Stress Elevates Frontal Midline Theta in Feedback-based Category Learning of Exceptions

Marcus Paul[1], Marie-Christin Fellner[1], Gerd T. Waldhauser[1], John Paul Minda[2], Nikolai Axmacher[1], Boris Suchan[1], and Oliver T. Wolf[1]

Abstract

■ Adapting behavior based on category knowledge is a fundamental cognitive function, which can be achieved via different learning strategies relying on different systems in the brain. Whereas the learning of typical category members has been linked to implicit, prototype abstraction learning, which relies predominantly on prefrontal areas, the learning of exceptions is associated with explicit, exemplar-based learning, which has been linked to the hippocampus. Stress is known to foster implicit learning strategies at the expense of explicit learning. Procedural, prefrontal learning and cognitive control processes are reflected in frontal midline theta (4–8 Hz) oscillations during feedback processing. In the current study, we examined the effect of acute stress on feedback-based category learning of typical category members and exceptions and the oscillatory correlates of feedback processing in the EEG. A computational modeling procedure was applied to estimate the use of abstraction and exemplar strategies during category learning. We tested healthy, male participants who underwent either the socially evaluated cold pressor test or a nonstressful control procedure before they learned to categorize typical members and exceptions based on feedback. The groups did not differ significantly in their categorization accuracy or use of categorization strategies. In the EEG, however, stressed participants revealed elevated theta power specifically during the learning of exceptions, whereas the theta power during the learning of typical members did not differ between the groups. Elevated frontal theta power may reflect an increased involvement of medial prefrontal areas in the learning of exceptions under stress. ■

INTRODUCTION

Learning to categorize objects, events, and people to discrete classes is a fundamental cognitive ability, which enables humans to make quick decisions in rewarding or threatening situations. Physical and psychosocial threats elicit stress responses, which are associated with endocrine changes in the body, such as the release of glucocorticoids. Stress has been shown to modulate feedback processing and learning processes (Glienke, Wolf, & Bellebaum, 2015; Lighthall, Gorlick, Schoeke, Frank, & Mather, 2013; Cavanagh, Frank, & Allen, 2011; Ossewaarde et al., 2011).

Different learning strategies have been identified in feedback-based category learning, which are associated with separate systems in the brain. Prototype-based procedural learning strategies (Reed, 1972; Posner & Keele, 1968), which rely on regions in PFC (Pan & Sakagami, 2012) and the striatal learning system (Cincotta & Seger, 2007; Ashby & Ennis, 2006), are distinguished from exemplar-based declarative learning strategies (Schenk, Minda, Lech, & Suchan, 2016; Nosofsky, 1986; Medin & Schaffer, 1978).

In prototype-based procedural learning, common characteristics of all category members are abstracted from the prototype of this category and are used to form a category representation. In exemplar-based learning, categories are learned on the basis of each stimulus. Whereas abstraction-based strategies are employed to learn typical members of a category, exceptions have to be learned by means of an exemplar-based strategy (Lech, Güntürkün, & Suchan, 2016; Cook & Smith, 2006).

The striatum is crucial for associative stimulus–response learning based on reward prediction errors, which are defined as violations to outcome predictions (Diederen, Spencer, Vestergaard, Fletcher, & Schultz, 2016; Rolls, McCabe, & Redoute, 2008). In feedback-based category learning, stimulus-category associations are acquired (Cincotta & Seger, 2007; Seger & Cincotta, 2005) by updating reward expectations after a reward prediction error was experienced (Nasser, Calu, Schoenbaum, & Sharpe, 2017; Seger, Peterson, Cincotta, Lopez-Paniagua, & Anderson, 2010; Sutton & Barto, 1981).

In electrophysiological studies with macaque monkeys, it was found that the striatum acts jointly with PFC during category learning, which was demonstrated by increases in the functional connectivity between the striatum and PFC during learning (Antzoulatos & Miller, 2011, 2014). Medial PFC regions, such as the dorsal ACC (dACC) and the dorsomedial PFC, represent the value of response options and reward expectancies and code reward prediction

[1]Ruhr-University Bochum, [2]University of Western Ontario

Journal of Cognitive Neuroscience 30:6, pp. 799–813
doi:10.1162/jocn_a_01241

errors. The dACC receives reinforcement learning signals (Alexander & Brown, 2011; Holroyd & Coles, 2002, 2008), and theta oscillations (4–8 Hz) are employed by medial frontal regions in the communication with lateral prefrontal and premotor regions to realize the adaptation of behavior (Smith et al., 2015; Oehrn et al., 2014; van de Vijver, Ridderinkhof, & Cohen, 2011).

In the scalp EEG, these frontal midline theta (FMT) oscillations have been linked with cognitive control processes (Cavanagh & Frank, 2014) and are involved in the evaluation of feedback and errors based on reinforcement learning signals. A larger FMT power is detected after negative feedback compared with positive feedback and after an error was made compared with a correct response (Cavanagh & Frank, 2014; Cohen, 2014; Cohen, Wilmes, & van de Vijver, 2011). The power of the FMT correlates with size of the reward prediction errors and the adaptation of behavior in subsequent trials (Mas-Herrero & Marco-Pallarés, 2014, 2016; van de Vijver et al., 2011; Cavanagh, Frank, Klein, & Allen, 2010). Inducing inphase theta oscillations synchronously in the medial and lateral PFC by means of transcranial alternating current stimulation has been shown to increase the feedback-based adaptation of behavior, whereas the induction of antiphase theta oscillations impairs the adaptation of behavior (Reinhart, 2017). Intracranial recordings in humans demonstrated directionality in this connectivity between medial and lateral prefrontal areas, such that theta oscillations propagate from medial prefrontal areas to the lateral PFC during feedback processing to adapt behavior (Smith et al., 2015). These studies suggest a causal role of the FMT in adaptive behavior. Theta oscillations are therefore a promising target to investigate the stress-related modulation of medial frontal processes in associative learning, such as abstraction-based categorization, with a high temporal resolution.

Stress and the stress-induced release of glucocorticoids are associated with changes in the feedback processing and learning (Lighthall et al., 2013; Cavanagh et al., 2011; Ossewaarde et al., 2011). The influence of stress on electrophysiological correlates of the feedback processing has been tested by recent studies using probabilistic reinforcement learning tasks. It was demonstrated that stress increases the amplitude of the feedback-related negativity (FRN; Wirz, Wacker, Felten, Reuter, & Schwabe, 2017; Glienke et al., 2015), a negative potential, which is more pronounced after negative feedback compared with positive feedback and which is the time domain representation of the FMT (Cohen et al., 2011).

Stress-related changes have also been found to influence categorization. In category learning, stress causes a switch between the relative use of learning strategies from declarative to nondeclarative strategies (Schwabe & Wolf, 2013). In a deterministic categorization task, dissociating striatal, procedural from prefrontal, rule-based learning, stress was found to enhance striatum-dependent procedural learning (Ell, Cosley, & McCoy, 2011). So far, it is unknown whether stress influences the learning of categories comprising typical, rule-following members and exceptions to the rule, which often occur in naturalistic categories.

The current study therefore tested the influence of stress on category learning of typical category members and exceptions. The analysis of frontal theta oscillations in the EEG during the feedback processing allows us to investigate the modulatory effects of stress on cortical oscillations that are involved in learning and feedback processing. Furthermore, a computational modeling procedure was applied to assess the use of prototype abstraction and exemplar-based strategies during learning. To investigate these questions, we exposed participants to either a stressful situation or a nonstressful control situation before they conducted a feedback-based category learning task composed of typical category members and exceptions (Cook & Smith, 2006). It was expected that stress might impair the exemplar-based learning of exceptions, whereas implicit, abstraction-based learning of typical stimuli should be unaffected. On the basis of previous studies (Schwabe & Wolf, 2012), it was expected that stress would foster the use of the abstraction-based strategies instead. Furthermore, stress was hypothesized to increase the FMT power reflecting changes in the frontal feedback processing. The FMT power increase may compensates for reduced hippocampal contribution or an increased need for cognitive control in the learning of exceptions.

METHODS

Participants

Forty-seven men participated in the study after a short screening for their general health status, with exclusion criteria such as smoking, current or history of psychiatric or neurological disorders, the intake of medication, a body mass index below 18 or above 29 kg/m^2, substance abuse, and color blindness. In addition, participants had to be naive to the socially evaluated cold pressor test (SECPT; Schwabe, Haddad, & Schachinger, 2008). Twenty-four participants were randomly assigned to the stress group (mean age = 24.8 years, *SEM* = 0.8 years), and 23 were assigned to the control group (mean age = 25.6 years, *SEM* = 0.9 years). Participants were excluded if their EEG data consisted of fewer than 10 artifact-free EEG trials in any of the experimental conditions ($n = 11$; see below for a detailed description). In addition, cortisol nonresponders (i.e., those not showing an increase in cortisol from baseline to peak in response to the stress of larger than 1.5 nmol/L; $n = 5$; Miller, Plessow, Kirschbaum, & Stalder, 2013) were excluded (Glienke et al., 2015; McCullough, Ritchey, Ranganath, & Yonelinas, 2015), because pharmacological work has repeatedly shown that glucocorticoids are crucial mediators in the impact of stress on different aspects of cognition (Vogel et al., 2017; Vogel, Fernández, Joëls, & Schwabe, 2016; Schwabe, Tegenthoff, Höffken, &

Wolf, 2013). The final sample consisted of 16 participants in the stress group and 16 in the control group.

The study was approved by the local ethics committee of the Faculty of Psychology at the Ruhr-University Bochum and is according to the Declaration of Helsinki. All participants gave written informed consent before participation and were reimbursed with €20.

Procedure

The experimental procedure was conducted between 2 and 6 p.m. to control for the diurnal cycle of endogenous cortisol concentrations (Kalsbeek et al., 2012). Participants refrained from alcohol and excessive exercise the day before the testing and refrained from eating and drinking anything except for water 2 hr before the testing. After giving written informed consent, electrodes were prepared for the EEG recordings. Subsequently, the first saliva sample was collected (−1 min) before the stress group underwent the stressful SECPT, whereas the control group was assigned to a nonstressful control procedure (both procedures are described in detail in the following section). Afterward, the second saliva sample (+1 min) was collected, and subjective stress ratings were obtained from the participants. Twenty minutes after the treatment, a saliva sample was collected (+20 min) and participants conducted the category learning task (see below). Afterward, a final saliva sample was collected (+55 min), and participants were shortly debriefed and reimbursed.

Stress Induction

Stress was induced by means of the SECPT, and the control group was assigned to a nonstressful control procedure. The stress group immersed their right hand in ice water (0–2°C) for up to 3 min. During the SECPT, they were filmed with a video camera and received instructions from a reserved female experimenter to keep their gaze fixated at the camera and to refrain from movement. The control group immersed their hands in warm water (35–37°C). They were neither filmed nor observed by a demure experimenter. Subjective ratings of the difficulty to keep the hand immersed, the discomfort, the pain, and the stress felt during the treatment were given on 11-point Likert scales from 0 (*not at all*) to 100 (*very much*). Systolic and diastolic blood pressure and heart rate were measured before, during, and after the treatment with the Dinamap system (Critikon, Tampa, FL). At each time point, three measurements of blood pressure and heart rate were obtained, which were used to calculate an average for each time point. Saliva was collected using Salivettes (Sarstedt, Nümbrecht, Germany). Samples were kept at −18°C until analysis. Cortisol concentrations were determined in duplicates using a cortisol enzyme-linked immunosorbent assay (Demeditec, Kiel, Germany). Interassay and intra-assay coefficients of variance were below 10%.

Category Learning Task

In the category learning task, participants learned, based on feedback to their responses, to assign circular stimuli (Figure 1A) to one of two categories (Schenk et al., 2016; Cook & Smith, 2006). Stimuli were constructed by systematically changing the colors of the six sections of a circle. Typical stimuli were derived from the prototype stimuli by changing the color of one section. Consequently, all typical stimuli in one category shared four of six colors (66.7%). One circle in each category was constructed as an exception. This exception shared zero to two colors (0%, 16.7%, or 33.3%) with the prototype and the typical circles of its category but instead shared four to five colors with the typical circles of the opposite category. In each trial, participants had to categorize one circle by a left or right button press and received a feedback about their accuracy.

The circles were presented for 800 msec, before the response options (Categories 1 and 2) were displayed concurrently with the stimulus for up to 1000 msec, until a response was made (Figure 1B). The chosen category was highlighted on the screen for 200 msec, followed by a blank screen for 500 msec, and feedback was presented for 1000 msec. As feedback, either "correct" or "wrong" was presented, depending on whether the prior stimulus was categorized correctly. If the response was too slow, a reminder for fast responses was presented. Intertrial intervals were jittered between 1000 and 2000 msec. Participants had to choose categories by button presses on the left (Category 1) or the right (Category 2) Ctrl keys on a standard keyboard. The learning task was composed of 420 trials in five blocks of 84 trials each, which were intermitted by self-paced breaks. Each block encompassed 72 typical trials and 12 exceptions. Prototypes and typical circles were pooled in the typical condition for all analyses (Lech et al., 2016; Schenk et al., 2016).

To analyze behavioral performance, the accuracy for both groups and for typical circles and exceptions was averaged over each block. Moreover, we applied a computational modeling procedure (described in detail in the following section) to assess the use of abstraction- and exemplar-based learning strategies in the categorization task.

Computational Modeling Procedures

To gain insight into how participants in each condition learned these category sets, we fit two well-known computational models to the data of stressed and control participants. The first model was based on the prototype model originally described by Minda and Smith (Minda & Smith, 2001, 2011; Smith & Minda, 1998). This model is ideally suited to model the current data as it was formulated and tested with the same categories used in the current study (Smith & Minda, 1998).

In the prototype model, each category is represented by a single prototype. Participants are assumed to abstract this

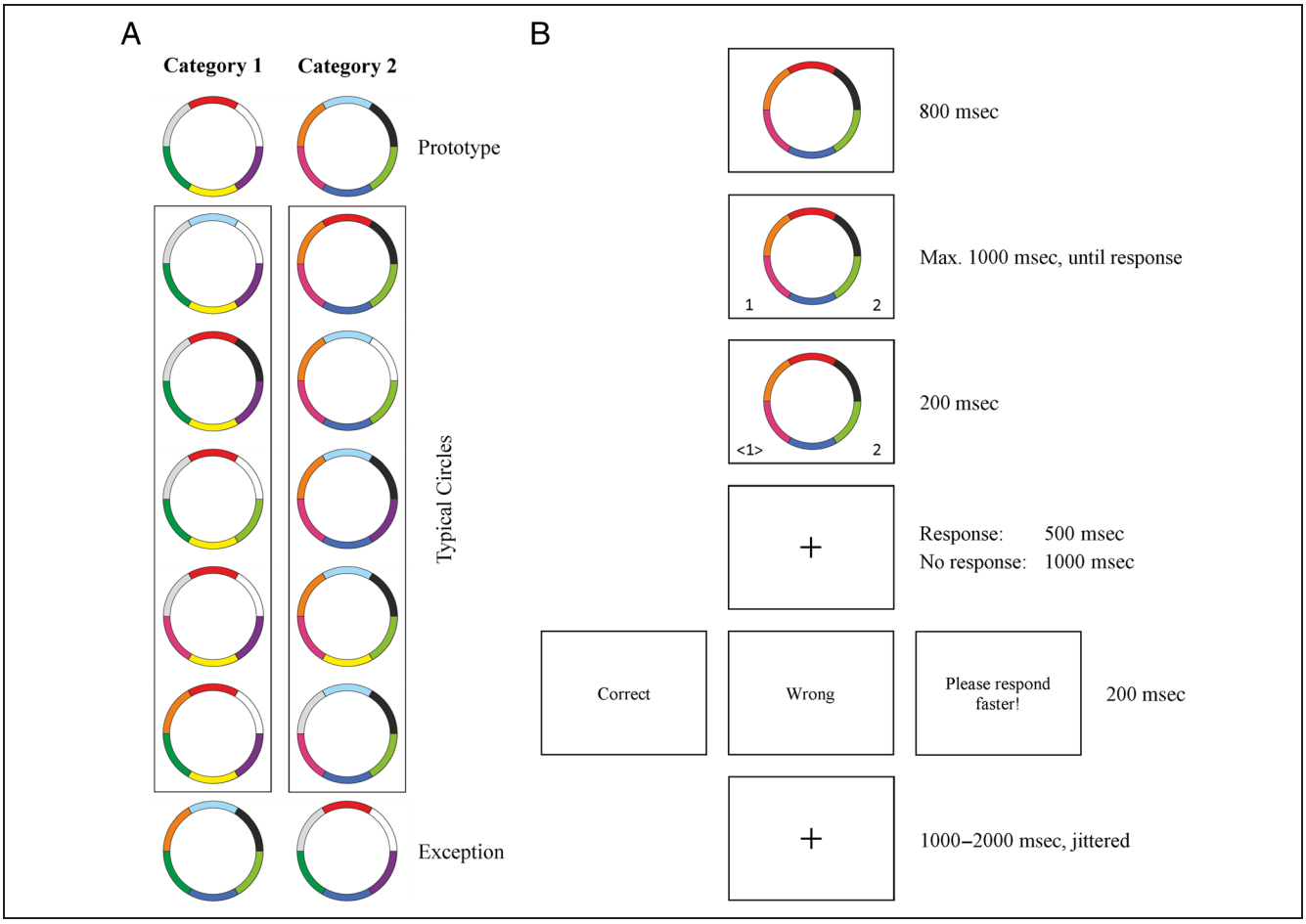

Figure 1. Schematic illustration of stimulus categories and the sequence of events in each trial. (A) The paradigm consisted of two categories. Typical circles were derived from the prototype of the respective category by changing one of the six colors. Exceptions were derived from the prototype of the opposite category. (B) In each trial, participants had to categorize one circle by a button press. On the basis of feedback, which was presented to the participants after their responses, they had to learn to categorize the circles correctly.

prototype from experience with individual exemplars, and they base a classification decision on the similarity of the to-be-classified item to each prototype. The item is then classified in accord with the most similar prototype. The prototype model predicts strong performance on prototypical and typical instances but poor performance on exception items.

The model is formulated as follows. First, the distance (d) between the item i and the prototype P is calculated by comparing the two stimuli along each weighted dimension k, as shown in Equation 1.

$$d_{iP} = \left[\sum_{k=1}^{N} w_k |x_{ik} - P_k|^r\right]^{\frac{1}{r}} \quad (1)$$

The value of r corresponds to the distance metric. When $r = 1$, the model uses a city-block distance metric, which is appropriate for separable-dimension stimuli (the present case). When $r = 2$, the model uses a Euclidean distance metric, which is appropriate for integral-dimension stimuli. Attentional weights (w) vary between 0.0 (*no attention*) and 1.0 (*exclusive attention*) and are constrained to sum to 1.0 across all the dimensions. The results of these weighted comparisons are summed across the dimensions to get the distance between the item and the prototype.

This distance (d_{iP}) between the item and the prototype is then converted into a measure of similarity (η_{iP}), following Shepard (1987), by taking

$$\eta_{iP} = e^{-cd_{iP}} \quad (2)$$

which gives a measure of similarity of an item i to prototype P. The exponent is distance (d_{iP}) multiplied by the scaling or sensitivity parameter c, which is a freely estimated parameter that reflects the steepness of the decay of similarity around the prototype. Prototype A similarity is then divided by the sum of Prototype A and Prototype B similarity to generate the model's predicted probability of a Category A response P (R_A) for stimulus (S_i), as shown in the probabilistic choice rule in Equation 3.

$$P(R_A|S_i) = \frac{\eta_{iP_A}}{\eta_{iP_A} + \eta_{iP_B}} \quad (3)$$

We also fit an exemplar model to the data, which was based on the Generalized Context Model of Nosofsky (1986, 2011). In this model, the learner is assumed to store representations of each category exemplar, and classification is based on the similarity of the to-be-categorized item to the entire collection of each exemplar in the category. The model is similar to the prototype model, except that the comparisons are made between exemplars. First, the distance between item i and exemplar j is calculated by comparing the two stimuli along each weighted dimension k, as shown in Equation 4. The value of r again corresponds to 1 in this case.

$$d_{ij} = \left[\sum_{k=1}^{N} w_k \left|x_{ik} - x_{jk}\right|^r\right]^{\frac{1}{r}} \tag{4}$$

Attentional weights (w) vary between 0.0 (*no attention*) and 1.0 (*exclusive attention*) and are constrained to sum to 1.0 across all the dimensions. The results of these weighted comparisons are summed across the dimensions to get the distance between the item and the exemplar.

This distance (d_{ij}) between the item and the exemplar is then converted into a measure of similarity (η_{ij}) by taking

$$\eta_{ij} = e^{-cd_{ij}} \tag{5}$$

which gives a measure of similarity of an item i to exemplar j. As with the prototype model, c is the freely estimated scaling parameter that reflects the steepness of the decay of similarity among exemplars.

The similarity of the item to the summed similarity to all the Category A exemplars is divided by the summed similarity of the item to exemplars in Categories A and B, as shown in the probabilistic choice rule in Equation 6.

$$P(R_A|S_i) = \frac{\sum_{j\in C_A} \eta_{ij}}{\sum_{j\in C_A} \eta_{ij} + \sum_{j\in C_B} \eta_{ij}} \tag{6}$$

Although the models differ in terms of their representational assumptions (abstracted prototype vs. stored exemplars), they are equivalent in terms of free parameters, attentional assumptions, and decision rules. Each model has a set of attentional weight parameters that correspond to the number of dimensions in a task (in this case, six dimensions) along with a psychological scaling parameter that corresponds to how close or distant the prototypes or exemplars are in psychological space. Because the attentional weights are constrained to sum to 1.0 (i.e., full attention), it leaves five free attentional parameters and the single, unconstrained scaling parameter for a total of six free parameters for each model.

Our analyses were based on fitting of each model to classification probabilities produced by each individual participant in each condition and at each block. The models were fit with a hill-climbing algorithm that adjusted the model's parameters to minimize the root mean square deviation (RMSD) between the observed data and the model's predictions (Minda & Smith, 2001). To find the best-fitting parameter settings of each model, a single parameter configuration (six attention weights and one scaling parameter) was chosen randomly and the predicted categorization probabilities for each of the 14 stimuli were calculated according to that configuration. The RMSD was calculated as shown in Equation 7.

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N} (O_i - P_i)^2}{N}} \tag{7}$$

The RMSD between the observed (O_i) and predicted (P_i) probabilities was then minimized with an algorithm that made a small adjustment to the provisional best-fitting parameter settings and chose the new settings if they produced a better fit (i.e., a smaller RMSD between predicted and observed performances). On each iteration of the algorithm, a weight parameter was adjusted by 0.01 (bounded by 0.00–1.00), or a scaling parameter was adjusted by 0.1 (bounded by 0.00–20.00). The algorithm continued to adjust until the fit could not be minimized further. To ensure that local minima were not a problem, the fitting procedure was repeated five times by choosing different random starting configurations of the model and hill-climbing from there. We chose the best-fitting parameters of the multiple fittings. The data for each block and each participant were fit independently, and each model (prototype or exemplar) was fit separately. This method, originally devised by Smith and Minda (1998), provides a static snapshot of performance, rather than an estimate of learning.

The resulting fit index (RMSD) from the model-fitting procedure provides information about how well (or not) each model fits the data of participants in each condition. On the basis of earlier work by Smith and Minda (1998) and more recent work by Minda and Smith (2011) and Schenk et al. (2016), we assumed that both models may fit moderately well early in learning but the exemplar model should provide a better fit later in the learning phase, because participants should have learned to classify all the exemplars, including the exception items. However, Smith and Minda (1998) also noted that the prototype model often fits the data better than the exemplar model early on, because many participants find it easy to learn to classify the prototype and typical items and systematically misclassify the exception items. In the present case, an advantage for the prototype model over the exemplar model might indicate that the participants have not been able to commit sufficient cognitive resources to learn the exceptions.

EEG Recording and Processing

Scalp EEG was recorded from 30 passive Ag/AgCl electrodes, which were distributed according to the 10–20 system (Pivik et al., 1993). Data were recorded at a sampling

rate of 500 Hz by a 32-channel BrainAmp Standard AC amplifier (Brain Products, Gilching, Germany), with a time constant of 10 sec. An electrode at the midfrontal position FPz was affixed to ground the participants, and data were referenced to linked mastoids. Impedances were kept below 10 kΩ.

Eye blinks were removed from continuous data using an independent component analysis as implemented in Brain Vision Analyzer 2 (Brain Products, Gilching, Germany; Lee, Girolami, & Sejnowski, 1999). For each participant, one independent component with a symmetrical frontal positive topography was removed, before the data were back-transformed. The successful removal of eye blinks was confirmed by visual inspection. Further analyses were performed using the FieldTrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011) and MATLAB 2016a (The MathWorks, Inc., Natick, MA).

Data were filtered with a 0.5-Hz high-pass, zero-phase Butterworth IIR filter and a band-stop filter from 48 to 52 Hz to eliminate line noise. Continuous data were epoched from −1.5 to 3 sec around feedback presentation. Epochs containing residual artifacts were removed during careful visual inspection. Subsequently, time–frequency decomposition was performed using 59 complex Morlet wavelets from 1 to 30 Hz, each having a width of five cycles. Power values were averaged across trials in each of the four following conditions: typicals/negative feedback, typicals/positive feedback, exceptions/negative feedback, and exceptions/positive feedback. Participants with less than 10 trials in any condition were excluded from further analyses. On average, typicals/negative feedback contains 63.44 (*SEM* = 5.29) trials, typicals/positive feedback contains 244.34 (6.37) trials, exceptions/negative feedback contains 26.75 (1.30) trials, and exceptions/positive feedback contains 23.25 (1.16) trials. Afterward, relative signal change to the condition-averaged baseline period (500 to 100 msec before stimulus presentation) was calculated for each condition to quantify feedback-related signal changes (Pfurtscheller & Aranibar, 1977). The baseline theta power did not differ between groups. To analyze differences in theta power, we averaged the power over the 200- to 600-msec postfeedback time interval (Chen, Zheng, Han, Chang, & Luo, 2017; Cunillera et al., 2012) and over the frequencies of 5–5.5 Hz. The frequencies of interest were determined using the difference between negative and positive feedback in averages across all groups and all conditions (Figure 2; Leicht et al., 2013). Group and condition differences were then analyzed in these peak frequencies, which were determined from group and condition averages.

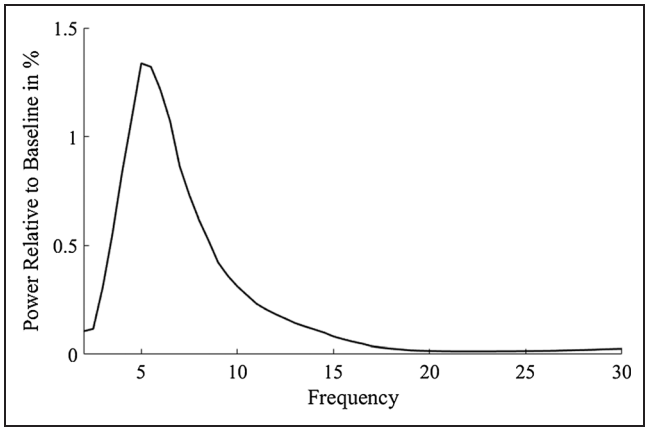

Figure 2. Power spectrum of the difference in the relative increase (200–600 msec postfeedback interval) to baseline between negative and positive feedback at electrode FCz. The frequencies of interest were defined as the frequencies with the maximal difference between negative and positive feedback pooled across both groups and both conditions.

Statistical Analyses

For the statistical analyses of the salivary cortisol concentrations and the behavioral data, repeated-measures ANOVAs were applied. All ANOVAs included the between-participant factor Group (stress and control). Group effects in salivary cortisol concentrations were analyzed statistically using repeated-measures ANOVA with the between-participant factor Group (stress and control) and the within-participant factor Time (−1, +1, +20, and +55). The analysis of the behavioral data included the within-participant factors Stimulus type (typical and exceptional items) and Block (1–5). The Greenhouse–Geisser correction was applied in cases of violations to the assumption of sphericity, and ε are reported. Significant interactions were resolved with Bonferroni-corrected post hoc pairwise *t* tests.

Differences in the subjective stress ratings and the cardiovascular measures of the stress response between the groups were tested by means of a multivariate ANOVA. Bonferroni-corrected post hoc pairwise comparisons were performed to test group differences on single subjective scales. The alpha level of .05 was applied to all parametric tests.

For the statistical analysis of the time–frequency data, nonparametric cluster-based permutation tests were applied to account for alpha error accumulations in the context of multiple comparison testing (Maris, 2012; Maris & Oostenveld, 2007). This cluster-based permutation test involves two steps: In the first step, spatially coherent clusters of electrodes exceeding a first level *t* threshold ($\alpha = .05$) are identified, and summed *t* values in each cluster are returned as test statistic. In the second step, this cluster-based test statistic is compared with a null distribution of the test statistic obtained by repeating Step 1 on data with randomly permuted condition affiliations for 1,000 iterations. For clusters, which reached the cluster α threshold, summed *t* values of the electrodes included in the cluster (t_{sum}) and cluster *p* values are reported. Additional to the whole electrode space covering cluster statistics, single-electrode *t* and *p* values of the frontocentral

electrode FCz are reported based on prior studies (e.g., Cavanagh, Figueroa, Cohen, & Frank, 2012; van de Vijver et al., 2011). Significant interactions in the cluster-based permutation tests were resolved using Wilcoxon signed-rank tests.

RESULTS

Subjective Stress Response

The subjective ratings revealed a successful stress induction by the SECPT ($F(4, 27) = 31.34, p < .001$, Wilk's $\Lambda = 0.177$, $\eta_p^2 = .82$). Stressed participants experienced more discomfort ($p < .001$, $d = 2.55$), more pain ($p < .001$, $d = 3.54$), and more stress ($p < .001$, $d = 2.15$) and reported more difficulty to keep the hand immersed ($p < .001$, $d = 3.38$) during the treatment (Table 1).

Cardiovascular Stress Responses

Increases in the blood pressure and heart rate of stressed participants during the stress induction revealed an activation of the sympathetic nervous system ($F(6, 116) = 15.33, p < .001$, Wilk's $\Lambda = 0.311$, $\eta_p^2 = .44$). Elevations of the systolic ($F(2, 60) = 30.92, p < .001, \eta_p^2 = .508$) and diastolic ($F(2, 60) = 39.06, p < .001, \eta_p^2 = .566$) blood pressure and of the heart rate ($F(2, 60) = 14.75, p < .001, \eta_p^2 = .330$) were confirmed by Time × Group interactions (Table 1).

Salivary Cortisol Concentrations

The stress group shows an elevation of salivary cortisol concentrations 20 min after the SECPT, whereas cortisol concentrations decrease over the course of the experiment in the control group (Figure 3A). The successful stress induction was confirmed by a significant Group × Time interaction ($F(3, 90) = 26.3, p < .001, \eta_p^2 = .467$) and a main effect of Time ($F(3, 90) = 19.86, p < .001, \eta_p^2 = .398$). The cortisol concentrations of the stress group were higher compared with those of the control group 20 min ($t(30) = -4.45, p < .001, d = 1.57$) and 55 min ($t(30) = -3.38, p = .002, d = 1.20$) after the treatment.

Learning Performance

The learning performance was assessed by calculating the percentage of correct responses of five blocks consisting of 72 typical trials or 12 trials with exceptions. Both stimulus classes, typical circles and exceptions, were successfully learned by the participants (Figure 3B), which was confirmed by a main effect of Block ($F(2.20, 66.01) = 44.21, p < .001, \eta_p^2 = .596, \varepsilon = .550$). The performance was better for typical circles throughout the whole task (main effect Stimulus type: $F(1, 30) = 192.15, p < .001, \eta_p^2 = .865$), but the increase in the performance was larger for the exceptions (last block − first block = 33.6%) compared with typical circles (24.44%; Stimulus type × Block interaction: $F(2.89, 86.80) = 3.960, p = .012, \eta_p^2 = .117, \varepsilon = .723$).

Stress, however, did not influence the learning performance, and neither the main effect of Group ($F(1, 30) = 0.20, p = .659, \eta_p^2 = .007$) nor any interaction including the factor Group revealed an influence of stress on the performance (Stimulus type × Group: $F(1, 30) = 2.58, p = .119, \eta_p^2 = .079$; Block × Group: $F(2.20, 66.01) = 1.00, p = .381, \eta_p^2 = .032, \varepsilon = .550$; Stimulus type × Block × Group: $F(2.89, 86.80) = 0.36, p = .775, \eta_p^2 = .012, \varepsilon = .723$).

A post hoc power analysis using G*Power (Version 3.1.9.2; Faul, Erdfelder, Lang, & Buchner, 2007) revealed that a medium-sized Stimulus type × Group interaction effect ($\eta_p^2 = .10$), which was expected based on stress-induced changes in category learning in previous studies (Schwabe & Wolf, 2012), would have been detectable with a power of $1 - \beta = .81$ ($\alpha = .05$; the average

Table 1. Differences in the Subjective and Cardiovascular Stress Responses between the Stress and Control Groups

	Control	*Stress*
Subjective stress response		
Discomfort	4.38 (2.03)	48.13 (5.72)**
Pain	0.00 (0.00)	60.00 (5.99)**
Stress	5.00 (2.24)	45.00 (6.19)**
Difficulty to keep hand immersed	1.88 (1.88)	57.50 (5.52)**
Systolic blood pressure (mm Hg)		
Pretreatment	123.27 (2.74)	125.92 (4.44)
During treatment	124.63 (2.58)	148.48 (4.05)**
Posttreatment	120.06 (2.65)	128.67 (3.73)
Diastolic blood pressure (mm Hg)		
Pretreatment	71.20 (2.82)	66.04 (2.88)
During treatment	73.98 (2.62)	87.48 (2.24)**
Posttreatment	66.44 (2.71)	69.27 (2.97)
Heart rate (BPM)		
Pretreatment	70.38 (2.65)	67.27 (3.30)
During treatment	69.46 (3.24)	75.19 (3.61)*
Posttreatment	68.48 (2.49)	65.46 (3.30)

Differences in subjective ratings were tested in planned *t* tests; differences in the cardiovascular responses were tested with Bonferroni-corrected post hoc *t* tests. Values represent mean (±*SEM*).

*$p < .1$.

**$p < .001$.

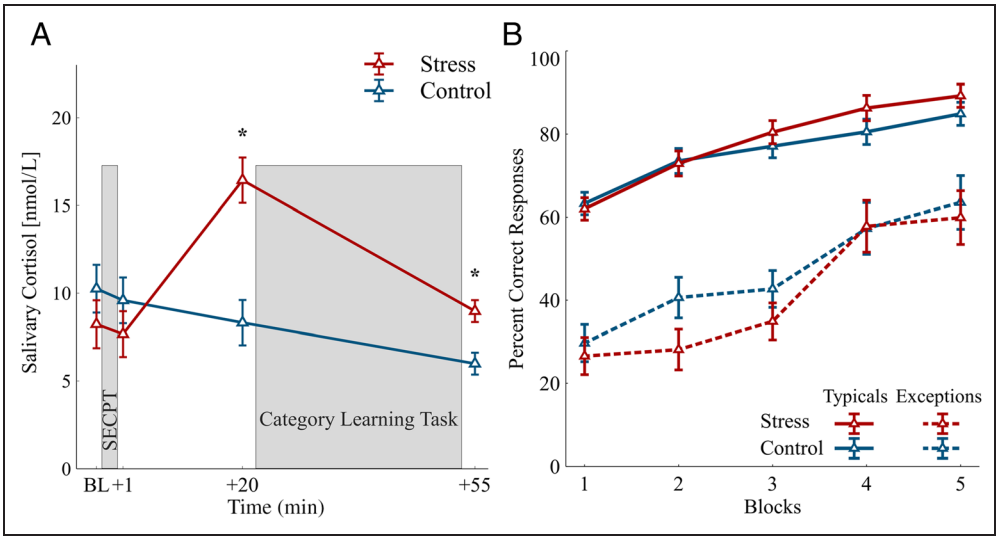

Figure 3. (A) Mean salivary cortisol concentrations at baseline (BL) as well as +1, +20, and +55 min after the SECPT/control procedure. Cortisol concentrations were elevated in the stress group compared with the control group 20 and 55 min after the SECPT. *$p < .01$. Error bars represent *SEM*. (B) Mean percent correct responses over the course of 30 blocks. The performance improved over the course of the experiments in typical category members and in exceptions as well as in both groups. Learning performance did not differ between the groups. Error bars represent *SEM*.

correlation between repeated measures in our sample was $r = .18$).

Computational Modeling

To examine the relative differences in model fit, we obtained the best fit (the RMSD) for each model fitting the data of each participant at each block, and we averaged across participants to obtain the average fit of each model. These average fits are plotted in Figure 4. Figure 4A shows the relative fits of the prototype model and the exemplar model for the participants in the stress condition, and Figure 4B shows the relative fit of the models when fitting data from the control condition. In both cases, the exemplar model fits better (has a lower RMSD) than the prototype model at the end of the learning phase. However, Figure 4A shows that the prototype model had an early advantage over the exemplar model at the second block for participants in the stress condition. This pattern was not observed in the data of the participants in the control condition, suggesting no early difficulty with exception items.

To examine these effects more directly, we conducted a 2 (Model) × 5 (Block) repeated-measures ANOVA on the RMSD values for each learning condition. For the Stress condition, we found find a main effect for Model ($F(1, 15) = 6.90$, $p = .019$, $\eta^2 = .315$), but not for Block ($F(4, 60) = 1.95$, $p = .115$, $\eta^2 = .115$). We also observed a significant interaction between Block and Model ($F(4, 60) = 6.16$, $p < .001$, $\eta^2 = .291$). To explore this interaction, we conducted five paired t tests (prototype fit vs. exemplar fit) at each block, using a Bonferroni correction to adjust the alpha level to .01 for the five comparisons. The models were not significantly different at the first to fourth blocks ($t(15) = -0.03$, -1.78, 0.70, and 2.25; $p = .973$, .095, .494, and .040, respectively), but the exemplar

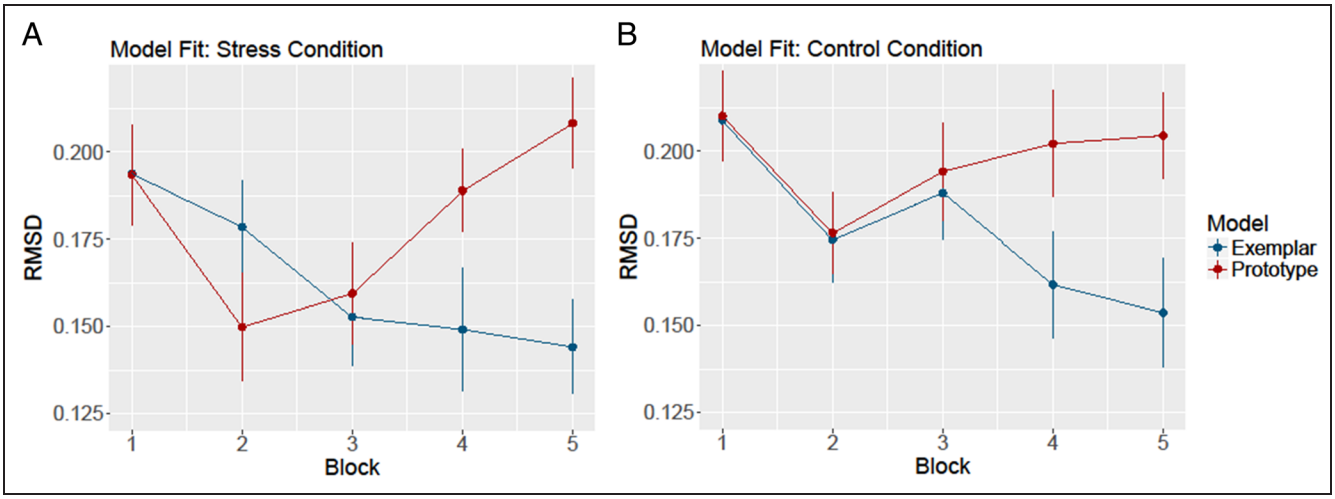

Figure 4. A shows the average fit (RMSD) across participants for the prototype and exemplar model fitting the data of participants in the stress condition. A lower RMSD indicates lower prediction error and a better fit for the model. B shows the average RMSD across participants for the prototype and exemplar model fitting the data of participants in the control condition. Error bars denote *SEM* for the calculated means.

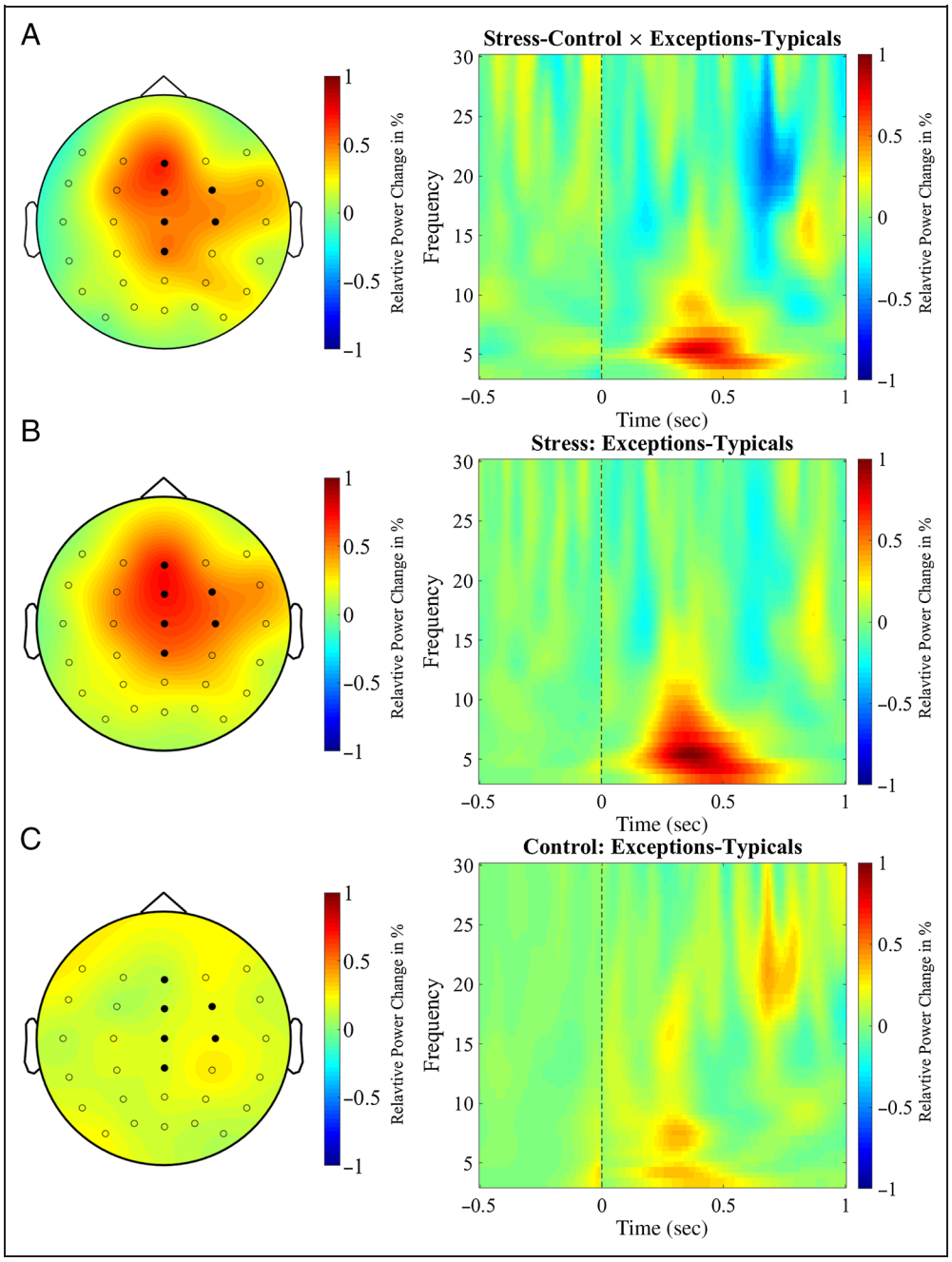

Figure 5. Topographical maps of the relative theta power (5–5.5 Hz, 200–600 msec postfeedback interval) and time–frequency plots at the electrode FCz relative to the presentation of negative feedback. On the topographical maps, filled circles represent electrodes of the cluster that revealed the significant Group × Stimulus type interaction. In the time–frequency plots, the prestimulus baseline interval (left of the dotted line) and the postfeedback activity (right of the dotted line) are shown. (A) A larger difference in the theta power between exceptions and typical circles was found in the stress group compared with the control group. (B) In the stress group, larger theta power was found for exceptions than for typical circles. (C) The control group did not show a difference in the theta power between exceptions and typical circles.

model fit significantly better than the prototype model in the fifth block, $t(15) = 3.57, p = .003$. Thus, although the prototype model appeared to have some advantage over the exemplar model in the second block, the advantage was not significant. The exemplar model did, however, significantly outperform the prototype model at the end of the study, and the main effect of model confirms that the exemplar model was fitting better overall.

For the Control condition, we again found a main effect for Model ($F(1, 15) = 12.51, p = .003, \eta^2 = .455$), but not for Block ($F(4, 60) = 2.00, p = .106, \eta^2 = .118$). We also observed a significant interaction between Block and Model ($F(4, 60) = 3.54, p = .012, \eta^2 = .191$). To explore this interaction, we conducted five paired t tests (prototype fit vs. exemplar fit) at each block, using a Bonferroni correction to adjust the alpha level to .01 for the five comparisons. The models were not significantly different at the first to third blocks ($t(15) = -0.14$, 0.34, and 0.62; $p = .889$, .738, and .543, respectively), but in the fourth block, the exemplar model fits significantly better than the prototype model ($t(15) = 2.95, p < .010$). In the fifth block, the t test revealed a trend toward the same direction ($t(15) = 2.58, p = .021$).

FMT

To assess whether the neural processes of feedback-based learning differ between the groups and between

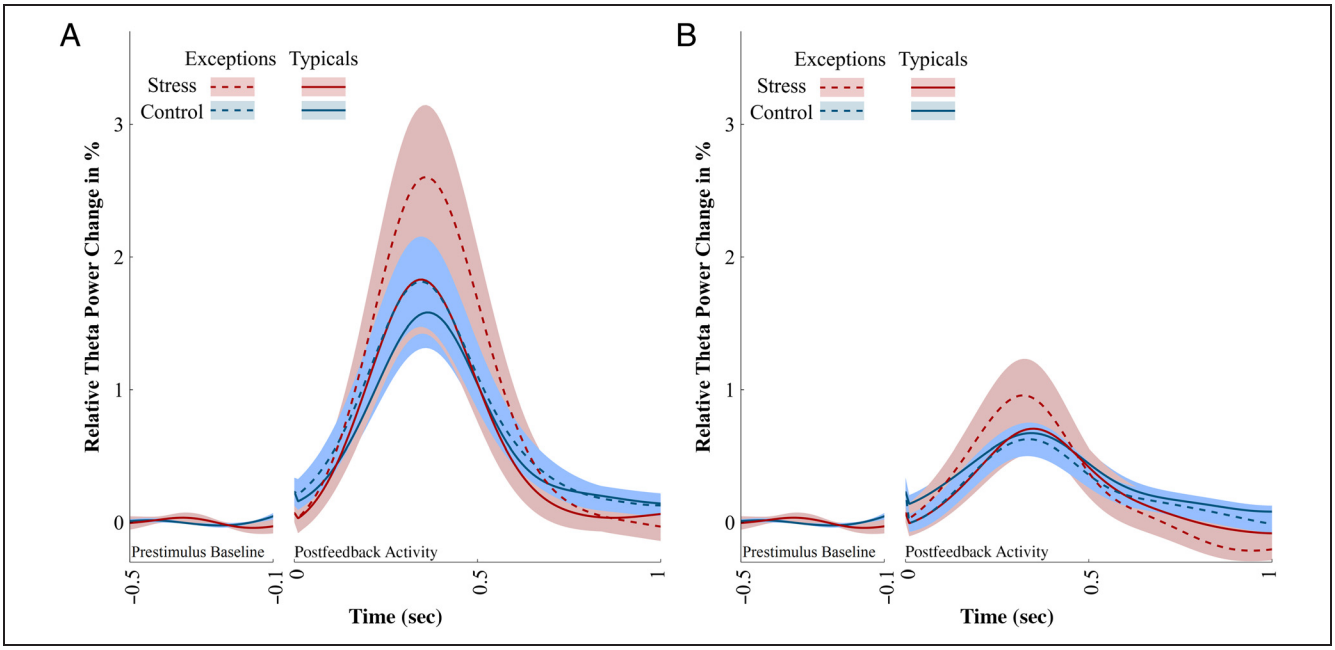

Figure 6. Theta power values (5–5.5 Hz) at electrode cluster Fz, FCz, FC4, Cz, C4, and CPz. The postfeedback theta power is plotted relative to the condition-averaged prestimulus baseline. A shows theta power after the presentation of negative feedback, and B shows theta power after positive feedback. Stress elevated the theta in exceptions after negative feedback. After positive feedback, theta power differs descriptively, but nonsignificantly, between groups or between exceptions and typical circles. Shaded areas represent *SEM*.

the learning of typical category members and exceptions, we analyzed the FMT power of feedback processing. First, the frequency window of the feedback-related theta power increase was determined by pooling the data of both groups and both conditions (Figure 2) for each frequency between 1 and 30 Hz and selecting the frequency showing the highest feedback-related change relative to baseline at electrode FCz. The highest feedback-related increase in theta power was found at 5–5.5 Hz. To investigate whether learning of typical and exceptional stimuli is related to different processes depending on stress level, differences in theta power between typical and exceptional trials were contrasted between groups. A cluster permutation test revealed that the Group × Stimulus type × Feedback valence interaction was not significant (all ts ≤ 1.6, all ps = 1). On the basis of previous studies, which showed that stress influences predominantly the use of negative feedback in reward learning and decision-making (Park, Lee, & Chey, 2017; Petzold, Plessow, Goschke, & Kirschbaum, 2010), we focused the analysis on negative feedback trials, where an effect of stress was expected. Indeed, a cluster permutation test revealed a significant Group × Stimulus type interaction ($t_{sum}(30) = 13.90$, $p = .042$) of FMT power (5–5.5 Hz) 200–600 msec after negative feedback at a frontocentral cluster of electrodes (Figure 5A), including the frontocentral electrode FCz ($t(30) = 2.25$, $p = .042$). This result reveals a difference in theta power during feedback processing during learning of typical category members and exceptions between the stress and control groups. To characterize this effect in depth, Wilcoxon signed-rank tests were used to contrast the theta power of this frontocentral electrode cluster in typical and exceptional stimuli separately for the stress and control groups. These follow-up tests demonstrated that stressed participants show an elevated theta power for trials with exceptions, compared with trials with typical circles ($Z = 3.00$, $p = .003$, $r = .750$; Figure 5B). This effect of stronger increases in theta power for exceptional compared with typical stimuli was missing in the control group ($Z = 0.88$, $p = .379$, $r = .220$; Figure 5C). These results indicate that the stress group shows an elevated FMT power after negative feedback, which is specific to the learning of exceptions (Figure 6A).

We focused the EEG analysis on negative feedback and found differences between the groups and stimulus types in frontal theta oscillations only in negative feedback. After positive feedback, the Group × Stimulus type interaction did not reach significance at any electrode site or cluster (all ts ≤ 1.88, all ps = 1; Figure 6B). Accordingly, neither in the stress group ($Z = 0.62$, $p = .535$, $r = .155$) nor in the control group ($Z = -1.40$, $p = .163$, $r = -.349$) did the typical circles and exceptions differ with respect to FMT oscillations after positive feedback.

DISCUSSION

In the current study, we examined the influence of acute stress on neural correlates of feedback processing during category learning of typical category members and exceptions. Participants were exposed to either a socially evaluative stress condition or a nonstressful control situation,

before conducting a feedback-based category learning task. The successful stress induction by the SECPT was demonstrated by increases in blood pressure, heart rate, and salivary cortisol concentrations. Both groups showed no substantial differences in behavioral performance. The stress and control groups successfully learned to assign typical stimuli and exceptions to the two categories and categorized typical category members more accurately than exceptions. Computational modeling revealed descriptively an increased use of the abstraction-based learning strategy in the early learning phase after stress. In the late learning phase, both groups made use of the exemplar learning strategy. Moreover, FMT power increases during negative feedback differed between the stress and control groups. During the learning of exceptions, midfrontal theta power after negative feedback increased in the stress group compared with the control group. After positive feedback, there was a descriptive but nonsignificant difference in the same direction.

Frontal theta oscillations are involved in the processing of feedback and are related to behavioral adaptation (Cohen et al., 2011; Cavanagh et al., 2010). Accordingly, an elevation of frontal theta power after stress suggests an increased involvement of frontal feedback processing in the learning of exceptions. This is in line with recent results of an increased FRN after stress (Wirz et al., 2017; Glienke et al., 2015), which is the time-locked reflection of feedback processing (Cavanagh, Zambrano-Vazquez, & Allen, 2012; Holroyd & Coles, 2002).

Other previous results, however, demonstrated an influence of anxiety on the FMT power (Mizuki et al., 1992; Mizuki, Hashimoto, Tanaka, Inanaga, & Tanaka, 1983) and illustrate a role of FMT oscillations in working memory maintenance (Hsieh & Ranganath, 2014) and cognitive control processes (Cavanagh & Frank, 2014; Cohen, 2014). Pastötter, Dreisbach, and Bäuml (2013), for instance, reported the largest FMT power in control-demanding incongruent task conditions.

The modeling analysis indicates that some participants in the stress condition may have relied more heavily on the abstracted prototypes early in the experiment relative to participants in the control condition, although this conclusion can only be tentative as the individual post hoc *t* tests were not significant. This may suggest that stressed participants had fewer cognitive resources available to learn the exceptions to the prototype, although participants in the stress condition still showed a clear, late advantage for exemplar learning, perhaps reflecting a recovery later in the learning phase.

The notion that stress could affect the learning of certain kinds of categories is consistent with studies demonstrating that performing a concurrent task that reduced the available working memory (Miles & Minda, 2011) or resource-depleting tasks (Minda & Rabi, 2015) impaired learning categories that required hypothesis testing and a rule selection, but not for tasks that relied on purely associative learning. The latter is assumed to be analogous to the prototype abstraction process (Minda & Miles, 2010). It is reasonable that stress might have reduced the availability of cognitive resources enough to allow interference with exception learning but not interference with prototype learning. Additional work is needed, however, with other category sets, such as those used by researchers showing impaired performance by older adults as a function of reduced working memory resources (Rabi & Minda, 2016). Alternatively, future research could also examine deterministic versus probabilistic responding and the possibility that the stress induction might have altered the choice strategy. Certain formulations of Nosofsky's Generalized Context Model can model this difference with a single additional parameter, but this same change cannot be instantiated in the prototype model without affecting the operation of the scaling parameter (Minda & Smith, 2001; Smith & Minda, 1998, 2000). Other possibilities for future work could examine alternative models, such as the drift diffusion class of models to make specific predictions about choice behavior and response time (Pedersen, Frank, & Biele, 2017).

In the current study, stress did not influence learning success. This is in line with previous studies, in which stress did not influence category learning performance per se but modulated the use of learning strategies and their neuronal basis (Wirz et al., 2017; Schwabe et al., 2013; Schwabe & Wolf, 2012). In light of the behavioral results, increases in theta power could reflect the increased recruitment of PFC regions, which have been linked with the abstraction of commonalities of category members (Mack, Preston, & Love, 2017; Pan & Sakagami, 2012). Medial prefrontal regions, such as the dACC and the dorsomedial PFC, have been shown to use theta oscillations to recruit regions in the lateral PFC in the realization of learning (Smith et al., 2015; van de Vijver et al., 2011). An alternative explanation could be that the early advantage of abstraction learning after stress increases the need of cognitive control, which has also been linked to midfrontal theta oscillations (Cavanagh & Frank, 2014), to switch to the exemplar learning strategy. Previous studies demonstrated that successful learning of exceptions relies on hippocampus activations during learning (Lech et al., 2016; Davis, Love, & Preston, 2012). It is known that stress impairs hippocampus-dependent memory processes, such as long-term memory retrieval (Wolf, 2017) as well as long-term memory encoding (Shields, Sazma, McCullough, & Yonelinas, 2017). We suggest that, after stress, an increased involvement of medial prefrontal regions might be responsible for adequate learning of exceptions in the current category learning task compensating a potential decline in hippocampal functionality.

Increases in electrophysiological correlates of the feedback processing after stress are thought to be related to an increased availability of dopamine in the striatum and PFC (Grace, 2016; Holly & Miczek, 2016), which is closely

linked to the processing of feedback (Frank, Seeberger, & O'Reilly, 2004; Holroyd & Coles, 2002; Schultz & Dickinson, 2000). Other studies, however, found an impairing effect of stress and glucocorticoids on the feedback-related activity of the reward system (Porcelli & Delgado, 2017; Montoya, Bos, Terburg, Rosenberger, & van Honk, 2014; Ossewaarde et al., 2011) and on the FRN and frontal theta power (Banis, Geerligs, & Lorist, 2014; Banis & Lorist, 2012). These inconsistent findings of stress effects on electrophysiological correlates of the feedback processing might depend on the timing of the stress, because enhancing effects of stress on the FRN (Wirz et al., 2017; Glienke et al., 2015) and theta power were found in the aftermath of stress, when cortisol levels are elevated (de Kloet, Joëls, & Holsboer, 2005), whereas attenuations were found when participants were stressed during the task (Banis et al., 2014; Banis & Lorist, 2012). Future studies are needed, which explicitly address the issue of timing of the stressor, similar to previous experiments in the domain of decision-making (e.g., Pabst, Brand, & Wolf, 2013).

We demonstrated an increase of the frontal theta power by stress after negative feedback. This is in line with the finding that FMT differentiates between positive and negative feedback, such that negative feedback elicits larger increases in theta power than positive feedback (Cohen, Elger, & Ranganath, 2007). Accordingly, theta oscillations are thought to reflect the processing of negative feedback and to be related to learning based on negative feedback (Andreou et al., 2017; Mas-Herrero, Ripollés, HajiHosseini, Rodríguez-Fornells, & Marco-Pallarés, 2015; van de Vijver et al., 2011).

In the current study, we could show that stress elevates the feedback-related theta power. The findings have implications for future fMRI or simultaneous EEG/fMRI studies (Hauser et al., 2015), which have to elucidate the relationship and timing of the learning systems involved in category learning. Importantly, future studies have to investigate the neural source of the observed increase in frontal theta oscillations and whether this increase is accompanied by reduced hippocampal activations.

Limitations

There are some limitations to the current study that need to be addressed. First, 15 participants had to be excluded from the study sample because of missing cortisol responses or too few trials for the EEG analysis. As a consequence, the absence of a significant stress effect on category learning performance and used learning strategy might be due to the somewhat reduced statistical power. Finally, the current study controlled for confounding gender differences by testing only male participants. Former studies reported sex differences in the responsiveness to acute stressors (Reschke-Hernández, Okerstrom, Bowles Edwards, & Tranel, 2017), in the effects of stress on emotional learning (Merz & Wolf, 2017), and in the effects of glucocorticoids on the reward system (Kinner, Wolf, & Merz, 2016). Future studies have to elucidate possible sex differences in the effects of stress on category learning and frontal theta power.

Conclusion

In summary, the current study reveals that stress has an enhancing effect on frontal theta oscillations in the processing of negative feedback during the category learning of exceptions. These results illustrate that stress modulates the neural basis of the learning of exceptions. The enhanced frontal theta oscillations might reflect a compensatory mechanism allowing preserved categorization performance of exceptions in the immediate aftermath of stress.

Acknowledgments

This research was supported by the Deutsche Forschungsgemeinschaft Projects B4, B8, and B11 of the Collaborative Research Centre (SFB) 874 "Integration and Representation of Sensory Processes." We thank Osman Akan, Alessa de Vries, Farina Helmke, Eve Hessas, Alexander Quent, and Theresa Wortmeier for assistance with data collection.

Reprint requests should be sent to Oliver T. Wolf, Department of Cognitive Psychology, Institute of Cognitive Neuroscience, Ruhr-University Bochum, Universitätsstr. 150, Bochum, Germany, 44780, or via e-mail: oliver.t.wolf@rub.de.

REFERENCES

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action–outcome predictor. *Nature Neuroscience, 14,* 1338–1344.

Andreou, C., Frielinghaus, H., Rauh, J., Mußmann, M., Vauth, S., Braun, P., et al. (2017). Theta and high-beta networks for feedback processing: A simultaneous EEG–fMRI study in healthy male subjects. *Translational Psychiatry, 7,* e1016.

Antzoulatos, E. G., & Miller, E. K. (2011). Differences between neural activity in prefrontal cortex and striatum during learning of novel abstract categories. *Neuron, 71,* 243–249.

Antzoulatos, E. G., & Miller, E. K. (2014). Increases in functional connectivity between prefrontal cortex and striatum during category learning. *Neuron, 83,* 216–225.

Ashby, F. G., & Ennis, J. M. (2006). The role of the basal ganglia in category learning. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 46, pp. 1–36). New York: Elsevier.

Banis, S., Geerligs, L., & Lorist, M. M. (2014). Acute stress modulates feedback processing in men and women: Differential effects on the feedback-related negativity and theta and beta power. *PLoS One, 9,* e95690.

Banis, S., & Lorist, M. M. (2012). Acute noise stress impairs feedback processing. *Biological Psychology, 91,* 163–171.

Cavanagh, J. F., Figueroa, C. M., Cohen, M. X., & Frank, M. J. (2012). Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. *Cerebral Cortex, 22,* 2575–2586.

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences, 18,* 414–421.

Cavanagh, J. F., Frank, M. J., & Allen, J. J. B. (2011). Social stress reactivity alters reward and punishment learning. *Social Cognitive and Affective Neuroscience, 6,* 311–320.
Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. B. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *Neuroimage, 49,* 3198–3209.
Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. B. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology, 49,* 220–238.
Chen, X., Zheng, T., Han, L., Chang, Y., & Luo, Y. (2017). The neural dynamics underlying the interpersonal effects of emotional expression on decision making. *Scientific Reports, 7,* 46651.
Cincotta, C. M., & Seger, C. A. (2007). Dissociation between striatal regions while learning to categorize via feedback and via observation. *Journal of Cognitive Neuroscience, 19,* 249–265.
Cohen, M. X. (2014). A neural microcircuit for cognitive conflict detection and signaling. *Trends in Neurosciences, 37,* 480–490.
Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage, 35,* 968–978.
Cohen, M. X., Wilmes, K., & van de Vijver, I. (2011). Cortical electrophysiological network dynamics of feedback learning. *Trends in Cognitive Sciences, 15,* 558–566.
Cook, R. G., & Smith, J. D. (2006). Stages of abstraction and exemplar memorization in pigeon category learning. *Psychological Science, 17,* 1059–1067.
Cunillera, T., Fuentemilla, L., Periañez, J., Marco-Pallarès, J., Krämer, U. M., Càmara, E., et al. (2012). Brain oscillatory activity associated with task switching and feedback processing. *Cognitive, Affective & Behavioral Neuroscience, 12,* 16–33.
Davis, T., Love, B. C., & Preston, A. R. (2012). Learning the exception to the rule: Model-based fMRI reveals specialized representations for surprising category members. *Cerebral Cortex, 22,* 260–273.
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., & Schultz, W. (2016). Adaptive prediction error coding in the human midbrain and striatum facilitates behavioral adaptation and learning efficiency. *Neuron, 90,* 1127–1138.
Ell, S. W., Cosley, B., & McCoy, S. K. (2011). When bad stress goes good: Increased threat reactivity predicts improved category learning performance. *Psychonomic Bulletin & Review, 18,* 96–102.
Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191.
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.
Glienke, K., Wolf, O. T., & Bellebaum, C. (2015). The impact of stress on feedback and error processing during behavioral adaptation. *Neuropsychologia, 71,* 181–190.
Grace, A. A. (2016). Dysregulation of the dopamine system in the pathophysiology of schizophrenia and depression. *Nature Reviews Neuroscience, 17,* 524–532.
Hauser, T. U., Hunt, L. T., Iannaccone, R., Walitza, S., Brandeis, D., Brem, S., et al. (2015). Temporally dissociable contributions of human medial prefrontal subregions to reward-guided learning. *Journal of Neuroscience, 35,* 11209–11220.
Holly, E. N., & Miczek, K. A. (2016). Ventral tegmental area dopamine revisited: Effects of acute and repeated stress. *Psychopharmacology, 233,* 163–186.
Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109,* 679–709.
Holroyd, C. B., & Coles, M. G. H. (2008). Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. *Cortex, 44,* 548–559.
Hsieh, L.-T., & Ranganath, C. (2014). Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval. *Neuroimage, 85,* 721–729.
Kalsbeek, A., van der Spek, R., Lei, J., Endert, E., Buijs, R. M., & Fliers, E. (2012). Circadian rhythms in the hypothalamo–pituitary–adrenal (HPA) axis. *Molecular and Cellular Endocrinology, 349,* 20–29.
Kinner, V. L., Wolf, O. T., & Merz, C. J. (2016). Cortisol alters reward processing in the human brain. *Hormones and Behavior, 84,* 75–83.
Lech, R. K., Güntürkün, O., & Suchan, B. (2016). An interplay of fusiform gyrus and hippocampus enables prototype- and exemplar-based category learning. *Behavioural Brain Research, 311,* 239–246.
Lee, T. W., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation, 11,* 417–441.
Leicht, G., Troschütz, S., Andreou, C., Karamatskos, E., Ertl, M., Naber, D., et al. (2013). Relationship between oscillatory neuronal activity during reward processing and trait impulsivity and sensation seeking. *PLoS One, 8,* e83414.
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.
Mack, M. L., Preston, A. R., & Love, B. C. (2017). Medial prefrontal cortex compresses concept representations through learning. *bioRxiv, 178145,* 1–22.
Maris, E. (2012). Statistical testing in electrophysiological studies. *Psychophysiology, 49,* 549–565.
Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164,* 177–190.
Mas-Herrero, E., & Marco-Pallarés, J. (2014). Frontal theta oscillatory activity is a common mechanism for the computation of unexpected outcomes and learning rate. *Journal of Cognitive Neuroscience, 26,* 447–458.
Mas-Herrero, E., & Marco-Pallarés, J. (2016). Theta oscillations integrate functionally segregated sub-regions of the medial prefrontal cortex. *Neuroimage, 143,* 166–174.
Mas-Herrero, E., Ripollés, P., HajiHosseini, A., Rodríguez-Fornells, A., & Marco-Pallarés, J. (2015). Beta oscillations and reward processing: Coupling oscillatory activity and hemodynamic responses. *Neuroimage, 119,* 13–19.
McCullough, A. M., Ritchey, M., Ranganath, C., & Yonelinas, A. (2015). Differential effects of stress-induced cortisol responses on recollection and familiarity-based recognition memory. *Neurobiology of Learning and Memory, 123,* 1–10.
Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85,* 207–238.
Merz, C. J., & Wolf, O. T. (2017). Sex differences in stress effects on emotional learning. *Journal of Neuroscience Research, 95,* 93–105.
Miles, S. J., & Minda, J. P. (2011). The effects of concurrent verbal and visual tasks on category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 588–607.
Miller, R., Plessow, F., Kirschbaum, C., & Stalder, T. (2013). Classification criteria for distinguishing cortisol responders from nonresponders to psychosocial stress: Evaluation of

salivary cortisol pulse detection in panel designs. *Psychosomatic Medicine, 75,* 832–840.
Minda, J., & Miles, S. (2010). The influence of verbal and nonverbal processing on category learning. In B. H. Ross (Ed.), *Psychology of learning and motivation—Advances in research and theory* (Vol. 52, pp. 117–162). Burlington, VT: Academic Press.
Minda, J. P., & Rabi, R. (2015). Ego depletion interferes with rule-defined category learning but not non-rule-defined category learning. *Frontiers in Psychology, 6,* 1–9.
Minda, J. P., & Smith, J. D. (2001). Prototypes in category learning: The effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 775–799.
Minda, J. P., & Smith, J. D. (2011). Prototype models of categorization: Basic formulation, predictions, and limitations. In E. Pothos & A. J. Wills (Eds.), *Formal approaches in categorization* (pp. 40–64). Cambridge: Cambridge University Press.
Mizuki, Y., Hashimoto, M., Tanaka, T., Inanaga, K., & Tanaka, M. (1983). A new physiological tool for assessing anxiolytic effects in humans: Frontal midline theta activity. *Psychopharmacology, 80,* 311–314.
Mizuki, Y., Kajimura, N., Kai, S., Suetsugi, M., Ushijima, I., & Yamada, M. (1992). Differential responses to mental stress in high and low anxious normal humans assessed by frontal midline theta activity. *International Journal of Psychophysiology, 12,* 169–178.
Montoya, E. R., Bos, P. A., Terburg, D., Rosenberger, L. A., & van Honk, J. (2014). Cortisol administration induces global down-regulation of the brain's reward circuitry. *Psychoneuroendocrinology, 47,* 31–42.
Nasser, H. M., Calu, D. J., Schoenbaum, G., & Sharpe, M. J. (2017). The dopamine prediction error: Contributions to associative models of reward learning. *Frontiers in Psychology, 8,* 1–17.
Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115,* 39–57.
Nosofsky, R. M. (2011). The generalized context model: An exemplar model of classification. In E. Pothos & A. J. Wills (Eds.), *Formal approaches in categorization* (pp. 18–39). Cambridge, UK: Cambridge University Press.
Oehrn, C. R., Hanslmayr, S., Fell, J., Deuker, L., Kremers, N. A., Do Lam, A. T., et al. (2014). Neural communication patterns underlying conflict detection, resolution, and adaptation. *Journal of Neuroscience, 34,* 10438–10452.
Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011,* 156869.
Ossewaarde, L., Qin, S., Van Marle, H. J. F., van Wingen, G. A., Fernández, G., & Hermans, E. J. (2011). Stress-induced reduction in reward-related prefrontal cortex function. *Neuroimage, 55,* 345–352.
Pabst, S., Brand, M., & Wolf, O. T. (2013). Stress and decision making: A few minutes make all the difference. *Behavioural Brain Research, 250,* 39–45.
Pan, X., & Sakagami, M. (2012). Category representation and generalization in the prefrontal cortex. *European Journal of Neuroscience, 35,* 1083–1091.
Park, H., Lee, D., & Chey, J. (2017). Stress enhances model-free reinforcement learning only after negative outcome. *PLoS One, 12,* e0180588.
Pastötter, B., Dreisbach, G., & Bäuml, K.-H. T. (2013). Dynamic adjustments of cognitive control: Oscillatory correlates of the conflict adaptation effect. *Journal of Cognitive Neuroscience, 25,* 2167–2178.
Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule in reinforcement learning. *Psychonomic Bulletin & Review, 24,* 1234–1251.
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.
Pfurtscheller, G., & Aranibar, A. (1977). Event-related cortical desynchronization detected by power measurements of scalp EEG. *Electroencephalography and Clinical Neurophysiology, 42,* 817–826.
Pivik, R. T., Broughton, R. J., Coppola, R., Davidson, R. J., Fox, N., & Nuwer, M. R. (1993). Guidelines for the recording and quantitative analysis of electroencephalographic activity in research contexts. *Psychophysiology, 30,* 547–558.
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.
Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353–363.
Rabi, R., & Minda, J. P. (2016). Category learning in older adulthood: A study of the Shepard, Hovland, and Jenkins (1961) tasks. *Psychology and Aging, 31,* 185–197.
Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology, 3,* 382–407.
Reinhart, R. M. G. (2017). Disruption and rescue of interareal theta phase coupling and adaptive behavior. *Proceedings of the National Academy of Sciences, U.S.A., 114,* 11542–11547.
Reschke-Hernández, A. E., Okerstrom, K. L., Bowles Edwards, A., & Tranel, D. (2017). Sex and stress: Men and women show different cortisol responses to psychological stress induced by the Trier social stress test and the Iowa singing social stress test. *Journal of Neuroscience Research, 95,* 106–114.
Rolls, E. T., McCabe, C., & Redoute, J. (2008). Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cerebral Cortex, 18,* 652–663.
Schenk, S., Minda, J. P., Lech, R. K., & Suchan, B. (2016). Out of sight, out of mind: Categorization learning and normal aging. *Neuropsychologia, 91,* 222–233.
Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience, 23,* 473–500.
Schwabe, L., Haddad, L., & Schachinger, H. (2008). HPA axis activation by a socially evaluated cold-pressor test. *Psychoneuroendocrinology, 33,* 890–895.
Schwabe, L., Tegenthoff, M., Höffken, O., & Wolf, O. T. (2013). Mineralocorticoid receptor blockade prevents stress-induced modulation of multiple memory systems in the human brain. *Biological Psychiatry, 74,* 801–808.
Schwabe, L., & Wolf, O. T. (2012). Stress modulates the engagement of multiple memory systems in classification learning. *Journal of Neuroscience, 32,* 11042–11049.
Schwabe, L., & Wolf, O. T. (2013). Stress and multiple memory systems: From "thinking" to "doing". *Trends in Cognitive Sciences, 17,* 60–68.
Seger, C. A., & Cincotta, C. M. (2005). The roles of the caudate nucleus in human classification learning. *Journal of Neuroscience, 25,* 2941–2951.
Seger, C. A., Peterson, E. J., Cincotta, C. M., Lopez-Paniagua, D., & Anderson, C. W. (2010). Dissociating the contributions of independent corticostriatal systems to visual categorization learning through the use of reinforcement learning modeling and Granger causality modeling. *Neuroimage, 50,* 644–656.
Shepard, R. (1987). Toward a universal law of generalization for psychological science. *Science, 237,* 1317–1323.
Shields, G. S., Sazma, M. A., McCullough, A. M., & Yonelinas, A. P. (2017). The effects of acute stress on episodic memory:

A meta-analysis and integrative review. *Psychological Bulletin, 143,* 636–675.
Smith, D. J., & Minda, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 3–27.
Smith, E. H., Banks, G. P., Mikell, C. B., Cash, S. S., Patel, S. R., Eskandar, E. N., et al. (2015). Frequency-dependent representation of reinforcement-related information in the human medial and lateral prefrontal cortex. *Journal of Neuroscience, 35,* 15827–15836.
Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1411–1436.
Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review, 88,* 135–170.
van de Vijver, I., Ridderinkhof, K. R., & Cohen, M. X. (2011). Frontal oscillatory dynamics predict feedback learning and action adjustment. *Journal of Cognitive Neuroscience, 23,* 4106–4121.
Vogel, S., Fernández, G., Joëls, M., & Schwabe, L. (2016). Cognitive adaptation under stress: A case for the mineralocorticoid receptor. *Trends in Cognitive Sciences, 20,* 192–203.
Vogel, S., Klumpers, F., Schröder, T. N., Oplaat, K. T., Krugers, H. J., Oitzl, M. S., et al. (2017). Stress induces a shift towards striatum-dependent stimulus–response learning via the mineralocorticoid receptor. *Neuropsychopharmacology, 42,* 1262–1271.
Wirz, L., Wacker, J., Felten, A., Reuter, M., & Schwabe, L. (2017). A deletion variant of the α2b-adrenoceptor modulates the stress-induced shift from "cognitive" to "habit" memory. *Journal of Neuroscience, 37,* 2149–2160.
Wolf, O. T. (2017). Stress and memory retrieval: Mechanisms and consequences. *Current Opinion in Behavioral Sciences, 14,* 40–46.