Acute stress impairs reward learning in men

Joana Carvalheiro [a,*], Vasco A. Conceição [b], Ana Mesquita [a], Ana Seara-Cardoso [a,*]

[a] Escola de Psicologia, CIPsi, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal
[b] Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa, Avenida Professor Egas Moniz, 1649-028 Lisboa, Portugal

ABSTRACT

Acute stress is ubiquitous in everyday life, but the extent to which acute stress affects how people learn from the outcomes of their choices is still poorly understood. Here, we investigate how acute stress impacts reward and punishment learning in men using a reinforcement-learning task. Sixty-two male participants performed the task whilst under stress and control conditions. We observed that acute stress impaired participants' choice performance towards monetary gains, but not losses. To unravel the mechanism(s) underlying such impairment, we fitted a reinforcement-learning model to participants' trial-by-trial choices. Computational modeling indicated that under acute stress participants learned more slowly from positive prediction errors — when the outcomes were better than expected — consistent with stress-induced dopamine disruptions. Such mechanistic understanding of how acute stress impairs reward learning is particularly important given the pervasiveness of stress in our daily life and the impact that stress can have on our wellbeing and mental health.

1. Introduction

Learning to choose options that lead to rewards and to avoid those that result in punishments is crucial for adaptive behavior. Situational factors, such as stress, can have deleterious effects on the ability to make the best choices and learn from them (Porcelli & Delgado, 2017). Stress is present in our day-to-day life, but, notably, how acute stress affects reward and punishment learning remains largely unknown. A growing body of evidence suggests that acute stress impairs reward-seeking behavior (Berghorst, Bogdan, Frank, & Pizzagalli, 2013; Bogdan & Pizzagalli, 2006; Bogdan, Santesso, Fagerness, Perlis, & Pizzagalli, 2011; Ehlers & Todd, 2017; Morris & Rottenberg, 2015; Paret & Bublatzky, 2020; but see Lighthall, Gorlick, Schoeke, Frank, & Mather, 2013), but less is known about the impact of acute stress on punishment-avoidance behavior (Aylward et al., 2019; Petzold, Plessow, Goschke, & Kirschbaum, 2010). More critically, there is even less evidence on the mechanisms that underlie the behavioral effects of acute stress on reward and punishment learning. Here, we use a computational reinforcement-learning framework to investigate the impact of acute stress on reward and punishment learning in men.

In the past decades, the use of computational modeling approaches to describe behavior-brain relationships in healthy humans has played an influential role on cognitive science (Daw & Frank, 2009; Frank, 2015; Huys, Maia, & Frank, 2016; Maia & Frank, 2011; Maia, 2015). Computational models, such as reinforcement-learning models, are built and implemented to capture very specific cognitive and neural mechanisms, thus linking different levels of analysis, from cognitive and behavioral phenomena to neurobiological mechanisms (Chater, 2009; Daw & Frank, 2009; Nair, Rutledge, & Mason, 2020). Reinforcement-learning models are considered extremely useful tools to investigate the neural computations underpinning cognition and behavior (Collins & Frank, 2013; Daw, 2011; Daw & Frank, 2009; Huys et al., 2016; Maia & Frank, 2011; Nair et al., 2020), and can thus provide a mechanistic framework to disentangle the effects of acute stress on reward and punishment learning (Aylward et al., 2019; Huys, Pizzagalli, Bogdan, & Dayan, 2013; Luksys & Sandi, 2011; Otto, Raio, Chiang, Phelps, & Daw, 2013; Radenbach et al., 2015; Robinson, Overstreet, Charney, Vytal, & Grillon, 2013).

According to reinforcement-learning theory, individuals learn to gradually select more and more often the actions that maximize rewards and those that minimize punishments by learning the values of the executed actions, and such learning is driven by prediction errors (Maia & Frank, 2011; Schultz, Dayan, & Montague, 1997; Sutton & Barto, 1998). Specifically, prediction errors — which signal the difference between obtained and expected outcomes — are used to progressively update the values of the executed actions (Collins & Frank, 2013; Maia &

Frank, 2011; Schultz et al., 1997; Sutton & Barto, 1998). Prediction errors can be positive or negative. Positive prediction errors occur when outcomes are better than expected, and can thus signal unexpected rewards or avoidance of expected punishments (Daw & Tobler, 2014; Schultz et al., 1997; Watabe-Uchida, Eshel, & Uchida, 2017). Negative prediction errors occur when outcomes are worse than expected, thus signaling unexpected punishing outcomes or omission of expected rewards (Daw & Tobler, 2014; Schultz et al., 1997; Watabe-Uchida et al., 2017). This means that positive and negative prediction errors can be both present in reward and punishment learning. Interestingly though, recent evidence indicates that in simple reinforcement-learning tasks, reward learning should be mostly driven by positive prediction errors (Lefebvre, Lebreton, Meyniel, Bourgeois-Gironde, & Palminteri, 2017; Palminteri & Pessiglione, 2013), while punishment learning seems to be largely driven by negative prediction errors (Palminteri & Pessiglione, 2013). Studies that do not consider the positive and negative valence of prediction errors when modeling reward and/or punishment learning often assume that, during reward learning, obtaining unexpected rewards (positive prediction errors) is coded similarly to the omission of expected rewards (negative prediction errors), and that, during punishment learning, obtaining unexpected punishments (negative prediction errors) is coded similarly to the omission of expected punishments (positive prediction errors). However, extant evidence suggests that the brain encodes positive and negative prediction errors differently (Frank, 2005; Maia & Frank, 2011; O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003; Schultz et al., 1997). Thus, to better understand the mechanisms underlying reward and punishment learning, computational models could benefit from the integration of such neurobiological evidence and model positive and negative prediction errors separately. Importantly, reinforcement-learning models can capture how quickly unexpected positive and negative outcomes are integrated over time through distinct learning rates for positive and negative prediction errors, respectively (Frank, Moustafa, Haughey, Curran, & Hutchison, 2007). Relatedly, blunted signaling of positive and negative prediction errors can be captured by reinforcement-learning models as reduced learning rates.

Dopaminergic functioning plays a key role in prediction-error-based learning (Frank, Seeberger, & O'Reilly, 2004; Glimcher, 2011; Maia & Frank, 2011; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006). Prediction-error signals are known to be encoded in the phasic activity of dopamine neurons (Daw & Tobler, 2014; Maia & Frank, 2011; Schultz et al., 1997). Specifically, phasic bursts of dopaminergic neurons are thought to adaptively encode positive prediction errors, whereas dopamine dips have been associated with the adaptive encoding of negative prediction errors (Daw & Tobler, 2014; Maia & Frank, 2011; Schultz et al., 1997). However, phasic-dopamine responses do not seem to be always adaptive, and there is evidence that dopamine can be phasically released in an aberrant spontaneous manner (Belujon, Grace, & Grace, 2015; Maia & Frank, 2017; Sulzer, Cragg, & Rice, 2016). Studies with non-human male animals suggest that acute stress increases aberrant spontaneous phasic-dopamine release in the mesolimbic pathway, which originates in the ventral tegmental area and projects to terminal regions such as the striatum (Anstrom, Miczek, & Budygin, 2009; Anstrom & Woodward, 2005; Valenti, Lodge, & Grace, 2011). Crucially, exaggerated, aberrant spontaneous dopamine release seems to reduce adaptive phasic bursts that signal positive prediction errors in the striatum (Daberkow et al., 2013; Grace, 2016; Maia & Frank, 2017; Werlen et al., 2020). Additionally, though more speculative (Maia & Frank, 2017), aberrant spontaneous dopamine release might also block the effects of dopamine dips (Frank, 2005; Frank & O'Reilly, 2006) needed to signal negative prediction errors. Still, the extent to which stress-induced aberrant spontaneous dopamine release blunts signaling of positive and/or negative prediction errors remains poorly understood.

Stress can be defined as a subjective state of sensing potentially adverse changes in the environment (Joëls & Baram, 2009). Stressful

stimuli (i.e., stressors) can elicit many physiological and functional changes in the brain in response to perceived adverse changes in the environment (de Kloet, Joëls, & Holsboer, 2005; Hermans, Henckens, Joëls, & Fernández, 2014; Joëls & Baram, 2009). The first wave of the stress response occurs rapidly and involves enhanced release of catecholamines, such as noradrenaline and dopamine. The catecholamines release generally initiates almost instantly after the onset of the stressor and the rapid raise in catecholamines levels is quickly translated into changes in neural functioning (Hermans et al., 2014; Joëls & Baram, 2009; Sapolsky, Romero, & Munck, 2000). For instance, increased release of noradrenaline seems to contribute to the reorientation of attention to a more general scanning of the environment (Hermans et al., 2014) and dopamine, which is released during stress exposure in the prefrontal cortex and striatum (Abercrombie, Keefe, DiFrischia, & Zigmond, 1989; Anstrom et al., 2009; Anstrom & Woodward, 2005; Cabib & Puglisi-Allegra, 2012; Nagano-Saito et al., 2013; Pruessner, Champagne, Meaney, & Dagher, 2004), is thought to alter reward learning rein (Porcelli & Delgado, 2017). A second, and slower, wave of the stress response involves secretion of glucocorticoids. Cortisol usually has slow, prolonged actions on neural functioning (de Kloet et al., 2005; Hermans et al., 2014; Joëls & Baram, 2009; Sapolsky et al., 2000). Peak concentrations of cortisol are typically not reached before 15–30 min after the stressor onset (de Kloet et al., 2005; Hermans et al., 2014), which points to a limited role of cortisol in the immediate stress response.

Stress exerts various effects on cognitive functioning (de Kloet et al., 2005), namely via fast-initial release of catecholamines and slow-delayed release of cortisol, as mentioned above. Thus, one important factor when studying the impact of acute stress on learning is to consider whether learning is tested immediately after the onset of the stressor when catecholamines levels are higher or later when cortisol levels peak (Hermans et al., 2014; Joëls, Pu, Wiegert, Oitzl, & Krugers, 2006; Porcelli & Delgado, 2017; Schwabe, Joëls, Roozendaal, Wolf, & Oitzl, 2012). A significant number of stress and learning studies have considered only slower effects of cortisol in the stress response, adopting designs that may involve relatively longer latencies between stress induction and task solving (Byrne, Cornwall, & Worthy, 2019; de Berker et al., 2016; Ehlers & Todd, 2017; Kruse, Tapia León, Stalder, Stark, & Klucken, 2018; Lighthall et al., 2013; Otto et al., 2013; Paul, Bellebaum, Ghio, Suchan, & Wolf, 2019; Petzold et al., 2010). In those studies, participants are often exposed to a stress manipulation, such as the Trier social stress test or the (socially evaluated) cold pressor test, and the impact of stress on learning is assessed around 10–20 min after the stress manipulation has occurred, so that cortisol levels peak during the learning task. Such stress inductions can combine a social evaluation component, unpredictability, and uncontrollability, which are known to produce robust increases in cortisol levels (McRae et al., 2006; Schwabe, Haddad, & Schachinger, 2008). However, in those paradigms, it is less clear if individuals are actually perceiving stress at the time of learning, as the stressor is experienced outside of the learning context (McRae et al., 2006; Schwabe et al., 2012). Moreover, it is not unlikely that stress recovery occurs during the learning task (see de Berker et al., 2016; Paul et al., 2019; Radenbach et al., 2015; Schwabe & Wolf, 2009, for examples of studies that show decreasing cortisol levels during task solving) due to suppressive actions of glucocorticoids (Hermans et al., 2014; Sapolsky et al., 2000).

Fast-acting stress-induced catecholaminergic effects, such as increased release of dopamine (Abercrombie et al., 1989; Anstrom et al., 2009; Anstrom & Woodward, 2005; Cabib & Puglisi-Allegra, 2012; Nagano-Saito et al., 2013; Pruessner et al., 2004), can influence brain function in regions known to be critically associated with learning processes. Thus, by exposing participants to an acute stressor whilst they perform a learning task, one can conceivably capture the impact of stress-induced catecholamine (e.g., dopamine) release on learning. However, one potential drawback of using acute stressors, such as threat of shock, time pressure or noise, during a cognitive task, is that it can make it difficult to disentangle whether cognitive impairments occur as

a result of stress or because the stressor acted as a distractor (Robinson, Vytal, Cornwell, & Grillon, 2013). Indeed, the first wave of the stress response can trigger a sensory hypervigilant state and the redirection of attention to stressors, conceivably causing distraction in tasks that require selective attention (Hermans et al., 2014). Importantly though, distractibility seems to be more robustly induced by unpredictable changes (Parmentier, 2014; Parmentier & Hebrero, 2013; Sussman, Winkler, & Schröger, 2003). Therefore, one possible alternative to minimize the potential distractor effect of an acute stressor during a cognitive task could be to minimize the unpredictability of that stressor.

In this study, we chose to use an acute stressor (with reduced unpredictability) during a reinforcement-learning task, to better understand whether and how acute stress affects reward and/or punishment learning, very likely via putative dopaminergic mechanisms, which can be captured by computational models such as the one we used (Frank et al., 2007). The aim of the present work was twofold. First, to investigate how acute stress impacted behavioral performance during reward and punishment learning. Second, to inspect the computational mechanisms behind the effects of acute stress on reward and punishment learning. Given the putative roles of phasic-dopamine responses on prediction errors signaling and of acute stress on aberrant spontaneous phasic-dopamine release, we hypothesized that acute stress would impair reward learning and, more tentatively punishment learning. Relatedly, we hypothesized that acute stress would decrease the learning rate for positive prediction errors and, more tentatively, the learning rate for negative prediction errors.

To test those hypotheses, we used a well-established reinforcement-learning task involving monetary gains and losses (Pessiglione et al., 2006) combined with a novel acute stress manipulation. Using this task combined with pharmacological manipulations of the dopaminergic system, Pessiglione et al. (2006) showed that dopamine-related drugs modulate prediction errors expressed in the striatum during reward learning. The stress manipulation consisted of exposing participants to an uncontrollable sound (i.e., participants could not put an end to it) whilst they performed the reinforcement-learning task. We chose to use a repetitive alarm sound as stressor, because alarms are known to activate the nervous sympathetic system (e.g., increase heart rate and skin conductance), involved in the earlier phase of the stress response (Hall et al., 2016; Korhonen, 1981; Westman & Walters, 1981). Repetitive, uncontrollable sounds for long periods of time seem to be particularly effective in increasing dopamine release (Arnsten & Goldman-Rakic, 1998) and levels of adrenocorticotropic hormone (Arguelles, Ibeas, Ottone, & Chekherdemian, 1962; Breier et al., 1987). To minimize the potential distractor effect of the stressor, we used a sound that was always constant and repetitive, and therefore predictable, as unexpected changes in sound sequences robustly induce distraction and negatively affect performance in an unrelated task (Hughes, 2014; Parmentier, 2014; Parmentier, Elford, Escera, Andrés, & Miguel, 2008). To further minimize unpredictability, we also instructed participants, before they started the task, that they would hear an alarm, not associated with the task itself or with their performance, during blocks cued with a warning signal and a red background (Hughes, 2014; Parmentier & Hebrero, 2013; Sussman et al., 2003). Although the stressor was not unpredictable, it was uncontrollable (i.e., participants could not put an end to it). Uncontrollable auditory stimuli can elevate stress responses (Arguelles et al., 1962; Breier et al., 1987; Rylander, 2004; Westman & Walters, 1981) and may disrupt dopaminergic mechanisms (Arnsten & Goldman-Rakic, 1998) and cognitive functioning (Glass, Reim, & Singer, 1971). To check the success of the acute-stress manipulation, we collected self-report stress levels at the end of each block of the task and measured skin conductance response (SCR) rate throughout the task. Then, we inspected how acute stress altered reward and punishment learning during the reinforcement-learning task using both classical statistical analyses and computational-model-based analyses of participants' behavioral data. For the latter analyses, we fitted participants' choices with a previously established, biologically inspired, reinforcement-

learning model (Frank et al., 2007), which has been extensively used to investigate the cognitive and behavioral impact of pharmacological manipulations and genetic variations in the dopaminergic system in humans (Diederen et al., 2017; Doll, Hutchison, & Frank, 2011; Frank & Fossella, 2011; Frank et al., 2007; Grogan et al., 2017; Rutledge et al., 2009). The fitted reinforcement-learning model allowed us to examine the effects of acute stress on learning rates for positive and negative prediction errors. In this study, we included only male participants due to females' hormonal-dependent variations on stress responsivity, as well as on reward and punishment learning (Diekhof, Korf, Ott, Schädlich, & Holtfrerich, 2020; Dreher et al., 2007; Ossewaarde et al., 2010).

2. Material and Methods

2.1. Participants

Sixty-two healthy male participants (age range = 18 – 35 years; M = 21.9 years, SD = 3.7 years) were recruited at University of Minho, Portugal. Due to similarities in study design, we used the study of Pezold et al. (2010) as a reference to determine our minimum sample size. To allow for possible data exclusions due to technical problems and/or nonresponders in skin conductance data, participant enrollment in our study was stopped once we reached a sample size two-fold the sample size analyzed in Petzold et. al (2010). Data collection continued until all participants enrolled up to that point were tested. Data collection was stopped before data analyses, independently of the statistical significance of the data. Four participants were excluded from skin-conductance analyses due to poor signal quality. No participants were *a priori* excluded from any other data analyses, although we conducted confirmatory analyses excluding potential outliers to ensure that our results were not driven by extreme values.

All participants provided their informed consent before the experimental session. All experimental procedures were approved by the Ethics Committee of University of Minho.

2.2. Reinforcement-learning task

After a short practice (12 trials), to familiarize participants with the task timings and response keys, participants completed four blocks of an adapted version of a well-established reinforcement-learning task (Pessiglione et al., 2006) (Fig. 1). Each block included three pairs of abstract stimuli, and each pair of stimuli was presented 40 times, making 120 trials in total per block. Different abstract stimuli were used in each block. Each pair of stimuli was associated with a valence: one pair of stimuli was associated with gains (gain 0.5€ or no change), a second pair associated with losses (loss 0.5€ or no change), and a third pair associated with neutral, or non-financial outcomes (look at a 0.5€ coin or no change; for a depiction of neutral trials see Fig. S1a in the Supplementary Material). The outcome probabilities were reciprocally 0.8 and 0.2 for the stimuli in each of the three pairs. On each trial, one pair was randomly presented on the screen, with one stimulus above and the other below a central fixation cross (the stimuli position was counterbalanced across trials). Participants were instructed to choose between the two visual stimuli displayed on the computer screen to maximize payoffs. Missing choices occurred when participants did not press the response keys within 2000 ms (0.38% missing choices: 40 in the stress condition and 72 in the control condition, in a total of 29,760 trials across all participants) and were signaled with a "Missed" message (no other outcome was provided). Missing choices were excluded from data analyses. Before starting the task, participants were informed that they would be paid the amount of money obtained during a randomly selected block, although they all left with the same fixed compensation (15€) for their participation. The experiment was programmed and presented with Cogent 2000 (http://www.vislab.ucl.ac.uk/cogent.php) implemented in MATLAB R2015a (MathWorks).
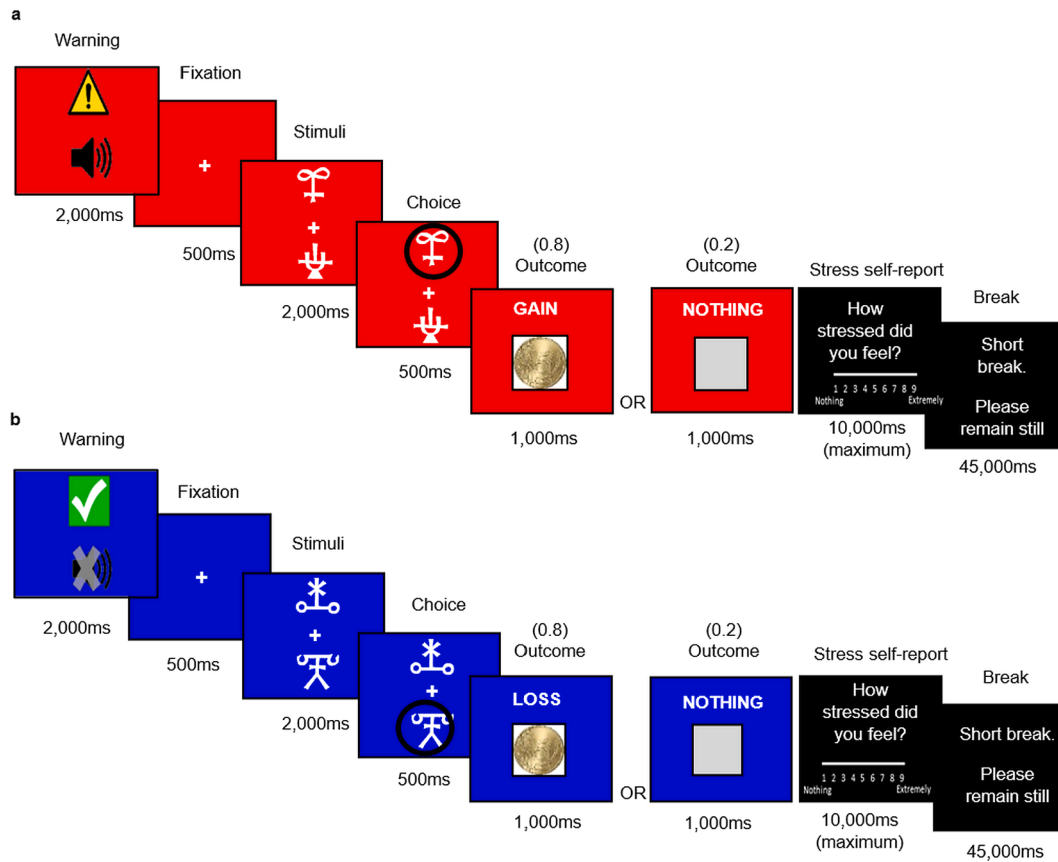
Fig. 1. Reinforcement-learning task. On each trial of the task, participants had to choose either the upper or the lower of two abstract visual stimuli presented on a computer screen and to subsequently observe the obtained outcome, whilst under acute stress and under control conditions. **(a)** In the depicted stress-condition example, the chosen stimulus was associated with a probability of 0.8 of winning 0.5€ and with a probability of 0.2 of obtaining nothing. The other (not chosen) stimulus was associated with a probability of 0.8 of obtaining nothing and with a probability of 0.2 of winning 0.5€. **(b)** In the depicted control-condition example, the chosen stimulus was associated with a probability of 0.8 of losing 0.5€ and with a probability of 0.2 of obtaining nothing. The other (not chosen) stimulus was associated with a probability of 0.8 of obtaining nothing and with a probability of 0.2 of losing 0.5€. Participants completed a total of four blocks, consisting of an alternation between two stress and two control blocks. To assess stress responses, self-reported stress levels were collected at the end of each block and skin conductance was measured throughout the task.

2.3. Acute stress manipulation

During the experimental session, participants performed two blocks of the reinforcement-learning task whilst exposed to a stressor (i.e., stress condition; Fig. 1a) and two blocks without the stressor (i.e., control condition; Fig. 1b). To elicit stress responses we exposed participants to a predictable, but uncontrollable auditory stimulus, a constant alarm ("Annoying modern office building alarm.wav", retrieved from freesound.org, and programmed to loop uninterruptedly), played through the same set of over-ear headphones (GOODIS, model GWH4093, with the volume set to the maximum), with the sound volume adjusted in the laptop to level 27 (in a scale from 0 to 100) for all participants. We conducted a brief pilot to qualitatively confirm that this volume was tolerable, yet stressful. Stress blocks were signaled by a warning sign and a red background (Fig. 1a), and control blocks were signaled by a safe sign and blue background (Fig. 1b). Stress and control blocks were administered alternately and in counterbalanced order. Each block of the task lasted 8 min, and the whole task took around 36 min (including warnings, breaks between blocks, and stress ratings with variable durations). The experiment was conducted in a soundproof room to avoid interference due to environmental noise, and between 12 pm and 6 pm to minimize diurnal variability in stress responses.

2.4. Manipulation check

Stress levels were assessed by asking participants at the end of each block to rate how stressed they felt during that block on a scale of 1 (nothing) to 9 (extremely). To further assess the impact of the acute stressor on autonomic responses, we acquired skin conductance responses (SCRs) using BIOPAC MP150 and a pair of finger electrodes. Electrodes were attached to participants' left index and ring fingers; the gain was set to 5, the low pass filter to 10 Hz, and the high pass filters to DC. Recordings were performed using Acknowledge 4.4. Data were acquired at 200 Hz, downsampled to 62 Hz, and smoothed with a median filter in order to remove outliers. Each participant's SCRs were detected using a threshold of 0 μs and a rejection rate of 10% (Kim, Bang, & Kim, 2004), meaning that we excluded SCR with an amplitude smaller that 10% of the maximum SCR amplitude in each subject, rather than using an arbitrary threshold that would not account for inter-individual differences in SCRs. The SCR rate was calculated by dividing the number of SCRs detected in each block by the duration of that block (in minutes).

Self-reported stress levels and SCR rates were analyzed using repeated-measures analyses of variance (ANOVAs), with condition (stress and control) and block (1 and 2) as within-subject factors, and post-hoc paired *t*-tests. ANOVAs effect sizes are reported as eta-squared, η^2, and post-hoc paired *t*-tests' effect sizes are reported as Cohen's *d* and 95% confidence intervals. We further conducted non-parametric Wilcoxon signed-rank tests, which are more robust to outliers, to confirm

the results from post-hoc paired *t*-tests. Additionally, to confirm that our findings were robust to extreme values, we repeated the SCR rate analyses excluding participants with abnormally large SCR rates. Statistical analyses were conducted using JASP 0.9.

2.5. Task performance analyses

To examine the impact of acute stress on choice performance during the reinforcement-learning task, we applied a generalized linear mixed-effects (glme) model to participants' choice data (with correct and incorrect choices coded as 1 and 0, respectively). We used a "logit" link function to account for the binomial distribution of the data. As predictor variables in the glme model we included condition (stress or control), valence (gains or losses), block number (1 or 2), and trial number (1 to 40), and the interaction of interest (condition × valence; see Table S1 in the Supplementary Material for a full description of the glme model). The glme included a fixed intercept, as well as random intercepts for each participant. We fitted the glme model to the behavioral data using MATLAB's *fitglme* function and conducted post-hoc analyses via contrast matrices using the MATLAB's *coefTest* function. To assess the robustness of our findings, we also tested the significance of the interaction of interest (condition × valence) using confirmatory likelihood ratio tests (Daw, 2011) between the aforementioned full glme model and two nested models, which assumed equal performance in stress and control conditions during either gain or loss trials, through the MATLAB's *lratiotest* function. Additionally, we repeated the analyses excluding participants that performed below chance levels, which is indicative that participants did not learn to perform the task correctly and may additionally reflect non-compliance with the experimental setting.

To confirm that the choice probabilities estimated by the glme model showed a close correspondence with the actual observed choices, we used MATLAB's *predict* function. Then, we assessed the Pearson's correlation between the percentage of actual "correct" choices (i.e., choice of the stimuli associated with a probability of 0.8 of winning or a probability of 0.2 of losing) for each participant in each condition (averaged across blocks) and the percentage of "correct" choices as estimated by the glme model. We also performed confirmatory Spearman's correlations, which are more robust to outliers.

2.6. Computational modeling

2.6.1. Reinforcement-learning model

We modeled participants' trial-by-trial behavior in the stress and control conditions using a reinforcement-learning framework (Sutton & Barto, 1998) that has been extensively used to investigate the behavioral and neural impact of pharmacological manipulations and genetic variations in the dopaminergic system in humans (Diederen et al., 2017; Doll et al., 2011; Frank & Fossella, 2011; Frank et al., 2007; Grogan et al., 2017; Rutledge et al., 2009). Importantly, the fitted model included separate learning rates for positive (α^+) and negative (α^-) prediction errors, to account both for the differential firing of dopaminergic neurons for positive and negative prediction errors (Daw & Tobler, 2014; Maia & Frank, 2011; Maia & Conceição, 2017) and the differential effects of dopamine onto the plasticity of the corticostriatal synapses implicated in action-value learning (Frank & O'Reilly, 2006; Maia & Frank, 2011; Maia & Conceição, 2017; Möller & Bogacz, 2019). This model also included the inverse temperature parameter, β, which controls the stochasticity of choice selection, or the decision noise (Daw, 2011; Sutton & Barto, 1998), as detailed below.

In the context of our experimental study, the specific reinforcement-learning model used (Frank et al., 2007) assumes that each participant gradually learns the value of choosing a given stimulus (say A or B) from a given pair of stimuli (here, "gain" or "loss" stimuli pairs, as the "neutral" pair of stimuli always yielded null monetary outcomes) as a function of the outcome that was obtained on that trial following

stimulus selection. Specifically, each expected pair-stimulus value, or Q-value, was initialized to zero, and for each trial, *t*, within that pair of stimuli, the value of the chosen stimulus (say A was chosen) was updated according to:

$$Q_A(t+1) = Q_A(t) + \alpha * \delta(t)$$

where δ was the prediction error:

$$\delta(t) = r(t) - Q_A(t)$$

where $r(t)$ was 0.5 for gains, 0 for neutral outcomes, and -0.5 for losses. The learning rate, α, was given by:

$$\alpha = \begin{cases} \alpha^+, if\, \delta(t) > 0 \\ \alpha^-, if\, \delta(t) < 0 \end{cases},$$

where α^+ and α^- were the learning rates for positive and negative prediction errors, respectively (Frank et al., 2007).

The probability of choosing one stimulus over another (say A over B) was given by the *softmax* equation:

$$P_A(t) = \frac{e^{[Q_A(t)*\beta]}}{e^{[Q_A(t)*\beta]} + e^{[Q_B(t)*\beta]}}$$

where the β parameter, or inverse temperature, controlled the amount of decision noise. The lower the β, the higher the decision noise (i.e., decisions were made more at random).

2.6.2. Model fitting, parameter analyses, and model validation

We fitted the reinforcement-learning model to the trial-by-trial choice data from each participant in each condition. Model fitting involved estimating the values of the parameters (α^+, α^-, and β) that best accounted for the respective trial-by-trial choices in each condition.

We estimated the best-fitting model parameters (α^+, α^-, and β) for each subject in each condition using maximum *a posteriori* estimation (Daw, 2011). Specifically, to optimize model parameters, we drew the learning rates from Beta distributions [*Beta* (1.1, 1.1)] and the inverse temperature from a Gamma distribution [*Gamma* (1.2, 5)] (Palminteri, Khamassi, Joffily, & Coricelli, 2015). We then used the MATLAB's *fmincon* function, initialized at 100 random starting points of the parameter space, to search for the parameter values that minimized the negative log posterior of the observed sequence of choices, given the previously observed outcomes, with respect to different settings of the model parameters (Daw, 2011).

To assess the effects of acute stress on the parameters estimated by the reinforcement-learning model, we conducted repeated-measures ANOVAs with condition (stress and control) and prediction error valence (positive and negative) as within-subject factors, and post-hoc paired *t*-tests. ANOVAs' effect sizes are reported as eta-squared, η^2, and post-hoc paired *t*-tests effect sizes are reported as Cohen's *d* and 95% confidence intervals. We also confirmed all results from post-hoc tests using non-parametric tests (Wilcoxon signed-rank tests), which are more robust to outliers. Additionally, we repeated all analyses excluding participants that performed below chance levels to test whether the significance of the results remained unchanged. These statistical analyses were conducted using JASP 0.9.

To validate the used reinforcement-learning model, we computed trial-by-trial choice probabilities for all participants using the best-fitting set of parameters in each condition. The actual observed choices and outcomes were used to update the choice probabilities. To assess whether the choice probabilities estimated by the reinforcement-learning model (for each subject, the choice probabilities were averaged across gains or loss trials in each condition) followed the same pattern as the actual observed choices, we conducted Pearson's correlations and confirmatory Spearman's correlations for both conditions, using the respective mean percentages.

To further validate the robustness of our model-fitting procedure, we examined the capacity of recovering subject-condition-specific

parameters using simulated datasets. Specifically, we simulated the task-choice behavior of 62 virtual participants using the parameter values that we had estimated for each of the 62 participants in each condition. We ran 100 simulations. Then, for each simulation, we fitted the model to the virtual participants' data to estimate new (recovered) parameters. Finally, we tested the correlations between the original and these recovered parameters using Pearson's and confirmatory Spearman's correlations.

3. Results

3.1. Manipulation check

First, we confirmed that the acute stress manipulation successfully elicited stress responses in the participants. Self-reported stress levels differed significantly between conditions, $F(1, 61) = 107.67$, $p < .001$, $\eta^2 = 0.64$, as participants reported higher levels of stress in the stress condition ($M = 5.16$, $SEM = 0.21$) than in the control condition ($M = 3.31$, $SEM = 0.20$), $t(61) = 10.38$, $p < .001$, $d = 1.32$, 95% confidence interval (CI) = [1.49, 2.20] (Fig. 2a). SCR rate also differed significantly between conditions, $F(1, 57) = 20.61$, $p < .001$, $\eta^2 = 0.27$, being higher in the stress condition ($M = 2.89$, $SEM = 0.27$) than in the control condition ($M = 2.46$, $SEM = 0.22$), $t(57) = 4.54$, $p < .001$, $d = 0.57$, 95% CI = [0.24, 0.62] (Fig. 2b). The condition × block interactions were non-significant for the self-reported stress levels, $F(1, 61) = 0.004$, $p = .95$, $\eta^2 = 0$, and for the SCR rate, $F(1, 57) = 0.46$, $p = .50$, $\eta^2 = 0.0080$, suggesting that both the self-reported stress levels and the SCR rate remained stable across blocks within conditions. Furthermore, confirmatory non-parametric Wilcoxon signed-rank tests replicated the effects of acute stress on self-reported stress levels, when comparing the stress condition with the control condition, $Z = 6.29$, $p < .001$, and SCR rate, $Z = 4.08$, $p < .001$. Additionally, to check whether our results were robust to extreme values, we identified two participants with abnormally large SCR rate and reanalyzed the data without those participants. The significance of the results remained unchanged (main effect of condition: $F(1, 55) = 19.47$, $p < .001$, $\eta^2 = 0.26$), with higher SCR rate in the stress condition than in the control condition, $t(55) = 4.41$, $p < .001$, $d = 0.59$, 95% CI = [0.21, 0.56], even after excluding the two participants.

In sum, these results suggest that the acute stress manipulation successfully elicited stress responses.

3.2. Task performance

After confirming that self-reported stress levels and SCR rate were augmented in the stress condition, we examined the impact of acute stress on choice performance during the reinforcement-learning task (Fig. 2c) using a glme model, which accounted for the binomial distribution of the data (correct or incorrect responses) and included as predictor variables condition (stress or control), valence (gains or losses), block number (1 or 2), trial number (1 to 40), and the interaction of interest (condition × valence). We found a significant condition × valence interaction, $\beta = -0.19$, $p = .018$, 95% CI = [-0.34, −0.031] (Fig. 2d), and post-hoc analyses revealed that under stress, relative to the control condition, participants performed significantly worse when seeking monetary gains, $F(1, 19755) = 12.87$, $p < .001$, but not when avoiding losses, $F(1, 19755) = 0.14$, $p = .71$. As performance below chance levels might be indicative of non-compliance with the experimental setting, we also inspected whether each participant's behavioral performance across gains and losses was below chance levels (i.e., less than 50% of correct choices) in the stress or control conditions. One participant performed below chance levels in the stress condition, and two participants performed below chance levels in the control condition. Thus, we repeated the aforementioned analyses excluding these three participants. We found that the condition × valence interaction was still significant after excluding participants that did not learn how to perform the task, $\beta = -0.28$, $p < .001$, 95% CI = [-0.44, −0.12]. We also

confirmed the robustness of our findings using likelihood ratio tests. We found that the full glme model (which assumed different performance towards gains between the stress and control conditions) had a significantly better fit than a model that assumed no differences in performance towards gains between conditions, $\chi^2(1) = 12.80$, $p < .001$, and that the same full model did not have a significantly better fit than a model that assumed no differences in performance towards losses between the stress and control conditions $\chi^2(1) = 0.20$, $p = .66$. These analyses indicate that behavioral performance towards gains, but not losses, significantly differed between the stress and control conditions.

In sum, acute stress selectively impaired choice performance towards monetary gains during the reinforcement-learning task. As an additional check, we confirmed that the choices estimated by the glme model showed close correspondence with the observed choices across trials in both conditions, Pearson's $r > 0.65$, $p < .001$, Spearman's $r > 0.52$, $p < .001$ (Fig. 2e, see Fig. S2 in the Supplementary Material for scatterplots).

Additionally, given that acute stress is thought to increase aberrant spontaneous phasic-dopamine release (Anstrom et al., 2009; Anstrom & Woodward, 2005; Valenti et al., 2011), which, in turn, may lead to augmented learning and behavioral responding for neutral stimuli (Maia & Frank, 2017; Roiser, Howes, Chaddock, Joyce, & McGuire, 2013), we performed an exploratory analysis of participants' choices for the neutral stimuli pairs (the pairs not associated with financial outcomes; for further details see Table S2 and "Analyses of neutral trials" in the "Supplementary Analyses" section of the Supplementary Material). We found that, within the pairs of neutral stimuli, acute stress increased behavioral responding towards the stimuli that more often yielded as outcome a coin with no financial value (the high-probability "look" stimuli) relative to the stimuli that yielded no outcome at all (Fig. S1b in the Supplementary Material). This tentative finding seems consistent with the idea that acute stress might bias behavioral responding for neutral stimuli due to augmented aberrant spontaneous phasic-dopamine release.

3.3. Computational modeling

To further probe the nature of the effects of acute stress on reward and punishment learning, we fitted a biologically inspired reinforcement-learning model (Frank et al., 2007) to participants' trial-by-trial choices (see subsection 2.6.1. in "Materials and Methods" for a full description of the model). The fitted model included separate learning rates for positive (α^+) and negative (α^-) prediction errors, to account for the differential firing of dopaminergic neurons for positive and negative prediction errors (Schultz et al., 1997). The model also included the inverse temperature parameter, β, which controls the decision noise (Daw, 2011; Sutton & Barto, 1998). We then analyzed the best-fitting model parameters using ANOVAs.

These analyses revealed a main effect of condition on learning rates, $F(1, 61) = 4.69$, $p = .034$, $\eta^2 = 0.071$, but a non-significant condition × valence interaction, $F(1, 61) = 2.13$, $p = .15$, $\eta^2 = 0.034$ (Fig. 3a). For completeness, we performed paired t-tests. Despite not formally well-justified (due to the lack of a significant interaction), such paired t-tests yielded results consistent with those from task-performance analyses and those from ensuing computational modeling analyses (cf. analyses that excluded participants who performed below chance levels and analyses of α±*β). Paired t-tests revealed that α^+ was significantly lower in the stress condition ($M = 0.40$, $SEM = 0.033$) than in the control condition ($M = 0.51$, $SEM = 0.038$), $t(61) = -2.25$, $p = .028$, $d = -0.29$, 95% CI = [-0.21, −0.013], while α^- was not significantly different between the stress ($M = 0.25$, $SEM = 0.025$) and control conditions ($M = 0.27$, $SEM = 0.027$), $t(61) = -0.72$, $p = .47$, $d = -0.092$, 95% CI = [-0.098, 0.046] (Fig. 3a). Additionally, as for task-performance analyses, we repeated all statistical analyses excluding the three participants that performed below chance levels. Exclusion of those three participants revealed that the condition × valence interaction reached significance, $F(1, 58) = 4.61$, $p = .036$, $\eta^2 = 0.074$), such that α^+ was significantly
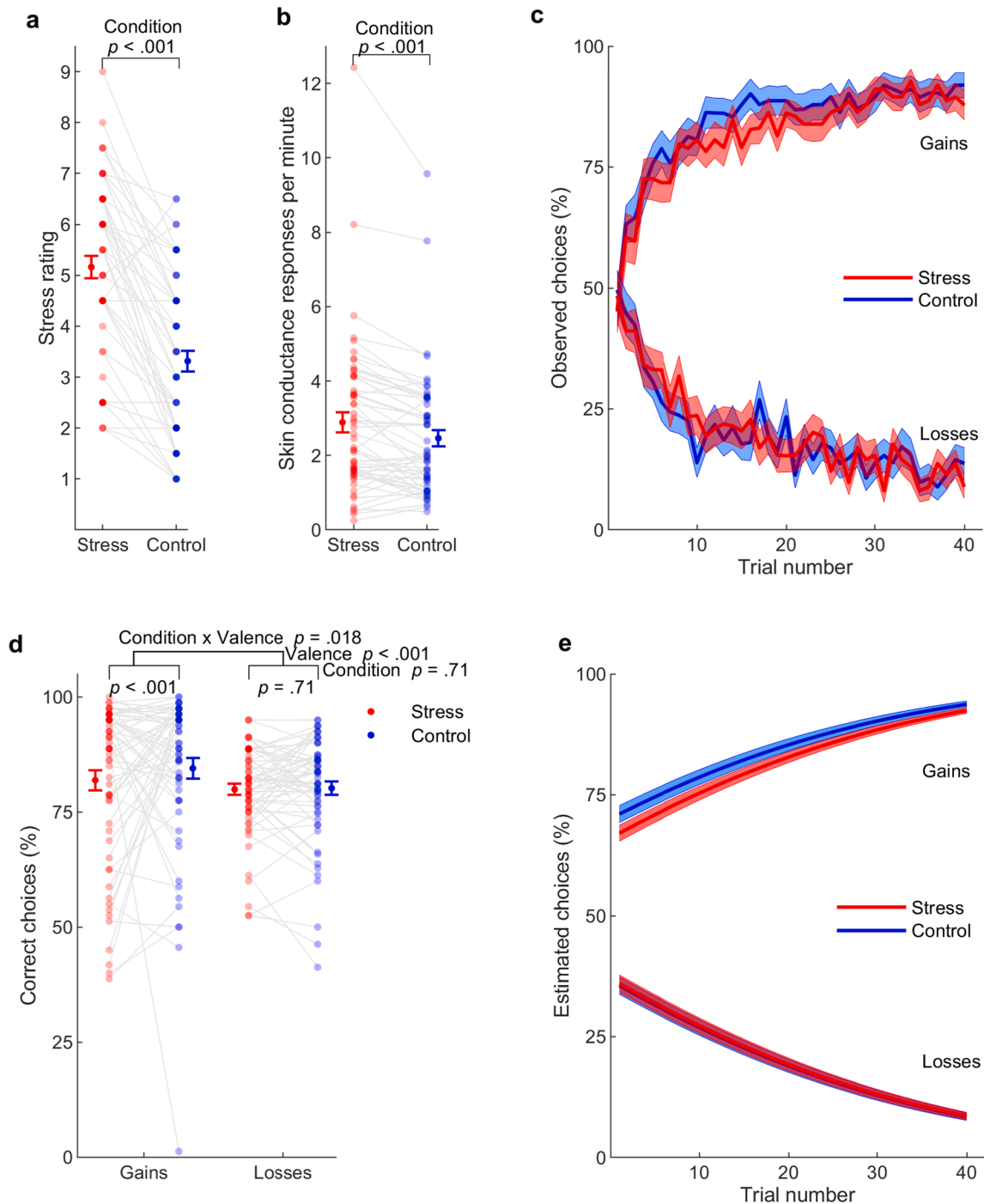
Fig. 2. Manipulation check and task performance. **(a)** Participants (n = 62) reported higher stress levels in the stress condition (red) compared with the control condition (blue). **(b)** Skin conductance response rate (n = 58) was significantly higher in the stress condition than in the control condition. **(c)** Learning curves represent the trial-by-trial percentage of participants (n = 62) who chose the "correct" gain stimulus (associated with a probability of 0.8 of winning 0.5€; upper part of the graph) and the "incorrect" loss stimulus (associated with a probability of 0.8 of losing 0.5€; lower part of the graph), in the stress and control conditions. **(d)** Participants performed significantly worse when seeking monetary gains, but not when avoiding monetary losses, in the stress condition relative to the control condition. The reported *p*-values are from the generalized linear mixed-effects model that included each participant's trial-by-trial choices and respective post-hoc tests. **(e)** The choices estimated by the generalized linear mixed-effects model captured the evolution of the actual observed choices during the reinforcement-learning task (to compare the choices estimated by the model with the actual observed choices, compare the overlap of the curves between the stress and control conditions depicted here with the overlap of the curves depicted in Fig. 2c). In panels a, b, and d, connected dots represent data points from the same participant, and more transparent (opaque) dots represent less (more) overlapping data points; the error bar displayed on the side of the scatter plots indicate the sample mean ± standard error of the mean. In panels c and e, each central line represents the mean and the filled area represents the ± standard error of the mean. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
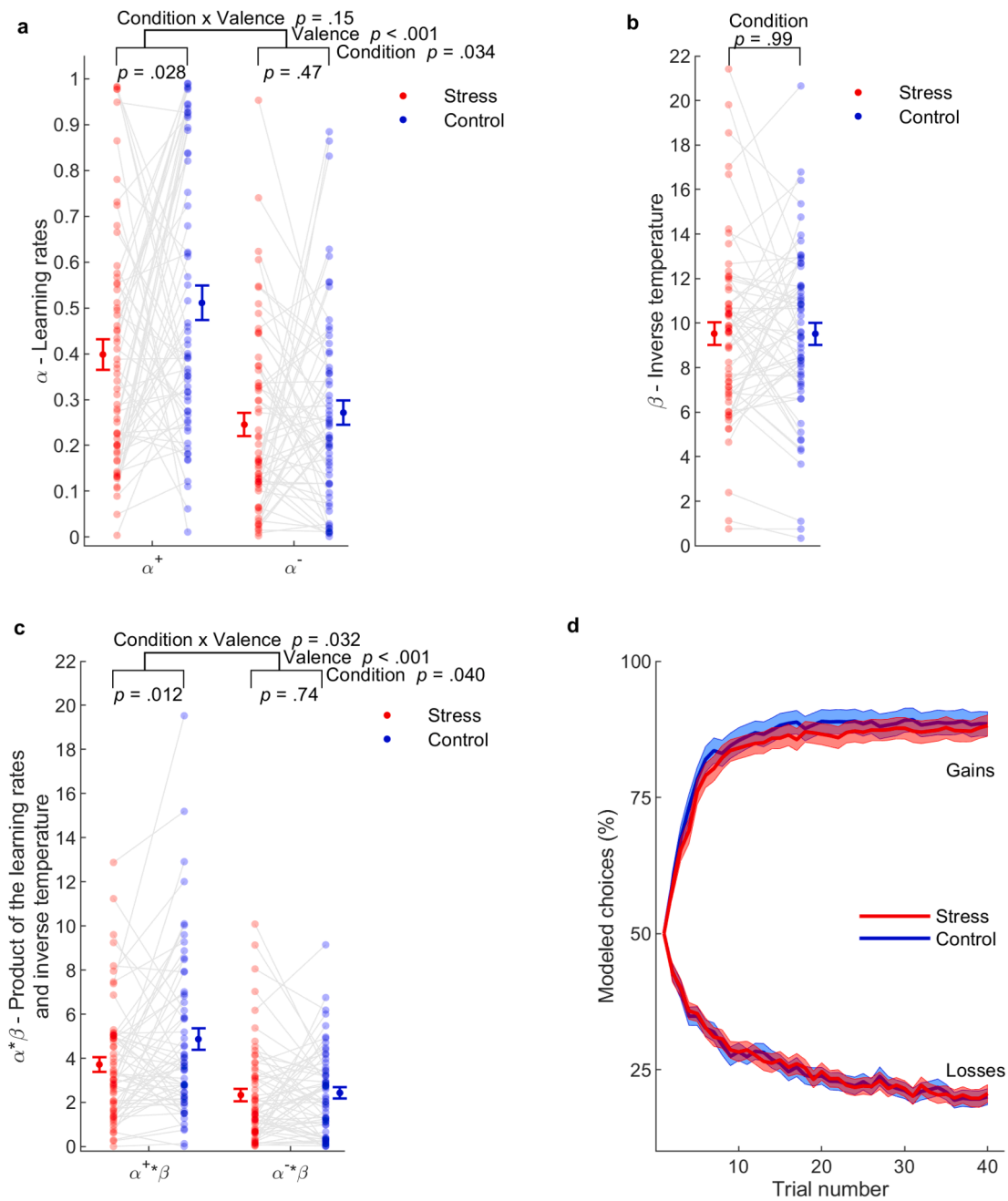
Fig. 3. Model-fitting results. **(a)** Participants (n = 62) had lower learning rates (α^+ and α^-) in the stress condition (red) compared with the control condition (blue). Paired t-tests revealed a significantly lower learning rate for positive prediction errors (α^+) in the stress condition comparatively to the control condition. **(b)** The inverse temperature (β) did not differ between conditions. **(c)** $\alpha^+*\beta$, but not $\alpha^-*\beta$, was significantly reduced in the stress condition compared with the control condition. In panels a, b and c, connected dots represent data points from the same participant, and more transparent (opaque) dots represent less (more) overlapping data points. The error bar displayed on the side of the scatter plots indicate the sample mean ± standard error of the mean. **(d)** The probabilities of choosing the "correct" gain (upper part of the graph) and the "incorrect" loss (lower part of the graph) stimuli estimated by the reinforcement-learning model followed the same pattern of the actual observed choices (compare the overlap of the curves between the stress and control conditions depicted here with the overlap of the curves that represent the actual observed choices depicted in Fig. 2c). Each central line represents the mean and the filled area represents the ± standard error of the mean. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

lower in the stress condition than in the control condition, $t(58) = -2.48$, $p = .016$, $d = -0.32$, 95% CI = [-0.22, −0.024], and α^- did not differ significantly between conditions $t(58) = -0.16$, $p = .88$, $d = -0.021$, 95% CI = [-0.073, 0.063]. Thus, our findings suggest that acute stress decreases α^+.

The parameter β, which controls the amount of noise in choice selection, also did not differ significantly between the stress ($M = 9.52$, $SEM = 0.49$) and control ($M = 9.53$, $SEM = 0.51$) conditions, $t(61) =$

-0.016, $p = .99$, $d = 0.0020$, 95% CI = [-0.98, 0.99] (Fig. 3b).

We further tested whether stress had a differential effect on reward compared with punishment learning by analysing the products between each learning rate ($\alpha^\pm$) and the inverse temperature (β). In reinforcement-learning models, $\alpha^\pm$ and β tend to be inversely coupled (Daw, 2011; Supplementary Material of Maia & Conceição, 2017) because $\alpha^\pm$ multiply by state-action values and the state-action values themselves are multiplied by β to compute choice probabilities (see

equations in subsection 2.6.1). As a result, the parameters viewed separately can have larger estimation errors, while their product tends to be more reliably estimated (Daw, 2011; Schonberg et al., 2007; Waltz et al., 2018; Zhang, Lengersdorff, Mikus, Gläscher, & Lamm, 2020), and thus better recovered. Note that the products $\alpha^{\pm}*\beta$ control how strongly the outcomes impact subsequent choice preferences (Schonberg et al., 2007). Statistical analyses revealed a significant condition × valence interaction, $F(1, 61) = 4.85$, $p = .032$, $\eta^2 = 0.074$, meaning that acute stress significantly decreased $\alpha^{+}*\beta$ (stress: $M = 3.72$, $SEM = 0.34$, control: $M = 4.87$, $SEM = 0.48$), $t(61) = -2.58$, $p = .012$, $d = -0.33$, 95% CI = [-2.04, −0.26]), while not significantly affecting $\alpha^{-}*\beta$ (stress: $M = 2.33$, $SEM = 0.28$, control: $M = 2.44$, $SEM = 0.25$), $t(61) = -0.33$, $p = .74$, $d = -0.042$, 95% CI = [-0.71, 0.51]) (Fig. 3c). Exclusion of the three participants that performed below chance levels did not change the significance of the results for the parameter β, $t(58) = 0.15$, $p = .88$, $d = 0.020$, 95% CI = [-0.94, 1.10], nor for the $\alpha^{\pm}*\beta$ condition × valence interaction, $F(1, 58) = 4.45$, $p = .039$, $\eta^2 = 0.071$. Notably, the previous results were fully replicated when comparing the estimated parameters between the stress and control conditions using non-parametric Wilcoxon signed-rank, rather than parametric, post-hoc tests (α^{+}: $Z = -2.22$, $p = .026$; α^{-}: $Z = -0.17$, $p = .86$; β: $Z = -0.11$, $p = .91$; $\alpha^{+}*\beta$: $Z = -2.19$, $p = .028$; $\alpha^{-}*\beta$: $Z = -0.35$, $p = .73$).

To validate the used reinforcement-learning model, we confirmed that the probability of choices estimated under the reinforcement-learning model had a close correspondence with the actual observed choices across trials in both conditions, Pearson's $r > 0.76$, $p < .001$, Spearman's $r \geq 0.70$, $p < .001$ (Fig. 3d, see Fig. S3 in the Supplementary Material for scatterplots). Next, to further validate that our previous model-fitting results were reliable, we conducted parameter-recovery analyses. Those parameter-recovery analyses demonstrated that the results of our model-fitting procedure were robust both in the stress and control conditions, Pearson's $r > 0.67$, $p < .001$, Spearman's $r > 0.69$, $p < .001$ (see Fig. S4 in the Supplementary Material). Additionally, Bayesian model averaging analyses (Hoeting, Madigan, Raftery, & Volinsky, 1999) using both the aforementioned, neurobiologically inspired reinforcement-learning model (with separate learning rates for positive and negative prediction errors) and a nested, alternative candidate model (with a single learning rate) provided further support for the robustness of the parameters estimated by the model with separate learning rates (see Fig. S5, Table S3 and "Model comparison and Bayesian model averaging" in the "Supplementary Analyses" section of the Supplementary Material).

Taken together, these computational findings indicate that under acute stress participants incorporated positive prediction errors at a lower rate, which seems to explain why acute stress impaired behavioral performance towards monetary gains during the reinforcement-learning task.

4. Discussion

Acute stress is present in day-to-day life, and people recurrently need to make choices and learn from the rewarding or punishing outcomes of those choices whilst under stress. In this study we investigated whether and how acute stress impacted reward and punishment learning in men using a reinforcement-learning framework. Given the putative roles of phasic-dopamine responses on prediction errors signaling and of acute stress on aberrant spontaneous phasic-dopamine release, we hypothesized that acute stress would impair reward learning and, more tentatively punishment learning. Relatedly, we hypothesized that acute stress would decrease the learning rate for positive prediction errors and, more tentatively, the learning rate for negative prediction errors. We found that acute stress impaired behavioral performance towards monetary gains, but not losses, and that this impaired performance could be explained by a decreased learning rate for positive prediction errors.

4.1. Effect of acute stress on reward learning

Our finding that acute stress impaired reward-seeking performance is consistent with several previous studies which found reduced reward responsiveness under acute stress (Bogdan & Pizzagalli, 2006; Bogdan et al., 2011; Morris & Rottenberg, 2015; Paret & Bublatzky, 2020), particularly in high stress-reactive individuals (Berghorst et al., 2013). At a first sight, however, some of the extant literature may seem equivocal, possibly due to critical methodological differences related to stress operationalization (Porcelli & Delgado, 2017). A significant number of studies have investigated the effects of stress on learning using different paradigms, such as the Trier social stress test (Boyle, Stanton, Eisenberger, Seeman, & Bower, 2019; Kruse et al., 2018; Petzold et al., 2010; Radenbach et al., 2015), or the (socially evaluated) cold pressor test (Byrne et al., 2019; Ehlers & Todd, 2017; Glienke, Wolf, & Bellebaum, 2015; Lighthall et al., 2013; Otto et al., 2013; Paul et al., 2019), in which acute stress is induced before the learning task. In these paradigms, stress induction precedes any learning processes, thus the stress-induced emotional state may be less concurrent with the cognitive processes that operate during the task. As such, it is unclear whether these paradigms probe the effects of acute stress or of recovery from stress on learning (Hermans et al., 2014). In contrast, in our study, we induced stress during the learning task, and we used relatively short task blocks to avoid the slow cortisol response from peaking and subsequent recovery within blocks. The timing of the stress induction relative to the task is an important factor to take into account when comparing studies, as early-fast catecholamine release and delayed-slow glucocorticoids actions might account for different effects on brain function (de Kloet et al., 2005; Hermans et al., 2014; Joëls & Baram, 2009). Critically, studies that induce stress before the learning task may suggest divergent behavioral results from ours (Byrne et al., 2019; Lighthall et al., 2013), whereas studies that have induced stress during the learning task (as ours) seem to be in agreement with our findings of impaired reward learning during acute-stress exposure (Berghorst et al., 2013; Bogdan & Pizzagalli, 2006; Bogdan et al., 2011; Morris & Rottenberg, 2015; Paret & Bublatzky, 2020). Therefore, cross-study comparisons suggest that individuals may perform more poorly when learning to maximize rewards whilst exposed to an acute stressor.

In this study we further inspected the computational mechanisms behind impaired reward learning under acute stress. Using a reinforcement-learning model with separate learning rates for positive and negative prediction errors (Frank et al., 2007), we found that acute stress reduced the rate at which participants learned from positive prediction errors. A reduced learning rate for positive predictions errors means that under acute stress individuals learned more slowly about unexpected rewards and therefore took longer to adapt their behavior on the basis of the more recent rewarding outcomes of their choices. Other studies have successfully applied this reinforcement-learning model to analyze data from distinct tasks to describe the genetic (Doll et al., 2011; Frank et al., 2007) and neural (Diederen et al., 2016; Lefebvre et al., 2017; Niv, Edlund, Dayan, & Doherty, 2012) correlates of cognition and behavior. However, to the best of our knowledge, no studies had attempted to model how acute stress affects cognition, particularly reward and punishment learning, using such model. Another reinforcement-learning model commonly used in the stress literature includes distinct learning rates for gain and loss trials regardless of the valence of the prediction error (e.g., Aylward et al., 2019; Robinson et al., 2013; Treadway et al., 2017), meaning that learning from rewards in gain trials (or punishments in loss trials) is modeled the same way as learning from omission of rewards in gain trials (or omission of punishments in loss trials, respectively), which may be misleading and not be so well supported by neurobiological evidence as a model that assumes distinct learning rates based on the valence of the prediction error (as the one we used) (Maia & Frank, 2011; O'Doherty et al., 2003; Schultz et al., 1997). In simple reinforcement-learning tasks, such as ours, reward learning seems to mostly occur via positive prediction

errors (Lefebvre et al., 2017). But, given the structure of our reinforcement-learning model, stress-induced impaired reward learning could have conceivably been underpinned by a higher learning rate for negative prediction errors — which could have led to unlearning of the correct stimulus — and/or by lower inverse temperature parameter (β) — which could have been reflected in choices more at random. The fact that our data indicated that the learning rate for negative prediction errors was not higher in the stress condition and that the inverse temperature parameter did not differ between conditions, but that the learning rate for positive prediction errors was decreased in the stress condition, suggests that acute stress specifically impaired reward learning via disruption of positive prediction errors. Thus, our findings contribute to a better mechanistic understanding of how acute stress may impact reward learning.

Given that the (quantifiable) parameters from the reinforcement-learning model that we used seem to reflect specific dopaminergic-related neural mechanisms (e.g., Frank et al., 2007; Frank & O'Reilly, 2006) our computational findings may shed light on putative neural mechanisms underlying the impact of acute stress on cognition and behavior. Specifically, our computational findings seem broadly consistent with the proposed neurobiological account of dopaminergic neurons functioning under acute stress. Acute stress is thought to induce aberrant spontaneous dopamine release (Anstrom et al., 2009; Anstrom & Woodward, 2005; Cabib & Puglisi-Allegra, 2012; Valenti et al., 2011); stress-induced spontaneous dopamine release, in turn, may disrupt the adaptive striatal phasic-burst dopamine responses that signal positive prediction errors (Daberkow et al., 2013; Grace, 2016; Maia & Frank, 2017; Werlen et al., 2020), which would explain the stress-induced impairment of reward learning. While our computational results suggest that acute stress disrupts striatal responses to prediction errors during reward learning (see also Huys et al., 2013), we cannot disregard, however, that acute stress may also affect reward learning via prefrontal cortex disturbances (Arnsten, 2009; Lighthall et al., 2013; Otto et al., 2013).

4.2. Effect of acute stress on punishment learning

Acute stress impaired reward learning, but we found no evidence for an effect of acute stress on punishment learning. Computational-modeling analyses did not provide evidence that acute stress affected the learning rate for negative prediction errors to the same extent that it affected the learning rate for positive prediction errors, which might explain why behavioral performance during loss trials — which might rely more on negative prediction errors compared with performance during gain trials (Palminteri & Pessiglione, 2013) — was not impaired under acute stress.

According to a long standing influential loss aversion framework (Kahneman & Tversky, 1979), losses can have more debilitative potential than gains; therefore, as an adaptive strategy, it is possible that individuals may be more attuned to losses than to gains (Lejarraga, Hertwig, & Gonzalez, 2012; Yechiam, 2019; Yechiam & Hochman, 2013), explaining why punishment learning was spared under acute stress. However, recent evidence suggests that individuals learn gain associations better than loss associations in reinforcement-learning tasks despite symmetrical task structure and symmetrical outcome probabilities (Lin, Cabrera-Haro, & Reuter-Lorenz, 2020) (as in our task). Indeed, our behavioral data indicated that participants performed better when learning to seek gains than when learning to avoid losses in both the stress and control conditions. In addition, our computational data indicated that the learning rate for positive prediction errors was significantly higher than the learning rate for negative prediction errors, which might suggest underestimation of negative outcomes relative to positive outcomes (see also Lefebvre et al., 2017). Thus, our data suggest that it is unlikely that acute stress spared the mechanisms involved in punishment learning, but not in reward learning, due to heightened attention towards losses relative to gains.

One potential neurocognitive explanation for such lack of effect of acute stress on loss avoidance is that D2 dopamine receptors, which mediate punishment learning (Frank & O'Reilly, 2006; Maia & Frank, 2011), are already mostly activated at baseline dopamine levels, so their activation might be affected by decreases, but less so by increases, in dopamine levels (Maia & Conceição, 2017; Möller & Bogacz, 2019). Finally, non-dopaminergic mechanisms may also be involved in punishment learning (Boureau & Dayan, 2011; Moran et al., 2018), which may partially explain why previous studies using the same reinforcement-learning task also did not find significant effects of pharmacological manipulations of the dopaminergic system on punishment learning (Eisenegger et al., 2014; Pessiglione et al., 2006).

4.3. Acute stress and behavioral performance in neutral trials

We found preliminary evidence that acute stress biased behavioral responding to neutral stimuli, as the choice of the high-probability "look" stimuli — compared to the stimuli with a high probability of yielding "nothing" — was augmented under stress compared to the control condition (for a detailed discussion about this finding, see the Supplementary Material). Although very tentative, this finding might be of relevance as excessive, aberrant spontaneous dopamine release is thought to underlie increased behavioral responding and aberrant learning for neutral stimuli (Maia & Frank, 2017; Roiser et al., 2013). Still, whether acute stress promotes aberrant valuation of neutral stimuli via dopaminergic disturbances remains unknown and should be further investigated.

4.4. Limitations

In this study, we induced acute stress in participants, using a repetitive and uncontrollable sound, whilst they completed a reinforcement-learning task. We exposed participants to the sound during the task to ensure that acute stress was contingent on the learning processes, but at the expense of possibly confounding the induction of stress with distraction. To minimize this potential confound, we used a sound that was always constant and repetitive, as unpredictable changes in sound sequences induce distraction more robustly (Hughes, 2014; Parmentier, 2014; Parmentier et al., 2008; Sussman et al., 2003). Moreover, and also aiming to reduce the unpredictability of the sound, before starting the task, participants were instructed that they would hear a sound during blocks cued with a warning signal and a red background (Hughes, 2014; Parmentier & Hebrero, 2013; Sussman et al., 2003). Besides the experimental precautions we took to avoid that the stressor acted as a distractor, our data indicated a selective effect of the stress manipulation on behavioral performance during gain trials. If the sound was acting as a distractor, rather than as a stressor, it would be more likely to find instead a general behavioral impairment; that is, worse performance towards both gains and losses, rather than a selective effect on gains. Relatedly, if the manipulation had acted mostly as a distractor, it would be expected that participants behaved more at random. For example, previous data indicate that when participants are asked to generate novel combinations of numbers whilst exposed to an auditory distractor, they generate combinations more at random than participants who are not exposed to the auditory distractor, suggesting that enhanced distraction during a cognitive task may enhance the ability to generate randomness (Tsushima & Nakayama, 2010). Thus, if participants were distracted by our auditory stressor, we would have expected them to make choices more at random, which would likely be captured by a reduction of the inverse temperature parameter — reflecting increased random behavior — in the stress condition. Still, although our data and experimental design make it unlikely that the observed effect on reward learning was due to distractibility rather than stress, additional control conditions should be added in future studies to experimentally disentangle distraction from stress.

Finally, only men were included in this study, to avoid the potential

confounding effects of menstrual-cycle-dependent variation on stress responsivity (Ossewaarde et al., 2010), as well as on reward and punishment learning (Diekhof et al., 2020; Dreher et al., 2007). Our finding that acute stress disrupts reward learning in men seems to be in line with previous reports showing that acute stress disrupts reward-seeking behavior in women (Berghorst, Bogdan, Frank, & Pizzagalli, 2013; Bogdan, Perlis, Fagerness, & Pizzagalli, 2010; Bogdan, Santesso, Fagerness, Perlis, & Pizzagalli, 2011; Bogdan & Pizzagalli, 2006; Morris & Rottenberg, 2015; Paret & Bublatzky, 2020), but further studies are needed to assess whether acute stress has the same computational effects on reward and punishment learning in men and women.

5. Conclusions

We present evidence that acute stress reduces how quickly male adults integrate the unexpected rewarding outcomes of their choices over time. Our results are consistent with a neurobiological framework of stress-induced dopaminergic disturbances and can thus contribute to a better understanding of the computational mechanisms that underlie the deleterious impact of acute stress on reward learning. Ultimately, this study might offer key mechanistic insights into the impact of acute stress in everyday life.

Author contributions

J. Carvalheiro, A. Mesquita and A. Seara-Cardoso developed the study concept with critical input from V. Conceição. Testing and data collection were performed by J. Carvalheiro under the supervision of A. Mesquita and A. Seara-Cardoso. Data analyses and interpretation was carried out by J. Carvalheiro and V. Conceição. J. Carvalheiro drafted the manuscript. All authors revised and edited the manuscript.

Declaration of Competing Interest

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

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi. org/10.1016/j.bandc.2020.105657.

References

Abercrombie, E. D., Keefe, K. A., DiFrischia, D. S., & Zigmond, M. J. (1989). Differential effect of stress on in vivo dopamine release in striatum, nucleus accumbens, and medial frontal cortex. *Journal of Neurochemistry, 52*, 1655–1658.

Anstrom, K. K., Miczek, K. A., & Budygin, E. A. (2009). Increased phasic dopamine signaling in the mesolimbic pathway during social defeat in rats. *Neuroscience, 161*, 3–12.

Anstrom, K. K., & Woodward, D. J. (2005). Restraint increases dopaminergic burst firing in awake rats. *Neuropsychopharmacology, 30*, 1832–1840.

Arguelles, A. E., Ibeas, D., Ottone, J. P., & Chekherdemian, M. (1962). Pituitary-adrenal stimulation by sound of different frequencies. *The Journal of Clinical Endocrinology and Metabolism, 22*, 846–852.

Arnsten, A. F. T. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience, 10*, 410–422.

Arnsten, A. F. T., & Goldman-Rakic, P. S. (1998). Noise stress impairs prefrontal cortical cognitive function in monkeys: evidence for a hyperdopaminergic mechanism. *Archives of General Psychiatry, 55*, 362–368.

Aylward, J., Valton, V., Ahn, W.-Y., Bond, R. L., Dayan, P., Roiser, J. P., & Robinson, O. J. (2019). Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nature Human Behaviour, 3*, 1116–1123.

Belujon, P., Grace, A. A., & Grace, A. A. (2015). Regulation of dopamine system responsivity and its adaptive and pathological response to stress. *Proceedings of the Royal Society B: Biological Sciences, 282*, 20142516.

Bogdan, R., Perlis, R. H., Fagerness, J., & Pizzagalli, D. A. (2010). The impact of mineralocorticoid receptor ISO/VAL genotype (rs5522) and stress on reward learning. *Genes, Brain and Behavior, 9*, 658–667.

Berghorst, L. H., Bogdan, R., Frank, M. J., & Pizzagalli, D. A. (2013). Acute stress selectively reduces reward sensitivity. *Frontiers in Human Neuroscience, 7*, 1–15.

Bogdan, R., & Pizzagalli, D. A. (2006). Acute stress reduces reward responsiveness: Implications for depression. *Biological Psychiatry, 60*, 1147–1154.

Bogdan, R., Santesso, D. L., Fagerness, J., Perlis, R. H., & Pizzagalli, D. A. (2011). Corticotropin-releasing hormone receptor type 1 (CRHR1) genetic variation and stress interact to influence reward learning. *Journal of Neuroscience, 31*, 13246–13254.

Boureau, Y.-L., & Dayan, P. (2011). Opponency revisited: Competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology, 36*, 74–97.

Boyle, C. C., Stanton, A. L., Eisenberger, N. I., Seeman, T. E., & Bower, J. E. (2019). Effects of stress-induced inflammation on reward processing in healthy young women. *Brain Behavior and Immunity, 83*, 126–134.

Breier, A., Albus, M., Pickar, D., Zahn, T. P., Wolkowitz, O. M., & Paul, S. M. (1987). Controllable and uncontrollable stress in humans: Alterations in mood and neuroendocrine and psychophysiological function. *The American Journal of Psychiatry, 144*, 1419–1425.

Byrne, K. A., Cornwall, A. C., & Worthy, D. A. (2019). Acute stress improves long-term reward maximization in decision-making under uncertainty. *Brain and Cognition, 133*, 84–93.

Cabib, S., & Puglisi-Allegra, S. (2012). The mesoaccumbens dopamine in coping with stress. *Neuroscience and Biobehavioral Reviews, 36*, 79–89.

Chater, N. (2009). Rational and mechanistic perspectives on reinforcement learning. *Cognition, 113*, 350–364.

Collins, A. G. E., & Frank, M. J. (2013). Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychological Review, 121*, 337–366.

Daberkow, D. P., Brown, H. D., Bunner, K. D., Kraniotis, S. A., Doellman, M. A., Ragozzino, M. E., … Roitman, M. F. (2013). Amphetamine paradoxically augments exocytotic dopamine release and phasic dopamine signals. *Journal of Neuroscience, 33*, 452–463.

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. *Decision Making, Affect, and Learning: Attention and Performance XXIII*, 1–26.

Daw, N. D., & Frank, M. J. (2009). Reinforcement learning and higher level cognition: Introduction to special issue. *Cognition, 113*, 259–261.

Daw, N. D., & Tobler, P. N. (2014). Value learning through reinforcement: The basics of dopamine and reinforcement learning. In P. W. Glimcher, & E. Fehr (Eds.), *Neuroeconomics: Decision making and the brain* (pp. 283–298). Amsterdam: Elsevier.

de Berker, A. O., Tirole, M., Rutledge, R. B., Cross, G. F., Dolan, R. J., & Bestmann, S. (2016). Acute stress selectively impairs learning to act. *Scientific Reports, 6*, 29816.

de Kloet, E. R., Joëls, M., & Holsboer, F. (2005). Stress and the brain: From adaptation to disease. *Nature Reviews Neuroscience, 6*, 463–475.

Diederen, K. M. J., Spencer, T., Vestergaard, M. D., Fletcher, P. C., Diederen, K. M. J., Spencer, T., … Schultz, W. (2016). Adaptive prediction error coding in the human midbrain and striatum facilitates behavioral adaptation and learning efficiency. *Neuron, 90*, 1127–1138.

Diederen, K. M. J., Ziauddeen, H., Vestergaard, M. D., Spencer, T., Schultz, W., & Fletcher, P. C. (2017). Dopamine modulates adaptive prediction error coding in the human midbrain and striatum. *Journal of Neuroscience, 37*, 1708–1720.

Diekhof, E. K., Korf, S., Ott, F., Schädlich, C., & Holtfrerich, S. K. C. (2020). Avoidance learning across the menstrual cycle: A conceptual replication. *Frontiers in Endocrinology, 11*, 1–12.

Doll, B. B., Hutchison, K. E., & Frank, M. J. (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *Journal of Neuroscience, 31*, 6188–6198.

Dreher, J., Schmidt, P. J., Kohn, P., Furman, D., Rubinow, D., & Berman, K. F. (2007). Menstrual cycle phase modulates reward-related neural function in women. *Proceedings of the National Academy of Sciences, 104*, 2465–2470.

Ehlers, M. R., & Todd, R. M. (2017). Acute psychophysiological stress impairs human associative learning. *Neurobiology of Learning and Memory, 145*, 84–93.

Eisenegger, C., Naef, M., Linssen, A., Clark, L., Gandamaneni, P. K., Müller, U., & Robbins, T. W. (2014). Role of dopamine D2 receptors in human reinforcement learning. *Neuropsychopharmacology, 39*, 2366–2375.

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience, 17*, 51–72.

Frank, M. J. (2015). Linking across levels of computation in model-based cognitive neuroscience. In *An introduction to model-based cognitive neuroscience* (pp. 159–177). New York, NY, US: Springer.

Frank, M. J., & Fossella, J. A. (2011). Neurogenetics and pharmacology of learning, motivation, and cognition. *Neuropsychopharmacology, 36*, 133–152.

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences, 104*, 16311–16316.

Frank, M. J., & O'Reilly, R. C. (2006). A mechanistic account of striatal dopamine function in human cognition: Psychopharmacological studies with cabergoline and haloperidol. *Behavioral Neuroscience, 120*, 497–517.

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in Parkinsonism. *Science, 306*, 1940–1943.

Glass, D. C., Reim, B., & Singer, J. E. (1971). Behavioral consequences of adaptation to controllable and uncontrollable noise. *Journal of Experimental Social Psychology, 7*, 244–257.

Glienke, K., Wolf, O. T., & Bellebaum, C. (2015). The impact of stress on feedback and error processing during behavioral adaptation. *Neuropsychologia, 71*, 181–190.

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences, 108*, 15647–15654.

Grace, A. A. (2016). Dysregulation of the dopamine system in the pathophysiology of schizophrenia and depression. *Nature Reviews Neuroscience, 17*, 524–532.

Grogan, J. P., Tsivos, D., Smith, L., Knight, B. E., Bogacz, R., Whone, A., & Coulthard, E. J. (2017). Effects of dopamine on reinforcement learning and consolidation in Parkinson's disease. *ELife, 6*, 1–23.

Hall, S., Aisbett, B., Tait, J., Turner, A., Ferguson, S., & Main, L. (2016). The acute physiological stress response to an emergency alarm and mobilization during the day and at night. *Noise and Health, 18*, 150–156.

Hermans, E. J., Henckens, M. J. A. G., Joëls, M., & Fernández, G. (2014). Dynamic adaptation of large-scale brain networks in response to acute stressors. *Trends in Neurosciences, 37*, 304–314.

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian averaging models. *Statistical Science, 14*, 382–417.

Hughes, R. W. (2014). Auditory distraction: A duplex-mechanism account. *PsyCh Journal, 3*, 30–41.

Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience, 19*, 404–413.

Huys, Q. J. M., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. *Biology of Mood & Anxiety Disorders, 3*, 1–16.

Joëls, M., & Baram, T. Z. (2009). The neuro-symphony of stress. *Nature Reviews Neuroscience, 10*, 459–466.

Joëls, M., Pu, Z., Wiegert, O., Oitzl, M. S., & Krugers, H. J. (2006). Learning under stress: How does it work? *Trends in Cognitive Sciences, 10*, 152–158.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.

Kim, K. H., Bang, S. W., & Kim, S. R. (2004). Emotion recognition system using short-term monitoring of physiological signals. *Medical & Biological Engineering & Computing, 42*, 419–427.

Korhonen, O. (1981). Firefighters' reaction to alarm, an ECG and heart rate study. *Journal of Occupational Medicine, 23*, 762–766.

Kruse, O., Tapia León, I., Stalder, T., Stark, R., & Klucken, T. (2018). Altered reward learning and hippocampal connectivity following psychosocial stress. *NeuroImage, 171*, 15–25.

Lefebvre, G., Lebreton, M., Meyniel, F., Bourgeois-Gironde, S., & Palminteri, S. (2017). Behavioural and neural characterization of optimistic reinforcement learning. *Nature Human. Behaviour, 1*.

Lejarraga, T., Hertwig, R., & Gonzalez, C. (2012). How choice ecology influences search in decisions from experience. *Cognition, 124*, 334–342.

Lighthall, N. R., Gorlick, M. A., Schoeke, A., Frank, M. J., & Mather, M. (2013). Stress modulates reinforcement learning in younger and older adults. *Psychology and Aging, 28*, 35–46.

Lin, Z., Cabrera-Haro, L. E., & Reuter-Lorenz, P. A. (2020). Asymmetrical learning and memory for acquired gain versus loss associations. Cognition, 202.

Luksys, G., & Sandi, C. (2011). Neural mechanisms and computations underlying stress effects on learning and memory. *Current Opinion in Neurobiology, 21*, 502–508.

Maia, T. V., & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature Neuroscience, 14*, 154–162.

Maia, T. V., & Frank, M. J. (2017). An integrative perspective on the role of dopamine in schizophrenia. *Biological Psychiatry, 81*, 52–66.

Maia, T. V. (2015). Introduction to the series on computational psychiatry. *Clinical Psychological Science, 3*, 374–377.

Maia, T. V., & Conceição, V. A. (2017). The roles of phasic and tonic dopamine in tic learning and expression. *Biological Psychiatry, 82*, 401–412.

McRae, A. L., Saladin, M. E., Brady, K. T., Upadhyaya, H., Back, S. E., & Timmerman, M. A. (2006). Stress reactivity: Biological and subjective responses to the cold pressor and Trier Social stressors. *Human Psychopharmacology, 21*, 377–385.

Möller, M., & Bogacz, R. (2019). Learning the payoffs and costs of actions. *PLOS Computational Biology, 15*.

Moran, R. J., Kishida, K. T., Lohrenz, T., Saez, I., Laxton, A. W., Witcher, M. R., … Montague, P. R. (2018). The protective action encoding of serotonin transients in the human brain. *Neuropsychopharmacology, 43*, 1425–1435.

Morris, B. H., & Rottenberg, J. (2015). Heightened reward learning under stress in generalized anxiety disorder: A predictor of depression resistance? *Journal of Abnormal Psychology, 124*, 115–127.

Nagano-Saito, A., Dagher, A., Booij, L., Gravel, P., Welfeld, K., Casey, K. F., … Benkelfat, C. (2013). Stress-induced dopamine release in human medial prefrontal cortex-18F-fallypride/PET study in healthy volunteers. *Synapse, 67*, 821–830.

Nair, A., Rutledge, R. B., & Mason, L. (2020). Under the hood: Using computational psychiatry to make psychological therapies more mechanism-focused. *Frontiers in Psychiatry, 11*, 140.

Niv, Y., Edlund, J. A., Dayan, P., & Doherty, J. P. O. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *Journal of Neuroscience, 32*, 551–562.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron, 38*, 329–337.

Ossewaarde, L., Hermans, E. J., van Wingen, G. A., Kooijman, S. C., Johansson, I.-M., Bäckström, T., & Fernández, G. (2010). Neural mechanisms underlying changes in stress-sensitivity across the menstrual cycle. *Psychoneuroendocrinology, 35*, 47–55.

Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences, 110*, 20941–20946.

Palminteri, S., Khamassi, M., Joffily, M., & Coricelli, G. (2015). Contextual modulation of value signals in reward and punishment learning. *Nature Communications, 6*, 8096.

Palminteri, S., & Pessiglione, M. (2013). *Reinforcement learning and Tourette syndrome, in International Review of Neurobiology* (pp. 131–153). Elsevier Inc..

Paret, C., & Bublatzky, F. (2020). Threat rapidly disrupts reward reversal learning. *Behaviour Research and Therapy, 131*, Article 103636.

Parmentier, F. B. R. (2014). The cognitive determinants of behavioral distraction by deviant auditory stimuli: A review. *Psychological Research, 78*, 321–338.

Parmentier, F. B. R., Elford, G., Escera, C., Andrés, P., & Miguel, I. S. (2008). The cognitive locus of distraction by acoustic novelty in the cross-modal oddball task. *Cognition, 106*, 408–432.

Parmentier, F. B. R., & Hebrero, M. (2013). Cognitive control of involuntary distraction by deviant sounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 1635–1641.

Paul, M., Bellebaum, C., Ghio, M., Suchan, B., & Wolf, O. T. (2019). Stress effects on learning and feedback-related neural activity depend on feedback delay. *Psychophysiology, 57*.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature, 442*, 1042–1045.

Petzold, A., Plessow, F., Goschke, T., & Kirschbaum, C. (2010). Stress reduces use of negative feedback in a feedback-based learning task. *Behavioral Neuroscience, 124*, 248–255.

Porcelli, A. J., & Delgado, M. R. (2017). Stress and decision making: Effects on valuation, learning, and risk-taking. *Current Opinion in Behavioral Sciences, 14*, 33–39.

Pruessner, J. C., Champagne, F., Meaney, M. J., & Dagher, A. (2004). Dopamine release in response to a psychological stress in humans and its relationship to early life maternal care: A positron emission tomography study using [11C]raclopride. *Journal of Neuroscience, 24*, 2825–2831.

Radenbach, C., Reiter, A. M. F., Engert, V., Sjoerds, Z., Villringer, A., Heinze, H. J., … Schlagenhauf, F. (2015). The interaction of acute and chronic stress impairs model-based behavioral control. *Psychoneuroendocrinology, 53*, 268–280.

Robinson, O. J., Overstreet, C., Charney, D. R., Vytal, K., & Grillon, C. (2013). Stress increases aversive prediction error signal in the ventral striatum. *Proceedings of the National Academy of Sciences, 110*, 4129–4133.

Robinson, O. J., Vytal, K., Cornwell, B. R., & Grillon, C. (2013). The impact of anxiety upon cognition: Perspectives from human threat of shock studies. *Frontiers in Human Neuroscience, 7*, 203.

Roiser, J. P., Howes, O. D., Chaddock, C. A., Joyce, E. M., & McGuire, P. (2013). Neural and behavioral correlates of aberrant salience in individuals at risk for psychosis. *Schizophrenia Bulletin, 39*, 1328–1336.

Rutledge, R. B., Lazzaro, S. C., Lau, B., Myers, C. E., Gluck, M. A., & Glimcher, P. W. (2009). Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. *Journal of Neuroscience, 29*, 15104–15114.

Rylander, R. (2004). Physiological aspects of noise-induced stress and annoyance. *Journal of Sound and Vibration, 277*, 471–478.

Sapolsky, R. M., Romero, L. M., & Munck, A. U. (2000). How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Reviews, 21*, 55–89.

Schonberg, T., Daw, N. D., Joel, D., O'Doherty, J. P., Doherty, J. P. O., Schonberg, T., … O'Doherty, J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *Journal of Neuroscience, 27*, 12860–12867.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*, 1593–1599.

Schwabe, L., Haddad, L., & Schachinger, H. (2008). HPA axis activation by a socially evaluated cold-pressor test. *Psychoneuroendocrinology, 33*, 890–895.

Schwabe, L., Joëls, M., Roozendaal, B., Wolf, O. T., & Oitzl, M. S. (2012). Stress effects on memory: An update and integration. *Neuroscience and Biobehavioral Reviews, 36*, 1740–1749.

Schwabe, L., & Wolf, O. T. (2009). Stress prompts habit behavior in humans. *Journal of Neuroscience, 29*, 7191–7198.

Sulzer, D., Cragg, S. J., & Rice, M. E. (2016). Striatal dopamine neurotransmission: Regulation of release and uptake. *Basal Ganglia, 6*, 123–148.

Sussman, E., Winkler, I., & Schröger, E. (2003). Top-down control over involuntary attention switching in the auditory modality. *Psychonomic Bulletin and Review, 10*, 630–637.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Treadway, M. T., Admon, R., Arulpragasam, A. R., Mehta, M., Douglas, S., Vitaliano, G., … Pizzagalli, D. A. (2017). Association between interleukin-6 and striatal prediction-error signals following acute stress in healthy female participants. *Biological Psychiatry, 82*, 1–8.

Tsushima, Y., & Nakayama, K. (2010). Does an auditory distractor allow humans to behave more randomly? *Journal of Vision, 10*, 1028.

Valenti, O., Lodge, D. J., & Grace, A. A. (2011). Aversive stimuli alter ventral tegmental area dopamine neuron activity via a common action in the ventral hippocampus. *Journal of Neuroscience, 31*, 4280–4289.

Waltz, J. A., Xu, Z., Brown, E. C., Ruiz, R. R., Frank, M. J., & Gold, J. M. (2018). Motivational deficits in schizophrenia are associated with reduced differentiation between gain and loss-avoidance feedback in the striatum. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 3*, 239–247.

Watabe-Uchida, M., Eshel, N., & Uchida, N. (2017). Neural circuitry of reward prediction error. *Annual Review of Neuroscience, 40*, 373–394.

Werlen, E., Shin, S. L., Gastambide, F., Francois, J., Tricklebank, M. D., Marston, H. M., … Walton, M. E. (2020). Amphetamine disrupts haemodynamic correlates of prediction errors in nucleus accumbens and orbitofrontal cortex. *Neuropsychopharmacology, 45*, 793–803.

Westman, J. C., & Walters, J. R. (1981). Noise and stress: A comprehensive approach. *Environmental Health Perspectives, 41*, 291–309.

Yechiam, E. (2019). Acceptable losses: The debatable origins of loss aversion. *Psychological Research, 83*, 1327–1339.

Yechiam, E., & Hochman, G. (2013). Losses as modulators of attention: Review and analysis of the unique effects of losses over gains. *Psychological Bulletin, 139*, 497–518.

Zhang, L., Lengersdorff, L., Mikus, N., Gläscher, J., & Lamm, C. (2020). Using reinforcement learning models in social neuroscience: Frameworks, pitfalls and suggestions of best practices. *Social Cognitive and Affective Neuroscience, nsaa089*.