of the LTM freezing distribution). The predictive potential of R3 and A3 seems to be highly specific to extinction, as evidenced by the data shown in Fig. 8. A between subjects t test revealed that there was a significant difference between the lower and upper 40% LTM freezers for the post-extinction LTM ($p < 0.05$), but not during fear acquisition ($p > 0.2$). Furthermore, there was a significant difference between the lower and upper groups on their ambulation during the flush-out period (A3), $p = 0.019$, and their rearing behavior during that same time period approached significance, $p = 0.068$. Additional correlations between A3 and preCS freezing for the conditioning, extinction, and post extinction LTM sessions are shown in the Supplementary Materials. In teasing apart elements of CO_2 reactivity that account for extinction phenotype, it is important to remember that the

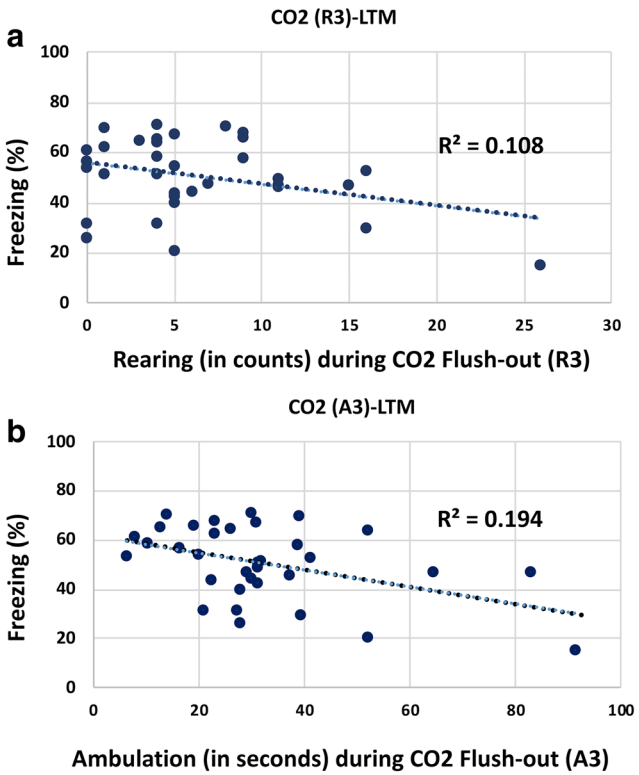

Fig. 5 Relationship of predictors R3 and A3 with extinction long-term memory (LTM). For our initial data analysis on extinction long-term memory (LTM), we used the best-subset approach to estimate our best model that is within 1 standard error, with the lowest cross-validation error, and with the smallest number of predictors. With those parameters, the model suggests that the best predictive effect of CO_2 reactivity is from predictor A3 ($R^2 = 0.19$, cross-validation (CV) = 0.08). If we query the best subset analysis for the best generated model for 2, or 3 predictors, we unveil a best 2-predictor solution of A3 + G3 ($R^2 = 0.31$, CV $R^2 = 0.18$), and a best 3-predictor solution of R3, G3, and L2 ($R^2 = 0.4$, CV $R^2 = 0.27$). Predictors R3 and A3's relationship with LTM are shown in A and B, respectively

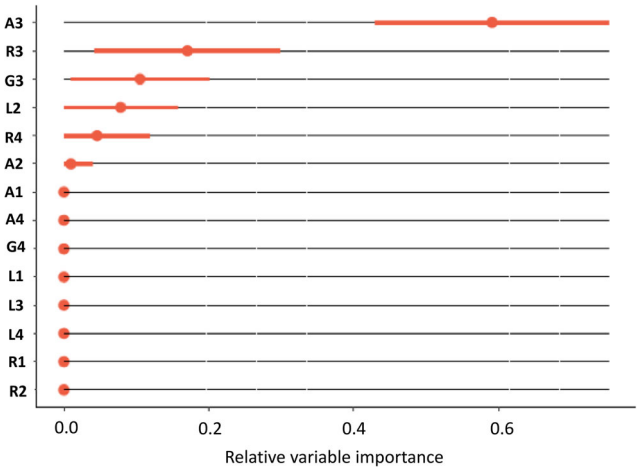

Fig. 6 Predictor importance. The nested cross-validation in the best subset analysis informs us that the most frequently selected best model is a single predictor model, A3, which is chosen 52% of the time. Our nested CV analysis also ranks our predictors in order of importance, and reveals the same predictors as those highlighted by the PCA (A3, R3, G3, and L2). Relative importance scores reflect the mean and standard error of the absolute value of the standardized regression coefficients from 100 different models, each fit to a different random subsample consisting of, on average, 27.5 out of the original 34 cases

current steps are exploratory, and that this study is a first step in determining which elements of CO_2 reactivity might reveal to be most important.

CO2 reactivity predicts cFos+Orexin co-labeling in the lateral hypothalamus

A subset of rats from the CO_2 reactivity experiment were processed and quantified for cFos and orexin immunohistochemistry ($n = 21$) (see Fig. 9). We found that there was a significant predictive relationship of CO_2 reactivity on cFos+Orexin co-labeled cells in the lateral hypothalamus. Our previous analysis identified A3 as the single best predictor of extinction LTM. Here, A3 gives us a significant prediction of cFos+Orexin, with an R^2 of 0.25, $F(1, 19) = 6.452$, $p = 0.01$.

Table 1 Correlations between A3, R3, G3, and L2 and the 4 first principal components. Our results revealed that the first 2 principal components accounted for ~ 50% of the overall variance, and the first four principal components accounted for nearly 80% of the variance. A3 was predominantly associated with principal components 2 and 3, G3 was predominantly associated with principle components 1 and 4, and R3 and L2 were associated with a combination of principal components 1 and 2

	PC 1	PC 2	PC 3	PC 4
A3	0.025	− 0.369	− 0.457	0.179
R3	0.267	− 0.354	− 0.237	− 0.031
G3	0.232	− 0.065	− 0.021	− 0.567
L2	− 0.380	− 0.288	0.109	0.027

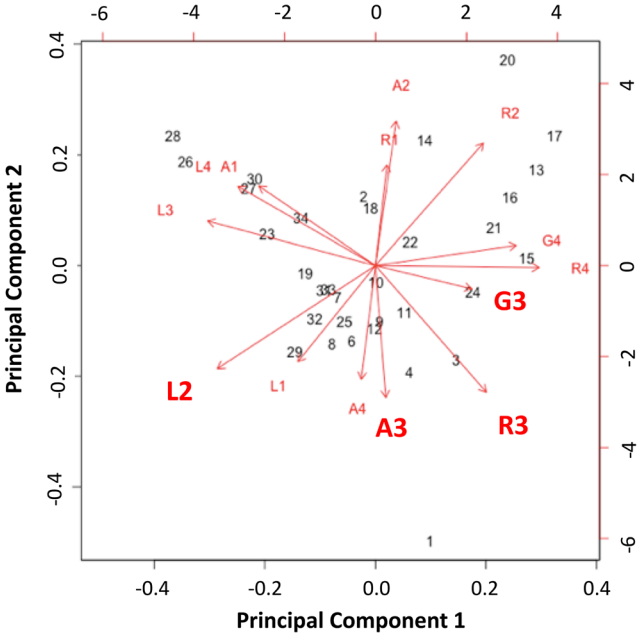

Fig. 7 Principal component analysis vectors for A3, R3, G3, and L2. Our results revealed that the first 2 principal components accounted for ~ 50% of the overall variance, and the first four principal components accounted for nearly 80% of the variance. Here, we show the relationship between the vectors for all the predictor variables, which conveys that A3 was predominantly associated with principal component 2, G3 was predominantly associated with principle component 1, and R3 and L2 were associated with a combination of the two

cFos+Orexin co-labeling in the lateral hypothalamus predict extinction LTM freezing

We first confirmed that our best subset model (A3) accounted for a significant portion of the variance in LTM freezing for just the subset of rats that we processed for cFos+orexin (21 rats that were processed for immunohistochemistry and had received a CO_2). We found that it was the case, with an R^2 of.5. For that subgroup, A3 accounted for 50% of the observed variance in LTM freezing, $F(1, 19) = 19.44$, $p < 0.001$. cFos+orexin co-labeling in the LH was found to account for a significant portion of the variance in LTM freezing, with an R^2 of 0.24, $F(1,19) = 5.925$, $p = 0.024$ (See Fig. 8). Additionally, there was a significant difference in the number of cells that expressed cFos and orexin between those that were the lower and upper 40% freezing, $p = 0.02$. There was also a significant difference between those sub-groups for LTM freezing, $p < 0.0001$ (See Fig. 10). Additional correlations of the 4 best predictors and orexin with cFos cell counts and extinction LTM freezing are shown in the Supplementary Materials.

Predictive effects of CO_2 reactivity are specific to extinction phenotype

There was no predictive effect of CO_2 reactivity (A3) in the elevated plus maze (Fig. 11a) or light dark box (Fig. 11b),

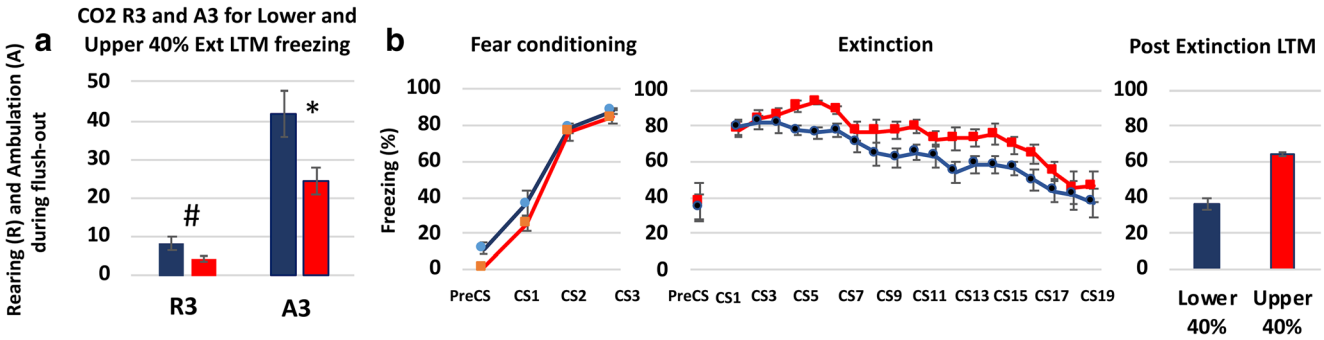

Fig. 8 CO_2 reactivity, fear conditioning, extinction, and post-extinction LTM for the rats in the Lower (Blue) and Upper (Red) 40% range of LTM freezing. CO_2 reactivity is shown here using 2 of the most important predictors of extinction phenotype: rearing and ambulation during the flush-out periods (R3 and A3). The CO_2 reactivity measures (**a**) between the lower and upper 40% freezers approached significance, and there were significant differences between the lower and upper 40% freezers during extinction, and at the post-extinction LTM test, but not during Fear conditioning, preCS freezing, or freezing during the first 3 trials of extinction (**b**), suggesting that CO_2 reactivity may be specific to differences in extinction rather than fear learning. $N = 22$, * = $p < 0.05$, # = $p < 0.1$, and p greater than .2 for all other comparisons

$R^2 = 0.0159$, $F(1,10) = .1612$, $p = 0.695$ and $R^2 = 0.052$, $F(1,20) = 1.091$, $p = 0.309$, respectively. Importantly, there was no interaction effect of prior CO_2 exposure (compared to prior normoxic air exposure) on fear conditioning, preCS freezing, early extinction, or post-extinction long-term memory (Fig. 11c). There was an overall main effect of CO2 group ($p < .05$). Follow-up tests revealed that while there were no differences between CO2 and normoxic air on fear conditioning, preCS freezing, early extinction, or post-extinction LTM (all comparisons, $p > 0.10$); there was a significant difference between the groups for the end of extinction ($p < 0.05$). Additional correlations between the light-dark box and elevated plus maze data with the extinction LTM freezing are shown in the Supplementary Materials.

Discussion

We set out to test whether CO_2 reactivity might be a predictor of extinction phenotype in rats. Here, we validated a rat model of CO_2 reactivity and showed that subcomponents of behavioral reactivity following acute CO_2 exposure explained a significant portion of the variance in extinction LTM. Furthermore, we showed evidence that variability in CO_2 reactivity was also significantly predictive of orexin activity in the LH, and that orexin activity in the LH significantly predicted extinction LTM freezing.

There is strong evidence suggesting that CO_2 exposure directly engages the fear system. In humans, adults with anxiety disorders display heightened emotional reactivity to a single inhalation of 35% CO_2, and Telch et al. (2012) reported that emotional reactivity to a single inhalation of 35% carbon dioxide was significantly predictive of later symptoms of post-traumatic stress disorder and anxiety in soldiers deployed to Iraq. In rats, exposure to moderate concentrations of hypercarbic gas (e.g., normoxic, 10 to 20% CO_2) elicits components of a panic-associated response as evidenced by in-

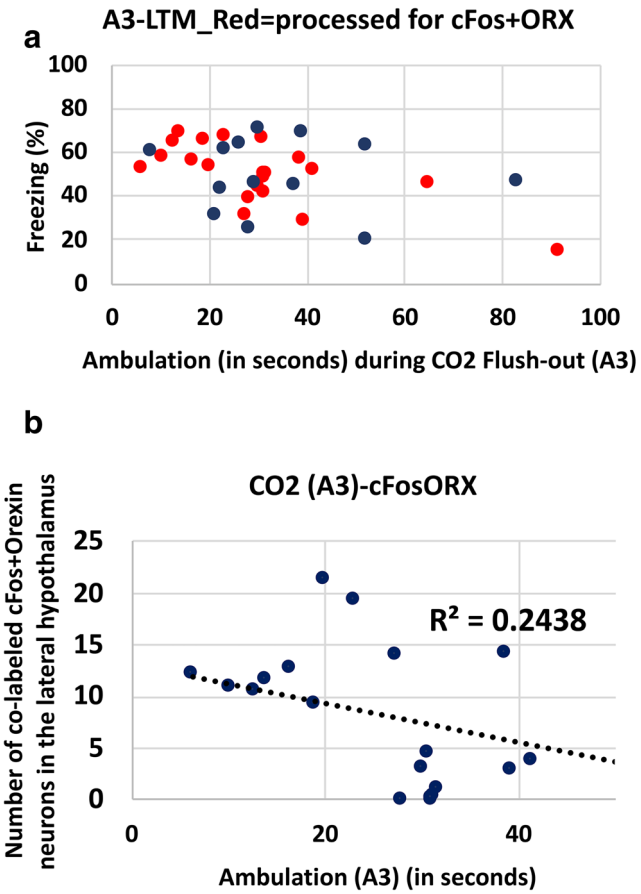

Fig. 9 CO_2 reactivity predicts cFos+Orexin co-labeling in the lateral hypothalamus. **a** In red, rats that were processed and quantified for cFos and orexin immunohistochemistry and were exposed to CO_2 (21 out of 34). Of that subset, our best predictor in the full sample (A3) accounted for 50% of the variance in post-extinction LTM freezing. **b** A3 also significantly predict cFOS+Orexin activity in the lateral hypothalamus, with a combined multiple R^2 of 0.25, $F(1,19) = 6.452$, $p = 0.01$

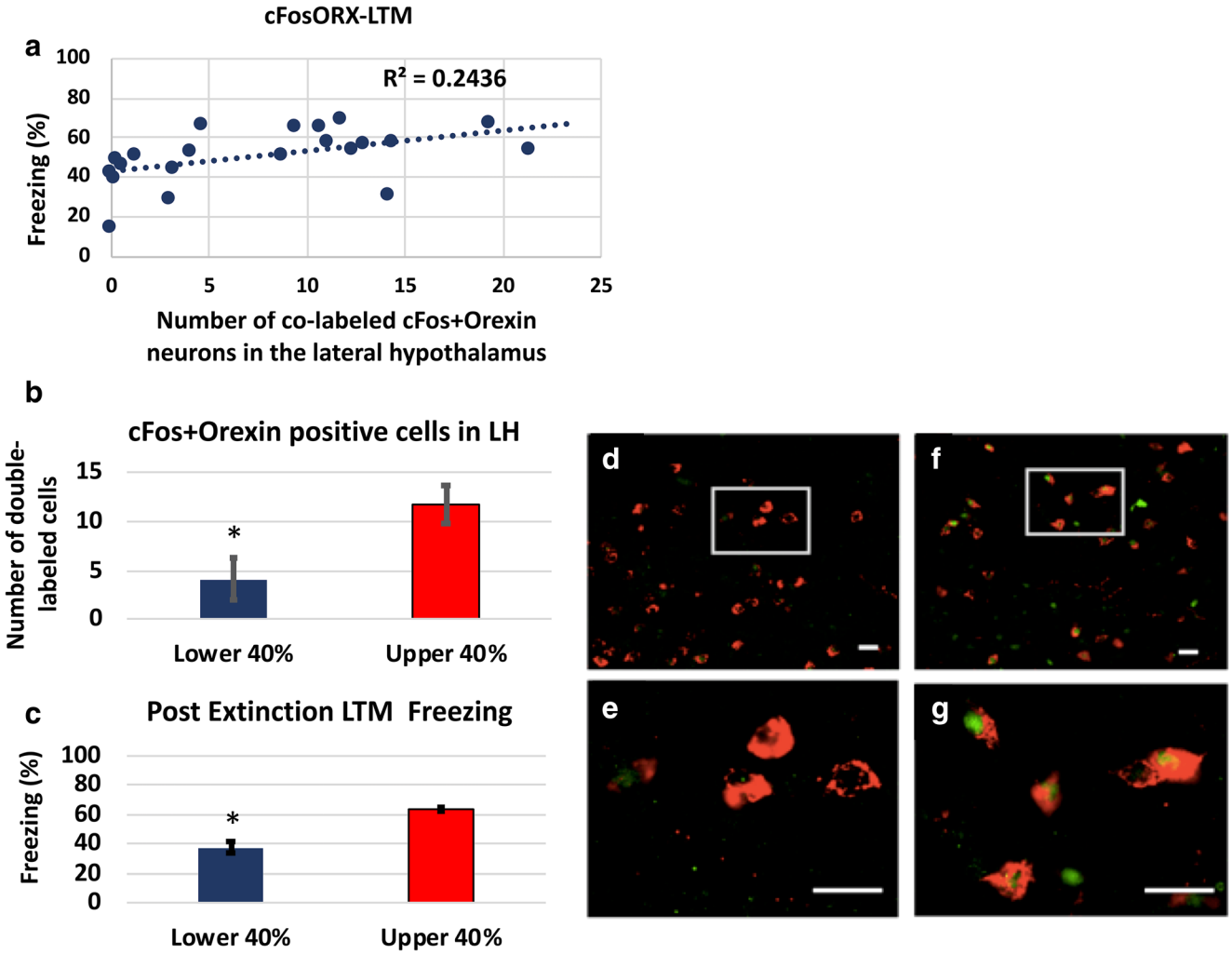

Fig. 10 cFos+Orexin predict extinction LTM freezing. **a** cFos+orexin co-labeling in the LH was found to account for a significant portion of the variance in LTM freezing, with an R^2 of 0.24, $F(1,19) = 5.925$, $p = 0.024$. **b** cFos+Orexin positive cells in the lateral hypothalamus in the lower and upper 40% of post-extinction LTM freezing. There was a significant difference in the number of cells that expressed cFos and orexin between those that were the lower and upper 40% freezing, $p = 0.02$. **c** Lower and upper 40% post-extinction LTM freezing in rats processed for cFos and Orexin ($N = 12$). Lower and upper groups were significantly different, $p < 0.0001$. Orexin (red) and cFos (green) immunohistochemistry in LH of good (**d**, **e**) and poor (**f**, **g**) extinguisher following CO2 reactivity. There appears to be more cFos-labeled orexin cells in the poor extinguisher, in support of our hypothesis. Scale bar = 20 μm

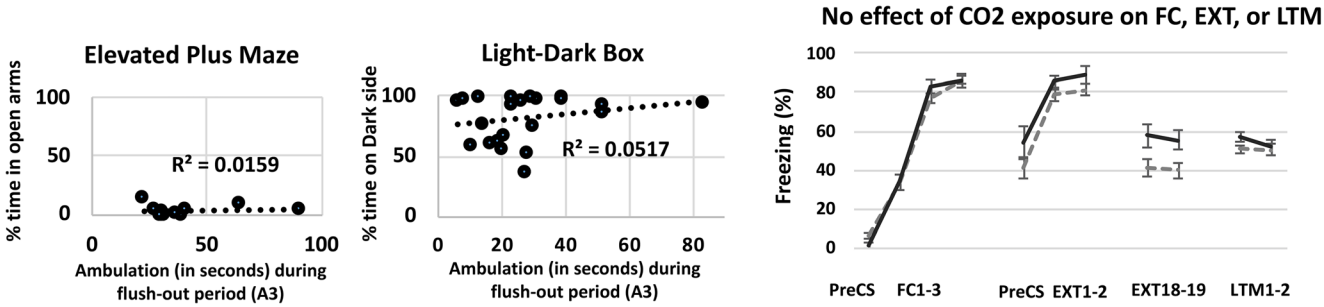

Fig. 11 Predictive effects of CO_2 reactivity are specific to extinction phenotype. There is no predictive effect in elevated plus maze (**a**) or light dark box (**b**). Importantly, there is no effect of prior CO_2 exposure (compared to prior control Air exposure) on fear conditioning, preCS freezing, early extinction, or post-extinction long-term memory. There was a significant difference between CO_2 and normoxic air at the end of extinction ($p < 0.05$) (**c**). All non-significant comparisons, $p > .1$

creases in sympathetic activity (Elam et al. 1981), blood pressure (Johnson et al. 2012; Walker and Brizzee 1990), and anxiety-like behaviors (e.g., increases in panic/escape/flight-associated locomotor responses) (Johnson et al. 2012; Cuccheddu et al. 1995; Johnson et al. 2011). Furthermore, exposing healthy rats to 20% CO_2 concentrations selectively increases cellular activity in panic and anxiety networks, such as the perifornical hypothalamus and dorsal periaqueductal gray (Johnson et al. 2011). When stimulated, those same areas are known to induce panic-associated symptoms in humans (Nashold et al. 1969; Wilent et al. 2011) and panic-associated behaviors and cardio-excitation in rats (Samuels et al. 2002; Shekhar and Katner 1995). Laboratory provocation of panic symptoms has been widely used as a means of investigating the pathogenesis of anxiety-related disorders (Margraf et al. 1986; McNally 1999); yet, to date, not one had tested whether emotional reactivity to a single CO_2 challenge might be predictive of fear reduction outcomes in response to standard extinction treatments. Rats, much like humans albeit to a lesser degree, show individual differences in responding to fear extinction. Individuals with PTSD show deficits in extinction (Pitman et al. 2012), and dysregulation of HPA axis (Yehuda 2009; Michopoulos et al. 2017). Individual differences in extinction as well as CO_2 exposure have been respectively found to activate orexin neurons in the lateral hypothalamus (Johnson et al. 2011). Orexin from the lateral hypothalamus modulates amygdala threat (fear) learning (Sears et al. 2013), and orexin receptor antagonism has been found to facilitate extinction from context and cued fear conditioning (Flores et al. 2014). Furthermore, antagonism of orexin receptors increases the recruitment of BLA neurons that project to the infralimbic cortex during extinction (Flores et al. 2017). Those very same neurons (the IL projecting BLA neurons) are the ones found to be active during extinction (Senn et al. 2014), supporting the notion that individual differences in orexin activation in the LH could account for individual differences in extinction (Sharko et al. 2017), and that those differences could be estimated, in vivo, through CO_2 reactivity. The findings from the present study show that this may indeed be the case.

The effect of CO_2 reactivity appears to be specific to extinction—the variability in CO_2 reactivity did not account for significant variability in either elevated plus maze or light-dark box behavior. Our findings further show that CO_2 exposure per se does not differentially affect fear acquisition, early extinction, or extinction long-term memory relative to normoxic air.

CO_2 exposure is inexpensive and can be safely administered in humans (e.g., Telch et al. 2010, 2012). Together, the results of our study open the possibility that we may be able to use CO_2 reactivity as a screening tool to determine if individuals are good candidates for an extinction/exposure-based approach. If they are not, other treatment avenues could be considered that may be better suited for these particular individuals. Evidence suggests that reconsolidation-based approaches may engage mechanisms that differ from those engaged during extinction (e.g., Tedesco et al. 2014; Lee et al. 2016; Monfils et al. 2009; Schiller et al. 2014). If CO_2 reactivity is reflective of individual differences in susceptibility to respond to extinction treatment, it could also indicate a differential engagement of networks during extinction, which may make a candidate better suitable for an approach other than extinction (such as reconsolidation blockade or updating). Indeed, this idea would also be very much in line with a recent article showing that inhaling CO2 enhances the ability of retrieval to render an aversive memory labile (Du et al. 2017). These findings appear in support of a reconsolidation-based hypothesis.

Our study illustrates that CO_2 reactivity is significantly predictive of Orexin activity in the LH. In recent years, a number of studies have found a relationship between Orexin activity and fear extinction. Our findings suggest the possibility that CO_2 reactivity could serve as a proxy to examine Orexin activity in an effective, rapid, and inexpensive way.

Overall, although CO_2 reactivity accounted for a significant portion of the variance in long-term memory after extinction, it remains to be seen whether we can extrapolate the results to a different subsample. Our cross-validation estimates from random sub-samples within our dataset combined with the positive test in our orexin sub-sample, as well as the fact that the same predictors were found to be of importance across a number of different validation estimates, increase our confidence that the selected model will generalize to a new sample. Still, our nested cross-validation suggests that models developed on this small data set may not yet generalize to new samples. Ultimately, it will be important to assess whether CO_2's predictive power translates into meaningful and practical predictive significance (i.e., improvement in overall treatment outcomes within a subsample).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Bush DE, Sotres-Bayon F, LeDoux JE (2007) Individual differences in fear: isolating fear reactivity and fear recovery phenotypes. J Trauma Stress 20:413–422. https://doi.org/10.1002/jts.20261

Cawley GC, Talbot NL (2010) On over-fitting in model selection and subsequent selection Bias in performance evaluation. J Mach Learn Res 11:2079–2107

Cuccheddu T, Floris S, Serra M, Porceddu ML, Sanna E, Biggio G (1995) Proconflict effect of carbon dioxide inhalation in rats. Life Sci 56: L321–L324. https://doi.org/10.1016/0024-3205(95)00093-3

Crawley J, Goodwin FK (1980) Preliminary report of a simple animal behavior model for the anxiolytic effects of benzodiazepines.

Pharmacol Biochem Behav 13(2):167–170. https://doi.org/10.1016/0091-3057(80)90067-2

Cunha C, Monfils MH, LeDoux JE (2010) $GABA_C$ receptors in the lateral amygdala: a possible novel target for the treatment of fear and anxiety disorders? Front Behav Neurosci 4:6. https://doi.org/10.3389/neuro.08.006.2010

Du J, Price MP, Taugher RJ, Grigsby D, Ash JJ, Stark AC, Hossain Saad MZ, Singh K, Mandal J, Wemmie JA, Welsh MJ (2017) Transient acidosis while retrieving a fear-related memory enhances its lability. eLife, 6. https://doi.org/10.7554/eLife.22564

Elam M, Yao T, Thore P, Svensson TH (1981) Hypercapnia and hypoxia: chemoreceptor-mediated control of locus coeruleus neurons and splanchnic, sympathetic nerves. Brain Res 222(2):373–381. https://doi.org/10.1016/0006-8993(81)91040-4

Flores Á, Valls-Comamala V, Costa G, Saravia R, Maldonado R, Berrendero F (2014) The hypocretin/orexin system mediates the extinction of fear memories. Neuropsychopharmacology 39(12):2732–2741. https://doi.org/10.1038/npp.2014.146

Flores Á, Herry C, Maldonado R, Berrendero F (2017) Facilitation of contextual fear extinction by orexin-1 receptor antagonism is associated with the activation of specific amygdala cell subpopulations. Int J Neuropsychopharmacol 20(8):654–659. https://doi.org/10.1093/ijnp/pyx029

Fox J, Weisberg S (2011) An {R} companion to applied regression, second edition. Thousand Oaks CA: sage. In: URL http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Galatzer-Levy IR, Ankri Y, Freedman S, Israeli-Shalev Y, Roitman P, Gilad M, Shalev AY (2013) Early PTSD symptom trajectories: persistence, recovery, and response to treatment: results from the Jerusalem trauma outreach and prevention study (J-TOPS). PLoS One 8(8):e70084. https://doi.org/10.1371/journal.pone.0070084

Johnson PL, Samuels BC, Fitz SD, Lightman SL, Lowry CA, Shekhar A (2012) Activation of the orexin 1 receptor is a critical component of CO_2-mediated anxiety and hypertension but not bradycardia. Neuropsychopharmacology 37:1911–1922. https://doi.org/10.1038/npp.2012.38

Johnson PL, Fitz SD, Hollis JH, Moratalla R, Lightman SL, Shekhar A, Lowry CA (2011) Induction of c-fos in "panic/defense"-related brain circuits following brief hypercarbic gas exposure. J Psychopharmacol 25:26–36. https://doi.org/10.1177/0269881109353464

Jones CE, Ringuet S, Monfils MH (2013) Learned together, extinguished apart: reducing fear to complex stimuli. Learn Mem 20(12):674–685. https://doi.org/10.1101/lm.031740.113

Jones CE, Monfils MH (2016) Dominance status predicts social fear transmission in laboratory rats. Anim Cogn 19(6):1051–1069. https://doi.org/10.1007/s10071-016-1013-2

Lee HJ, Groshek F, Petrovich GD, Cantalini JP, Gallagher M, Holland PC (2005) Role of amygdalo-nigral circuitry in conditioning of a visual stimulus paired with food. J Neurosci 25(15):3881–3888. https://doi.org/10.1523/JNEUROSCI.0416-05.2005

Lee HJ, Gallagher M, Holland PC (2010) The central amygdala projection to the substantia nigra reflects prediction error information in appetitive conditioning. Learn Mem 17(10):531–538. https://doi.org/10.1101/lm.1889510

Lee HJ, Haberman RP, Roquet RF, Monfils MH (2016) Extinction and retrieval+ extinction of conditioned fear differentially activate medial prefrontal cortex and amygdala in rats. Front Behav Neurosci 9:369. https://doi.org/10.3389/fnbeh.2015.00369

Margraf J, Ehlers A, Roth WT (1986) Biological models of panic disorder and agoraphobia–a review. Behav Res Ther 24(5):553–567. https://doi.org/10.1016/0005-7967(86)90036-7

McNally RJ (1999) On the experimental induction of panic. Behav Ther 30(2):331–339. https://doi.org/10.1016/S0005-7894(99)80012-4

Michopoulos V, Powers A, Gillespie CF, Ressler KJ, Jovanovic T (2017) Inflammation in fear-and anxiety-based disorders: PTSD, GAD, and beyond. Neuropsychopharmacology 42(1):254–270. https://doi.org/10.1038/npp.2016.146

Monfils MH, Cowansage KK, Klann E, LeDoux JE (2009) Extinction-reconsolidation boundaries: key to persistent attenuation of fear memories. Science 324(5929):951–955. https://doi.org/10.1126/science.1167975

Nashold BS Jr, Wilson WP, Slaughter DG (1969) Sensations evoked by stimulation in the midbrain of man. J Neurosurg 30:14–24. https://doi.org/10.3171/jns.1969.30.1.0014

Pitman RK, Rasmusson AM, Koenen KC, Shin LM, Orr SP, Gilbertson MW, Milad MR, Liberzon I (2012) Biological studies of post-traumatic stress disorder. Nat Rev Neurosci 13(11):769–787. https://doi.org/10.1038/nrn3339

R Core Team (2016) R: a language and environment for statistical computing. In: R Foundation for Statistical Computing, Vienna, Austria URL https://www.R-project.org/

Samuels BC, Zaretsky DV, DiMicco JA (2002) Tachycardia evoked by disinhibition of the dorsomedial hypothalamus in rats is mediated through medullary raphe. J Physiol 538:941–946. https://doi.org/10.1113/jphysiol.2001.013302

Schiller D, Kanen JW, LeDoux JE, Monfils MH, Phelps EA (2014) Extinction during reconsolidation of threat memory diminishes prefrontal cortex involvement. Proc Natl Acad Sci U S A 110(50):20040–20045. https://doi.org/10.1073/pnas.1320322110

Schwartze D, Barkowski S, Strauss B, Knaevelsrud C, Rosendahl J (2017) Efficacy of group psychotherapy for posttraumatic stress disorder: Systematic review and meta-analysis of randomized controlled trials. Psychother Res:1–17. https://doi.org/10.1080/10503307.2017.1405168

Sears RM, Fink AE, Wigestrand MB, Farb CR, de Lecea L, Ledoux JE (2013) Orexin/hypocretin system modulates amygdala-dependent threat learning through the locus coeruleus. Proc Natl Acad Sci U S A 110(50):20260–20265. https://doi.org/10.1073/pnas.1320325110

Senn V, Wolff SB, Herry C, Grenier F, Ehrlich I, Gründemann J, Fadok JP, Müller C, Letzkus JJ, Lüthi A (2014) Long-range connectivity defines behavioral specificity of amygdala neurons. Neuron 81(2):428–437. https://doi.org/10.1016/j.neuron.2013.11.006

Sharko AC, Fadel JR, Kaigler KF, Wilson MA (2017) Activation of orexin/hypocretin neurons is associated with individual differences in cued fear extinction. Physiol Behav 178:93–102. https://doi.org/10.1016/j.physbeh.2016.10.008

Jason Shumake (2018) beset: Best Subset Predictive Modeling. R package version 0.0.0.9402. https://github.com/jashu/beset

Shumake J, Furgeson-Moreira S, Monfils MH (2014) Predictability and heritability of individual differences in fear learning. Anim Cogn 17(5):1207–1221. https://doi.org/10.1007/s10071-014-0752-1

Shumake J, CE J, A A, Monfils MH (2018) Data-driven criteria to assess fear remission and phenotypic variability of extinction in rats. Phil Trans R Soc B 373(1742):20170035. https://doi.org/10.1098/rstb.2017.0035

Shekhar A, Katner JS (1995) Dorsomedial hypothalamic gaba regulates anxiety in the social interaction test. Pharmacol Biochem Behav 50:253–258. https://doi.org/10.1016/0091-3057(94)00307-5

Taylor J, Tibshirani RJ (2015) Statistical learning and selective inference. Proc Natl Acad Sci U S A 112(25):7629–7634. https://doi.org/10.1073/pnas.1507583112

Tedesco V, Roquet RF, DeMis J, Chiamulera C, Monfils MH (2014) Extinction, applied after retrieval of auditory fear memory, selectively increases zinc-finger protein 268 and phosphorylated ribosomal protein S6 expression in prefrontal cortex and lateral amygdala. Neurobiol Learn Mem. https://doi.org/10.1016/j.nlm.2014.08.015

Telch MJ, Rosenfield D, Lee HJ, Pai A (2012) Emotional reactivity to a single inhalation of 35% carbon dioxide and its association with later symptoms of posttraumatic stress disorder and anxiety in soldiers

deployed to Iraq. Arch Gen Psychiatry 69(11):1161–1168. https://doi.org/10.1001/archgenpsychiatry.2012.8

Telch MJ, Smits JAJ, Brown M, Dement M, Powers MB, Lee HJ, Pai A (2010) Effects of threat context and cardiac sensitivity on fear responding to a 35% CO2 challenge: a test of the context-sensitivity panic vulnerability model. J Behav Ther Exp Psychiatry 41:365–372. https://doi.org/10.1016/j.jbtep.2010.03.008

Walker BR, Brizzee BL (1990) Cardiovascular responses to hypoxia and hypercapnia in barodenervated rats. J Appl Physiol 68:678–686. https://doi.org/10.1152/jappl.1990.68.2.678

Wilent WB, Oh MY, Buetefisch C, Bailes JE, Cantella D, Angle C, Whiting DM (2011) Mapping of microstimulation evoked responses and unit activity patterns in the lateral hypothalamic area recorded in awake humans. Technical note J Neurosurg 115(2):295–300. https://doi.org/10.3171/2011.3.JNS101574

Yehuda R (2009) Status of glucocorticoid alterations in post-traumatic stress disorder. Ann N Y Acad Sci 1179(1):56–69. https://doi.org/10.1111/j.1749-6632.2009.04979.x