



Social Context and Rejection Expectations Modulate Neural and Behavioral Responses to Social Feedback

Pauline Petereit^{iD}, Sarah Jessen, Tatiana Goregliad Fjaellingsdal, and Ulrike M. Krämer

Abstract

■ When meeting other people, some are optimistic and expect to be accepted by others, whereas others are pessimistic and expect mostly rejections. How social feedback is evaluated in situations that meet or do not meet these biases and how people differ in their response to rejection and acceptance depending on the social situation are unknown. In this study, participants experienced rejection and acceptance by peers in two different social contexts, one with high (negative context) and the other with low probability of rejection (positive context). We examined how the neural and behavioral responses to rejection are altered by this context and whether it depends on the individual's sensitivity to rejection. Behavioral results show that, on average, people maintain an optimistic bias even when mostly experiencing rejection. Importantly, personality differences in

rejection sensitivity affected both prior expectations to be rejected in the paradigm and the extent to which expectations changed during the paradigm. The context also strongly modulated ERPs and theta responses to rejection and acceptance feedback. Specifically, valence effects on neural responses were enhanced in the negative context, suggesting a greater relevance to monitor social feedback in such a situation. Moreover, mid-frontal theta predicted how expectations were changed in response to prediction errors, stressing a role for theta in learning from social feedback. Surprisingly, interindividual differences in rejection sensitivity did not affect neural responses to feedback. Our results stress the importance of considering the interaction between subjective expectations and the social context for behavioral and neural responses to social rejection. ■

INTRODUCTION

When we interact with others, we constantly receive feedback about ourselves as persons and whether others like us or not. This feedback can take implicit forms, like a friendly smile, or more explicit forms, like a refused dating proposal. Social feedback gives vital information about the success of our social interactions and about our value for others and is crucial for feelings of self-worth (Gruenenfelder-Steiger, Harris, & Fend, 2016; Denissen, Penke, Schmitt, & van Aken, 2008; Williams, 2007; Pickett, Gardner, & Knowles, 2004). Acknowledging this relevance, social neuroscientists have started to characterize the neural processes involved when receiving social feedback (Kortink, Weeda, Crowley, Gunther Moor, & van der Molen, 2018; van der Molen, Harrewijn, & Westenberg, 2018; Will, Rutledge, Moutoussis, & Dolan, 2017; Cao, Gu, Bi, Zhu, & Wu, 2015; Jones et al., 2011). A characteristic aspect of social feedback, so far neglected in the neuroscientific literature, is that instances of rejection or acceptance usually do not happen in isolation but are embedded in a social context. Such context can be overall positive, like a dinner party with friends, or overall negative, like a debate with an opposing political group. Importantly, positive or negative feedback might be very differently evaluated depending on the social context. In

a mostly positive context, we might well disregard the few instances of rejection. In a mostly negative context, we might value the few instances of acceptance even more. Moreover, people encounter social situations with certain expectations about whether or not others will like them (Downey & Feldman, 1996). Research has shown that most people are overly optimistic and expect more acceptance than rejection when encountering friends or strangers (Loeb, Tan, Hessel, & Allen, 2018; van der Molen et al., 2018; Loeb, Hessel, & Allen, 2016; Cao et al., 2015; Hepper, Hart, Gregg, & Sedikides, 2011). A predominantly positive social situation would thus fulfill precisely these expectations. On the contrary, strong rejection experiences in childhood and adolescence can lead to high rejection sensitivity later in life (London, Downey, Bonica, & Paltin, 2007; Feldman & Downey, 1994), that is, high expectations to be rejected during a social encounter (Gutz, Roepke, & Renneberg, 2016). A mostly negative social context would thus match the expectations of a person high in rejection sensitivity. In the current study, our main question was therefore how the social context and a person's trait rejection sensitivity interact to shape the behavioral and neural response to social feedback.

As feedback processing is particularly relevant for how we learn from feedback in a given situation, that is, how we change our behavior afterward, we were interested in both the general response to feedback and the trial-by-trial

University of Lübeck, Lübeck, Germany

changes related to behavioral adaptations. People do not only differ in their expectations about the probability of being rejected (London et al., 2007; Feldman & Downey, 1994), these expectations presumably relate to differences in how people learn from social feedback and thus change their expectations once they participate in a social situation (Kube & Rozenkrantz, 2021). Research on the processing of self-related information suggests that healthy people show an optimistic learning bias and tend to change their opinions more in response to better-than-expected compared with worse-than-expected information (Koban et al., 2017; Korn, Prehn, Park, Walter, & Heekeren, 2012; Sharot, Korn, & Dolan, 2011). In two of the few studies that explicitly investigated learning from feedback about one's likeability as a person, Will et al. (2017, 2020) presented participants with feedback of ostensible peers. How expectations to be liked changed after receiving feedback was explained using so-called social prediction errors (PEs), that is, the difference between received feedback and one's expectation, weighed by a learning rate (Rescorla & Wagner, 1972). Generally, participants overestimated the probability to be liked by others by about 5%, which was explained by a positive response bias parameter in the learning model (Will et al., 2017, 2020). In contrast, people with social anxiety disorder (which is related to high rejection sensitivity; Staebler, Helbing, Rosenbach, & Renneberg, 2011) show a tendency to weigh negative feedback more when updating their feelings of self-worth based on performance feedback by others (Koban et al., 2017).

Based on these prior findings, we assume that the extent to which people are biased in their learning about the probability to be rejected or accepted in a given social situation is influenced by their initial expectations. This would also be in line with recent refinements to classical reinforcement learning theory, stating that outcome expectations with great certainty (e.g., after having learned that acceptance is very likely) are not easily changed, even when the PE is large (e.g., when unexpected rejection occurs; Spicer, Mitchell, Wills, & Jones, 2020). Specifically, as former research suggests, we assume that most people have rather positive expectations before entering a social situation and change these expectations with a positive bias (Will et al., 2017; Korn et al., 2012; Hepper et al., 2011). Importantly, this bias will presumably interact with the social context. In mostly positive contexts, it will lead to neglecting rejections; in mostly negative situations, an optimistic bias will counteract the experience of predominant rejections, which would result in using both rejection and acceptance signals for learning. Although such a pattern can be expected for people with an optimistic bias, it can be assumed to be less asymmetric or even inverted for individuals high in rejection sensitivity.

Decades of research have scrutinized the neural signatures of performance feedback and their sensitivity to expectations and learning (Glazer, Kelley, Pornpattananangkul, Mittal, & Nusslock, 2018; Luft, 2014; Gehring &

Willoughby, 2002). More recent studies have started to investigate how these neural signatures of performance feedback translate to processing of social feedback (e.g., van der Veen, Burdzina, & Langeslag, 2019; Kortink et al., 2018; van der Molen et al., 2018; Cao et al., 2015; Sun & Yu, 2014). In the current study, we therefore examined the effects of context and interindividual differences not only on behavioral but also on neural signatures of feedback processing. In this way, we aimed to clarify their functional role for behavioral responses to social feedback. Of particular relevance for this study are the P3 and the midfrontal theta response. The P3, a positive going ERP occurring at approximately 300 msec after feedback onset, has been shown to reflect the subjective probability and emotional salience of feedback and the updating of context information (Glazer et al., 2018; Hajcak, MacNamara, & Olvet, 2010; Polich, 2007). Results on the sensitivity of the P3 to subjective probability or valence of social feedback are mixed (Harrewijn, van der Molen, van Vliet, Tissier, & Westenberg, 2018; Cao et al., 2015; Dekkers, van der Molen, Gunther Moor, van der Veen, & van der Molen, 2015; Leitner, Hehman, Jones, & Forbes, 2014; Sun & Yu, 2014). In some studies, a stronger P3 for expected acceptance has been found and interpreted as neural marker of the rewarding value of the acceptance (Kortink et al., 2018; van der Molen et al., 2018). In others, only effects of expectancy (Dekkers et al., 2015; Gutz, Renneberg, Roepke, & Niedeggen, 2015) or valence (van der Veen et al., 2019; Harrewijn et al., 2018) but no interactions have been found. Some studies also reported no modulation of P3 by social feedback (Cao et al., 2015; Leitner et al., 2014). Regarding interindividual differences in rejection sensitivity, previous work found contradicting results, namely, stronger responses to either rejection or acceptance in people with high rejection sensitivity. The former was interpreted as indicative of higher salience of rejection, whereas the latter was interpreted as indicative of greater unexpectedness of acceptance (Leng, Qian, & Zhu, 2018; Gutz et al., 2015).

Another neural signature of feedback processing is midfrontal theta power, which is considered to reflect the processing of PEs in the ACC (Luu, Tucker, & Makeig, 2004) or dorsal ACC (dACC; Weiss et al., 2018). Changes in theta power in response to feedback have also been reported in the social domain. For instance, van der Molen, Dekkers, Westenberg, van der Veen, and van der Molen (2017) reported theta responses to unexpected rejection feedback, which were localized to the dACC. Other studies confirmed enhanced theta power for unexpected rejection, that is, for negative PEs (Harrewijn et al., 2018; Kortink et al., 2018; van der Molen et al., 2018). Increased dACC activity in response to negative feedback might be related to another line of research, which linked the dACC to the negative affect elicited by social rejection and exclusion (Eisenberger, 2015a, 2015b). Together, these previous studies suggest that midfrontal theta is sensitive to the valence and expectedness of social feedback.

However, in previous EEG studies on social feedback processing, acceptance and rejection were presented with equal probability (van der Veen et al., 2019; Kortink et al., 2018; van der Molen et al., 2017, 2018; Cao et al., 2015; Dekkers et al., 2015). Therefore, it is still unclear whether P3 and midfrontal theta reflect the overall probability of social rejection, the momentary subjective expectation expressed in a single trial, or simply the valence of the feedback, as the first factor was always kept constant. Studying if P3 and midfrontal theta are sensitive to either of these factors could help clarify the neural bases of expectation biases, namely, if these relate to a flawed (or absent) representation of objective probabilities.

Moreover, as changes in P3 have been shown to track learning from nonsocial feedback (Zioga, Hassan, & Luft, 2019; Bennett, Murawski, & Bode, 2015), we investigated whether P3 effects change over time. Previous research found that increases in theta responses to incorrect feedback predicted adequate performance adjustments in the next trial (van de Vijver, Ridderinkhof, & Cohen, 2011), and better learners showed stronger theta responses than poor learners (Luft, Nolte, & Bhattacharya, 2013; De Pascalis, Varriale, & Rotonda, 2012). Based on these findings, we assumed greater theta power after PEs to predict greater adjustments of expectations in the following trials (Luft, 2014), possibly dependent on interindividual differences. So far, interindividual differences in learning from social feedback could not be linked consistently to interindividual differences in neural responses to that feedback (Kortink et al., 2018). We aim to close this gap, as this link could provide valuable insight into the processes underlying dysfunctional expectations in people with high rejection sensitivity.

To address our research questions, we presented participants in the current study either predominantly with rejection or predominantly with acceptance from ostensible peers to manipulate the social context. We assessed participants' expectations before and during the experiment and analyzed the influence of social context and prior expectations on two levels: first, on a "macroscopic" level on average behavioral and neural responses to rejection and acceptance across the paradigm, and second, on a "microscopic" level on trial-by-trial dynamics to track changes in expectations and neural responses to rejection and acceptance.

METHODS

Participants

We recruited young adults in the age range from 18 to 35 years from the student population of Lübeck. Exclusion criteria were studying psychology (except students of the first year), current psychiatric illnesses, current or former neurological illnesses or injuries, and left-handedness. We invited 106 individuals to the laboratory, of whom 14 had to be excluded, either because of disbelief of the

cover story ($n = 3$), misunderstanding of instructions ($n = 1$), technical problems ($n = 3$), missing questionnaire data ($n = 2$), not enough artifact-free EEG trials ($n = 3$), or because they fulfilled some exclusion criteria ($n = 2$). This resulted in data of 92 participants for the analyses (18 men and 74 women, mean age = 21.85 years, $SD = 3.25$ years). The power to detect correlations between personality measures and neural measures in each context group separately was .8 when assuming a correlation of at least .39, which can be expected according to former studies on social rejection (Cao et al., 2015). All participants provided written informed consent before participating in the study. The experiment was carried out according to the Declaration of Helsinki and approved by the ethics committee of the University of Lübeck.

General Study Overview

Participants were invited to take part in a study introduced to examine "how people get to know each other." This was meant to enhance the credibility of the social feedback paradigm, which was the central experimental paradigm, during which participants received the alleged feedback of peers in the laboratory while their EEG was being recorded. Before the actual measurement, participants were first asked to create personal profiles and to rate others' profiles with regard to the question whether they would want to meet each other in real life. They were also told that they would be given the contact information of those other participants with whom they had matching positive ratings. This was done to enhance participants' personal involvement. In the actual social feedback task (see below), participants were presented with the alleged feedback of their peers (whether they wanted to meet them or not, in the following referred to as acceptance or rejection). They were instructed that they would receive feedback only from those whom they themselves had given a positive rating to ensure their interest in the feedback. As the main experimental manipulation, participants were randomly assigned to one of two groups, of which one group received mainly rejection and the other group received mainly acceptance. In the following, the personal profiles and the social feedback paradigm are explained in detail.

Personal Profiles

Because of logistic reasons, participants did not rate the profiles of real other participants but 200 fake profiles that had been created by the experimenters. The profile consisted of an avatar profile picture (chosen from 60 picture options), a profile name, and the answers to two personal questions ("In my next life, I will..." and "What do you think is most overrated?"). The avatar profile picture options were created with freely available online software (<https://j0e.org/tools/avatar-generator/>). They were

created to look androgynous to reduce possible gender biases and had a wide variety of skin and hair colors. All had a mildly smiling expression (see examples in Figure 1A). Profile names were also created in a way that made a clear gender attribution difficult. The answers in the ostensible profiles were taken from real pilot participants, recombined and completed by the authors. Pilot ratings confirmed that they made the impression to be authentic.

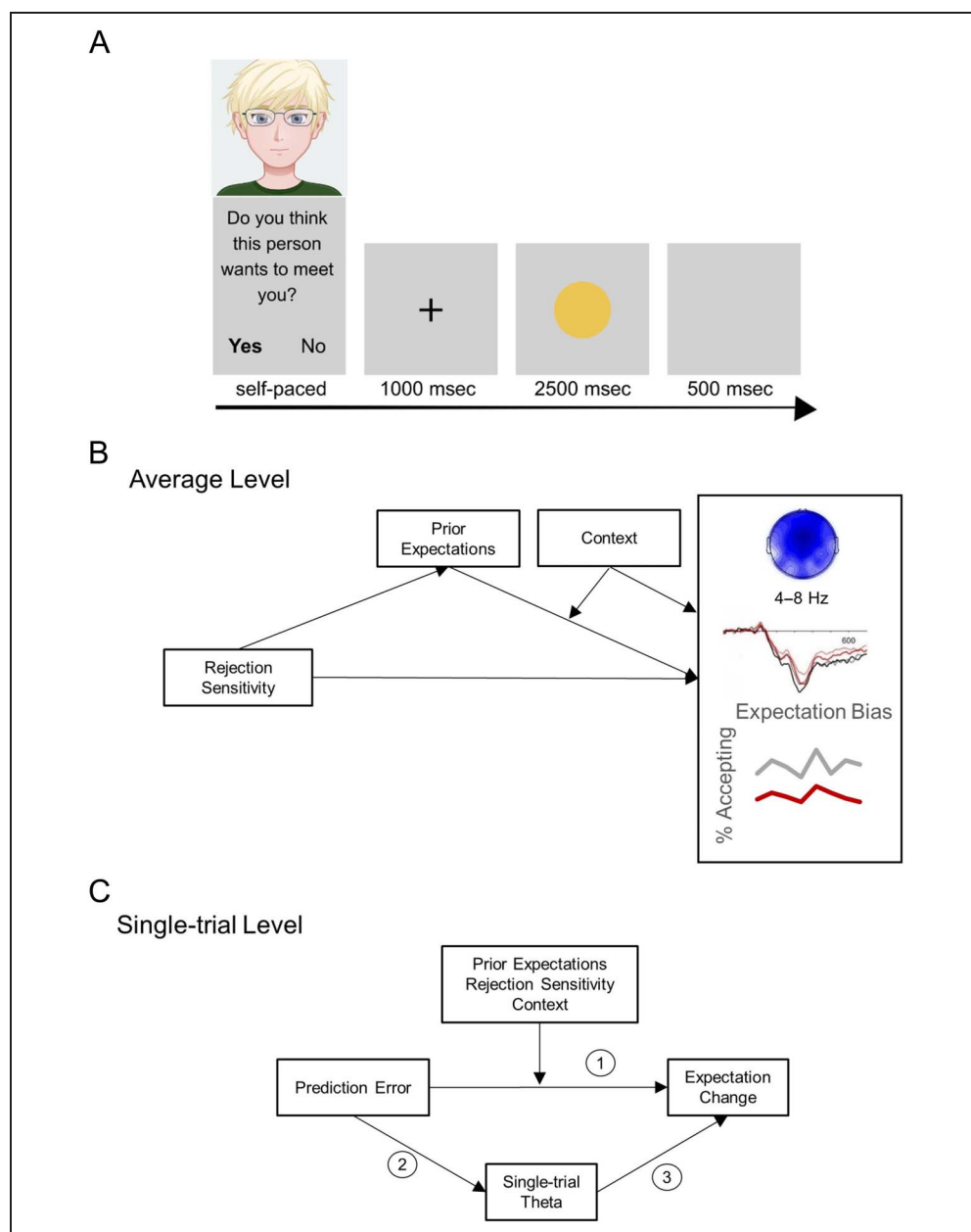
A few days before the laboratory experiment, participants created a profile (analogous to the profile described above, choosing their profile picture from the 60 picture options created by the experimenters) at home that was then ostensibly shown to and rated by other participants with regard to the question if they would like to meet them in real life. There were two answer options, yes and no, to create acceptance and rejection. At the same time, participants

themselves rated the profiles of 120 other ostensible participants, which had been created by the experimenters (80 additional profiles were rated later in the laboratory because of time constraints). Later in the laboratory task, participants were only shown the feedback of those other participants who they themselves wanted to meet according to their ratings. To achieve a sufficient number of trials, participants were therefore requested to rate at least half of the other profiles with “yes”. The creation of the profile and the rating of the others’ profiles were done using the online survey platform SoSciSurvey (Leiner, 2019).

Social Feedback EEG Paradigm

Participants were presented with the ostensible feedback of 100 other participating students. In each trial, the

Figure 1. (A) Trial structure of the social feedback paradigm. (B) Overview of analyses on the average data. (C) Pathways to be tested in the single-trial analyses.



profile of another ostensible participant was shown with the question, “Do you think this person wants to meet you?” and participants were instructed to press a button on the keyboard to indicate their feedback expectations (“yes” or “no”). There was no time restriction for providing the answer. After the button press, there was a fixation cross for 1000 msec, followed by the feedback that was presented for 2500 msec, then followed by an intertrial interval of 500 msec consisting of a gray screen. Acceptance and rejection was represented by orange and turquoise circles, respectively (the meaning of the colors was randomly assigned for each participant, resulting in a distribution of the color conditions of 51% vs. 49% in the negative context and 46.6% vs. 53.4% in the positive context). The feedback was presented centrally with an approximate visual angle of 5°. Participants were randomly assigned to one of two groups, one that received mainly rejections (negative context, 70% rejection, 30% acceptance) and one that received mainly acceptance (positive context, 30% rejection, 70% acceptance). The experiment was broken down into two blocks of 50 trials each. Between the two blocks, there was a short break to give participants the opportunity to relax their eyes. Trials were presented in random order. The paradigm lasted approximately 20 min.

Before the social feedback paradigm, participants stated their expectations, how many of the other participants (in percent) would want to meet them, on a visual analogue scale (in the following referred to as “prior expectations”). After the paradigm, participants were asked to estimate how many of the other participants had wanted to meet them (in percent) on the same visual analogue scale as before (referred to as “feedback estimations”).

During the task, participants sat in a dimly lit, sound-attenuated chamber, approximately 60 cm away from a 23-in. screen (Dell) on which feedback was presented. The task was programmed in Psychtoolbox (Version 3.0.14, Kleiner, Brainard, & Pelli, 2007) in MATLAB (Version 2015b, The MathWorks). For logistics reasons, some of the experiments were conducted in a different room, which was not sound-attenuated but otherwise very similar using a 27-in. screen (Dell) and MATLAB Version 2017b. The trial structure is depicted in Figure 1A.

Personality Questionnaires

To assess the rejection sensitivity of the participants, eight items from the German translation of the Rejection Sensitivity Questionnaire (Downey & Feldman, 1996) by Staebler et al. (2011) were used. These items consist of a short description of a situation, in which people might be rejected by others. Participants are asked how much they would expect to be rejected in the situation (expectancy scale) and how much they would worry about being rejected in the situation (anxiety scale), both on a 6-point Likert scale. We added a third question that asked how angry they would be about being rejected in the situation

(anger scale). To obtain scores for anxious and angry rejection sensitivity, the scores from the expectancy scale are multiplied with the scores of the anxiety and anger scales, respectively, and averaged over all situations. Because of time constraints, we did not use all items of the original Rejection Sensitivity Questionnaire but chose only those items that asked for situations with peer groups, unfamiliar peers, and friends (rather than romantic partners and relatives) as our study dealt with feedback from unfamiliar peers. Participants also filled in the German version of the Brief Fear of Negative Evaluation Scale (Reichenberger et al., 2016) and the Buss Perry Aggression Questionnaire (Buss & Perry, 1992; German Version: Herzberg, 2003). Moreover, before and after the social feedback paradigm, they answered a state emotion questionnaire that was developed for this study. These three questionnaires as well as the angry rejection sensitivity scores are not further evaluated here.

Experimental Procedure

As described above, participants rated 120 profiles at home, filled in the personality questionnaires (see the Personality Questionnaires section) and were then invited to the laboratory. To shorten the additional time needed to complete the rating of the total 200 profiles, participants rated another 80 profiles while the EEG recording was prepared. When the EEG equipment was ready for measurement, participants filled out an emotion questionnaire (not further evaluated here) and then underwent the social feedback paradigm. Afterward, participants filled in the same emotion questionnaire again. After the social feedback paradigm, participants performed a fight-or-escape paradigm (Beyer, Buades-Rotger, Claes, & Krämer, 2017), the results of which will be presented elsewhere. In the end, participants filled in a postexperimental questionnaire assessing their beliefs about the aim of the study (Questions were as follows: “Did you notice anything special about this part of the study?” and “What do you think was examined in this part of the study?”). Finally, participants were fully debriefed about the deception and the true aim of the study.

EEG Recordings

EEG data were recorded with 59 Ag/AgCl electrodes placed on an elastic cap according to the international 10–20 system using BrainVisionRecorder (Version 1.21.0102 or Version 1.20.081, BrainProducts GmbH). An online reference electrode was placed on the left earlobe, and an off-line reference electrode was placed on the right earlobe. Additionally, horizontal and vertical EOG was recorded with four electrodes placed next to the outer corners of the eyes and above and below the left eye. Sampling rate was 250 Hz, and data were recorded with an online high-pass filter of 0.016 Hz and a notch filter at 50 Hz. Impedances were kept below 5 k Ω .

EEG Data Analysis

All preprocessing and averaging was done in EEGLAB Version 14.1.2 and ERP-Lab Version 7.0.0 (Lopez-Calderon & Luck, 2014; Delorme & Makeig, 2004). Data were rereferenced to the right earlobe, and bipolar horizontal and vertical EOG channels were computed. Afterward, data were segmented into feedback-locked epochs of 3500-msec length (1000 msec before and 2500 msec after the feedback onset). Consistently bad channels (based on visual inspection) were interpolated (spherical interpolation, on average 0.5 channels per participant), and data segments with large artifacts were removed from the data. Next, an independent component analysis (implemented with the *runica* function in eeglab) was used for ocular artifact correction. Independent component analysis components that were clearly related to eye blinks or horizontal eye movements based on topography and time course were visually detected and removed (2.6 components on average per participant). Afterward, data were filtered with a 0.1-Hz high-pass filter and a 48-Hz low-pass filter (finite impulse response filter, filter order = 8250, hamming window). Next, data were baseline-corrected to the 1000 msec before stimulus onset. Nonocular artifacts were rejected using a semiautomatic procedure. A voltage threshold (between $-60/60 \mu\text{V}$ and $-85/85 \mu\text{V}$) was set for each participant in a way that all trials with artifacts were removed. If a large number of artifacts was caused by a single channel, the channel was interpolated (spherical interpolation, 0.14 channels per participant). The number of rejected trials varied from 0 to 54% per participant and valence and did not differ between conditions (means: negative context/rejection 16%, negative context/acceptance: 17%, positive context/rejection: 15%, positive context/acceptance: 18%).

Epochs time-locked to feedback presentation were averaged for each participant for the four conditions varying on the two factors valence (rejection vs. acceptance) and expectancy (expected vs. unexpected) separately. The P3 was calculated as the mean amplitude at CPz between 300 and 400 msec (Hajcak et al., 2010). Grand averages of the potentials were calculated for each of the context groups separately and additionally for the first and the second half of the paradigm.

For the time–frequency analysis, single-trial data of all electrodes were convolved with a complex Morlet wavelet as implemented in MATLAB (function *cut* with parameter specification “*cmor1-1.5*”):

$$\omega(t) = (\pi f_b)^{-0.5} e^{-\pi f_c t} e^{-\frac{t^2}{f_b}}$$

where $f_b = 1$ is the bandwidth parameter and $f_c = 1.5$ is the wavelet center frequency. Specifically, for each subject, changes in time-varying energy were computed and averaged (square of the convolution between wavelet and signal) in the frequencies (1–40 Hz, linear increase) for the 1500 msec after feedback onset with respect to a baseline of 500–50 msec before feedback onset (van der

Molen et al., 2017; Teolis, 1998). Power values were converted to decibel to make power values from different frequency bands comparable (Cohen, 2014).

Statistical Analyses

Statistical analyses were carried out using R Versions 3.5 and 4.0 (R Core Team, 2020) and MATLAB Version 18b. If assumptions for parametric tests were violated, non-parametric alternatives were used. In cases of multiple testing, Bonferroni correction was applied. All tests were carried out two-tailed with an alpha level of .05. First, analyses on the averaged behavioral and neural data are described, followed by the single-trial analyses for single-trial expectation changes and theta analyses (Figure 1B, C). All raw data and the code for the preprocessing of EEG data are available at https://osf.io/4j6c/?view_only=b56de4731a404a51a539f2e1966ec856.

Behavioral Data

As manipulation check, we tested if participants perceived the acceptance rate differently in the two contexts. To that end, the difference between prior expectations and the estimated acceptance rate after the paradigm were compared between the two contexts.

To assess possible biases in feedback expectations, a bias score was calculated by subtracting the actual frequency of acceptance (30% in the negative context, 70% in the positive context) from the frequency with which participants expected acceptance across trials. The bias scores were tested against zero using one-sample *t* tests for the two groups separately and compared between groups using independent *t* tests. Moreover, absolute bias scores were calculated for blocks of 10 trials. An ANOVA with the factors Context and Block was calculated to assess if possible differences in the bias scores were stable over the whole paradigm.

To test whether rejection sensitivity predicted the expectation bias during the feedback presentation, a linear regression model was used. To test whether this effect was mediated via the expectations before receiving the feedback, these were entered as a mediator into the model. As we assumed an interaction between prior expectations and the actually received feedback, the context was treated as a moderator of the relationship (Figure 1B).

ERP Data

To explore context group differences in P3 mean amplitude to feedback, a $2 \times 2 \times 2$ ANOVA with the within-subject factor Feedback Valence (rejection vs. acceptance) and Feedback Expectancy (unexpected vs. expected) and the between-subject factor Context Group (negative vs. positive) was conducted. To assess changes over time in the two groups, a $2 \times 2 \times 2$ ANOVA with the within-subject factor Valence and first versus second half of the paradigm (to maintain enough trials per condition, the task was only

separated in half) and the between-subject factor Context Group were computed. The factor Expectancy was dropped here because we had no specific hypotheses about the development of responses to unexpected versus expected feedback over time and to keep possible interactions interpretable. To assess if biases in expectations were reflected in changes in ERPs, correlation analyses were used: Residualized P3 amplitudes to rejection and acceptance in the first and the second half were obtained by calculating the residuals from regression models predicting each condition's amplitude from the other three conditions. These residualized P3 amplitudes were correlated with the expectation bias during the task, prior expectations, and rejection sensitivity. This was done to control for effects of the other variables (similar to a baseline correction) without the disadvantage of inflating the measurement error by calculating difference scores (Meyer, Lerner, Reyes, Laird, & Hajcak, 2017).

Time–Frequency Data

For the time–frequency data, we performed three analyses. (i) Average analyses: We tested for feedback valence and context differences in averaged power especially in theta but also across the frequency range of 1–25 Hz. (ii) Single-trial analyses: We assessed to what extent theta power predicted changes in expectations of acceptance or rejection. (iii) We did a control analysis to assess if theta effects were attributable to objective feedback probability only (single-trial analyses). We will explain each of these analyses in the following.

(i) Condition Effects on Theta

To assess power differences between rejection and acceptance and between the two contexts, a nonparametric permutation test was calculated for all time points from 500 msec before feedback onset until 1500 msec after feedback onset and for all frequencies between 1 and 25 Hz across all electrodes (Cohen, 2014). Nonparametric permutation tests are well suited to account for multiple comparison problems in testing EEG data (Maris & Oostenveld, 2007). A null hypothesis distribution of double difference values between the two contexts $((\text{rejection} - \text{acceptance})_{\text{neg-context}} - (\text{rejection} - \text{acceptance})_{\text{pos-context}})$ was created by randomly shuffling the context value and calculating difference values for this permutation and repeating this process 1000 times. The empirical difference value for a certain data point was considered significant if it was greater than the 97.5th or smaller than the 2.5th percentile of the distribution. To control for the large number of tests, a maximum value correction was used, which has been described to be more conservative than a cluster size correction (Cohen, 2014). In this way, we tested for significant differences in theta power (4–8 Hz) in the first 500 msec after feedback onset at frontocentral electrodes, where other studies found effects of valence (van der Molen et al., 2017, 2018). Moreover, we explored if there

were significant differences between rejection and acceptance in other frequency bands and time windows. If there were time–frequency clusters with significant differences, post hoc *t* tests on the average activity in these clusters (at the electrode with the largest effect) were performed. Finally, the residualized average activity was correlated with prior expectations, rejection sensitivity and expectation bias.

(ii) Theta Activity and Changes in Expectations

As laid out in the Introduction, we expected that the theta response to PEs relates to an adjustment of expectations afterward. Specifically, we expected PEs to lead to an expectation change mediated by theta power changes. We thus set up a series of (generalized) linear mixed models to test each path in the hypothesized model (Figure 1C). In all of these models, the single trials were treated as Level 1 units, which were nested in the participants (Level 2 units). First, we tested whether PEs predicted a change in expectations from one trial to the next and whether this was moderated by context, prior expectations, or rejection sensitivity (Pathway 1; Figure 1C). Second, we tested if theta was predicted by PE, possibly moderated by context, prior expectations, and rejection sensitivity (Pathway 2). Third, we tested whether theta itself predicted changes in expectations (Pathway 3). If all three pathways showed significant effects, this would suggest that theta responses relate to the mechanism that triggers changes in expectations.

An “expectation change” in trial *t* was coded as 1 if the expectation in trial *t* was different from *t* – 1 and otherwise coded as 0. The PE in trial *t* was coded as positive PE if the participant had expected rejection but received acceptance, as negative if the participant had expected acceptance but received rejection, and as zero if the expectations were fulfilled. (Note that this differs from classic operationalizations of PE, as those usually do not include subjective expectations directly but infer them from choice behavior. We chose this operationalization as it required no assumptions about how explicitly expressed expectations relate to underlying gradual expectations.) As this resulted in a factor with three levels, it was dummy-coded, and zero PE was used as a reference category. Note also that the PE combines the factors feedback valence and expectation (similar to other EEG studies on social feedback; e.g., van der Molen et al., 2017), and the linear mixed models therefore assess effects of expectations on neural measures, which were ignored in the analysis of averaged data (analysis (i)). For the theta response, we used the averaged theta power (4–8 Hz) in single trials in the time window in which significant differences between negative and positive contexts occurred in the permutation test at the electrode where the effect was largest. Single-trial power was calculated as explained above but without baseline correction, as differences in mean power between participants are controlled for by the random intercepts in the model.

To find the best model for each pathway, model averaging of (generalized) linear mixed models was employed. Model averaging helps to avoid difficult manual model selection processes and thereby to select more robust models in cases where a large number of models is plausible (Burnham & Anderson, 2004). Using the *MuMIn* (Barton, 2020) and the *lme4* (Bates et al., 2020) package in R, the best averaged model for each of the three different pathways was fitted as follows: (a) testing Pathway 1, (b) predicting theta power from the variables contained in the averaged model from Pathway 1, and (c) predicting changes in expectations from variables contained in 1 and 2.

For each pathway, a full model containing all plausible interactions and their subordinate effects was constructed. Random effects were first optimized for a model containing all fixed effects (Zuur, Ieno, Walker, Saveliev, & Smith, 2009). Then, each plausible subordinate model (leaving out one or more predictors from the full model) was fitted. Models were then compared using the Akaike information criterion (AIC), and all models that had an AIC difference of 2 or less from the best fitting model were averaged (Burnham & Anderson, 2004). The averaged parameters were then interpreted. For Pathway 1, a threefold interaction between prior expectations, PE, and context, as well as a threefold interaction between rejection sensitivity, PE, and context were entered as plausible predictors in the full model. The predictors for the other pathways resulted from this first analysis as explained above.

Predictive model fit of the averaged models was calculated in three ways: First, the number of trials predicted correctly was computed by binarizing predicted change (probabilities to change $> .5 \sim$ *change predicted*, probabilities $< .5 \sim$ *no change predicted*; Rana, Midi, & Sarkar, 2010). Second, the average absolute difference between the predicted probability to change expectations and the actual value of the trial ($1 =$ *change*, $0 =$ *no change*) was calculated as a more precise measure of model fit. For the models predicting theta activity, the correlation between measured and predicted theta activity was calculated as a measure of prediction accuracy (note that this is similar to R^2 ; Kvålseth, 1985). To test if these predictive fit measures were greater than chance, a null hypothesis distribution of model fits was calculated by randomly shuffling the outcome variable and calculating the same averaged models and their predictive fit 500 times. If the percentage of correctly predicted trials in the real model was greater than the 95th percentile of the null hypothesis distribution and the average absolute difference was smaller than the 5th percentile of the null hypothesis distribution, predictive model fit was considered greater than chance.

(iii) Control Analyses for Theta Activity

As the PE was confounded with the objective probability of the feedback, a control analysis was run to test whether theta activity could be better explained by this predictor than by PE. To this end, linear mixed models containing both PE and objective probability were calculated and

averaged as explained above. Again, trials (Level 1) were nested in participants (Level 2). Objective probability was calculated as the cumulative objective probability of the feedback in a given trial. To get a stable estimate of the first 10 trials, we added 10 artificial trials (50% accepting) previous to the actual feedback trials, mimicking a starting value of 0.5. To ensure the robustness of the findings, the same analyses were repeated with slightly alternative approaches for calculating the objective probability: cumulative probability without the artificial trials, objective probability smoothed over blocks of 10 trials, and objective probability smoothed over blocks of 10 trials with 10 artificial trials (50% accepting at the beginning). The interaction between objective probability and context was added in addition to the predictors revealed by the main analyses explained above.

RESULTS

Randomization and Manipulation Checks

Prior expectations ($Mean_{neg} = 42$, $SD = 17.2$; $Mean_{pos} = 41.7$, $SD = 16.6$; see Figure 2A, box plots on the left side of each context) and rejection sensitivity ($Mean_{neg} = 10.16$, $SD = 3.95$; $Mean_{pos} = 10.18$, $SD = 3.80$) were both normally distributed and showed a broad range in both context groups. There were no differences between participants in the context groups in expectations, $t(89) = 0.09$, $p = .93$, $r = .009$ (Figure 2A) or rejection sensitivity before the experiment, $t(90) = 0.03$, $p = .98$, $r = .003$. In the negative context, 31.9% of participants were currently in a romantic relationship, 61.7% were single, and 6.4% stated that they did not know their relationship status. In the positive context, 55.6% of participants were currently in a romantic relationship, 40% were single, and 4.4% stated that they did not know their relationship status. The change from prior expectations to post feedback estimations differed significantly between the two contexts. Participants in the negative context stated that they had received less acceptance than expected, and participants in the positive context stated that they had received more acceptance than expected ($M_{neg} = -18.91$, $SD = 13.9$; $M_{pos} = 20.96$, $SD = 17.6$), $t(83) = -12.00$, $p < .001$, $r = .76$ (see Figure 2A, right box plot in each context for mean acceptance estimations after the paradigm).

Participants in the negative context expected significantly less often acceptance than participants in the positive context already during the task, $t(77.8) = -5.80$, $p < .001$, $r = .55$ (Figure 2A, box plots in the middle of each context). However, the comparison of the bias scores between the two groups showed that participants in the negative context maintained a positive bias ($M = 14.79$, $SD = 14.34$; one-sample t test: $t(46) = 7.10$, $p < .001$) and participants in the positive context showed a negative bias ($M = -10.49$, $SD = 9.05$; one-sample t test: $t(44) = -7.70$, $p < .001$) and that these scores differed significantly, $t(78) = 10.20$, $p < .001$, $r = .75$ (Figure 2B).

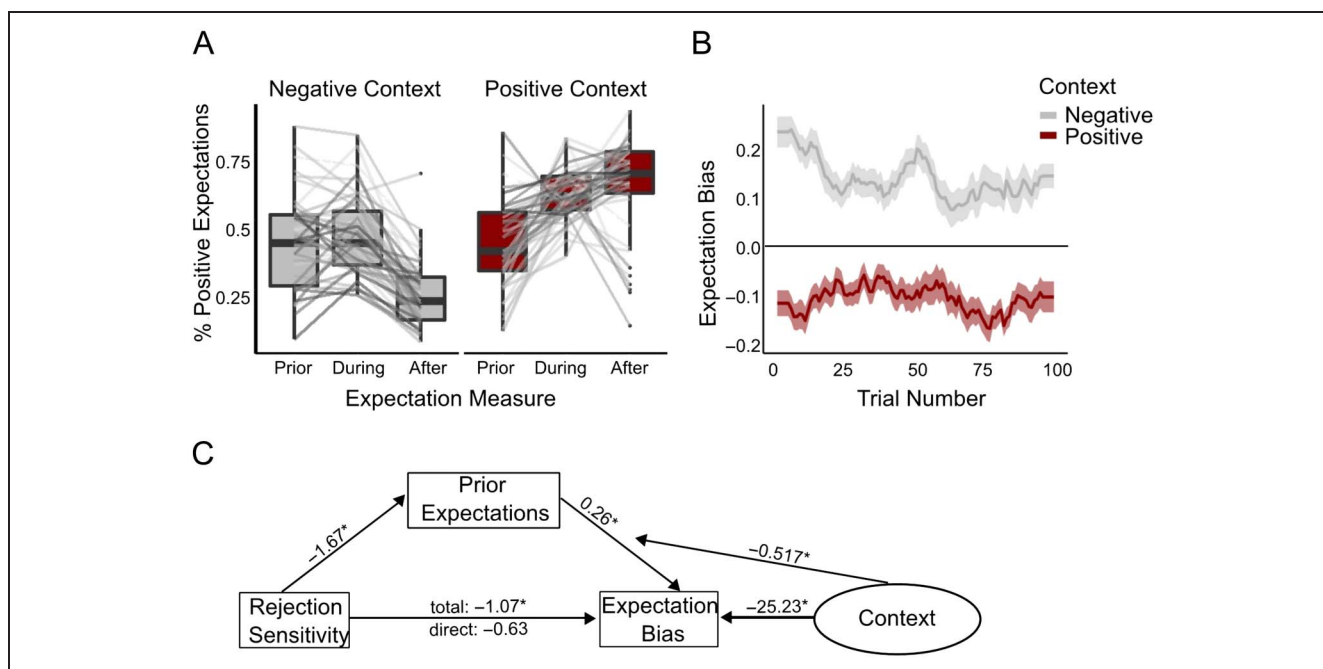


Figure 2. (A) Expectations of acceptance prior, during, and estimations of received acceptance after the feedback presentation separately for the two contexts. Gray lines depict scores of single participants. (B) Mean smoothed expectation bias (expected – received acceptance) over time in the two contexts. (C) Results of the moderated mediation models predicting bias scores. Numbers are the unstandardized betas. $*p \leq .001$.

Moreover, an ANOVA on the time course of the absolute bias scores, averaging responses in 10 blocks of 10 trials each, showed that the effect of Context was significant throughout all blocks, $F(1, 89) = 4.98, p = .028, \eta_g^2 = .021$, and there was no interaction between Context and Block, $F(9, 801) = 0.77, p = .60, \eta_g^2 = .005$ (Figure 2B). This result shows that participants in the negative context overestimated the probability of acceptance more than the participants in the positive context underestimated it. Moreover, it ensures that participants' bias in feedback perception persisted across the experiment and did not disappear with learning.

Prediction of Feedback Expectations by Rejection Sensitivity and Prior Expectations

The moderated mediation model (Figure 2C) showed that rejection sensitivity negatively predicted the expectation bias during the task (total effect = $-0.92, SE = 0.32, t(88) = -2.92, p = .004$). Rejection sensitivity also predicted expectations before the task (a-path: $b = -1.64, SE = 0.43, t(88) = -3.82, p < .001$, which in turn predicted the expectation bias during the task (b-path: $b = 0.28, SE = 0.07, t(88) = 3.85, p < .001$). Crucially, the indirect effect of rejection sensitivity over prior expectations on bias scores was also significant, indicating a mediated effect of rejection sensitivity (beta = $-0.46, SD = 0.17, CI [-0.81, -0.15]$). The direct effect of rejection sensitivity on the bias scores became nonsignificant when considering the mediator, indicating a full mediation ($b = -0.46, SE = 0.32, t(87) = -1.46, p = .15$). Participants with higher rejection sensitivity thus expected less acceptance

in the beginning, which in turn led to a negative expectation bias during the experiment (Figure 2C).

The context also predicted the bias score (total effect = $-25.26, SE = 2.42, t(88) = -10.44, p < .001$, but did not moderate the mediated effect of rejection sensitivity (direct effect = $0.36, SE = 0.59, t(87) = 0.61, p = .54$; indirect effect: beta = $0.06, SD = 0.22, CI [-0.39, 0.53]$). However, a separate moderation model showed that the b-path (predicting bias scores from expectations before the task) was moderated by context ($b = -0.52, SE = 0.12, t = -4.16, p < .001$). This indicated that prior expectations influenced the expectation bias during the task stronger in the negative than in the positive context (Figure 2C).

Condition Effects and Interindividual Differences in the P3

The ANOVA on the P3 amplitude between 300 and 400 msec (Figure 3, top) showed a significant main effect of Valence, $F(1, 90) = 11.90, p < .001, \eta_g^2 = .007$, and a significant interaction between Valence and Context, $F(1, 90) = 101.80, p < .001, \eta_g^2 = .054$. All other main effects and interactions were not significant (all $p > .121$, all $\eta_g^2 < .020$). Post hoc ANOVAs on the interaction between Valence and Context showed that, in the negative context, the P3 was lower after rejection than after acceptance, $F(1, 46) = 108.00, p < .001, \eta_g^2 = .109$, whereas in the positive context, the P3 was reduced after acceptance, $F(1, 44) = 18.90, p < .001, \eta_g^2 = .025$. Moreover, the P3 amplitude for rejection did not differ between contexts, $F(1, 90) = 4.59, p = .07, \eta_g^2 = .049$, whereas the P3 amplitude to acceptance was significantly higher in the negative than in the

positive context, $F(1, 90) = 7.06, p = .018, \eta_g^2 = .073$ (see Figure 3B, bottom). In summary, P3 amplitudes were differently sensitive to valence in the two contexts, which was mainly due to different reactions to acceptance. Note also that the effect size of the difference between rejection and acceptance in the negative context was four times higher than the effect size in the positive context. Measured in microvolt, the mean difference in the negative context was more than twice as high ($-4.5 \mu\text{V}$) as in the positive context ($2.2 \mu\text{V}$), demonstrating stronger neural differentiation between rejection and acceptance in the negative context.

The ANOVA including the factor Time revealed an interaction between Time and Condition, $F(1, 90) = 5.38, p = .023, \eta_g^2 = .004$. The main effect of Time, $F(1, 90) = 3.00, p = .085, \eta_g^2 = .002$, and all other interactions with Time (all $p > .26$, all $\eta_g^2 < .001$) were not significant. The effect of Time was present for the positive context, $F(1, 44) = 6.48,$

$p = .028, \eta_g^2 = .012$, but not for the negative context, $F(1, 46) = 0.22, p > 1, \eta_g^2 = .0003$. The effect in the positive context was caused by smaller P3 amplitudes in the second half of the task, indicating that in the positive context amplitudes decreased over time, whereas they remained stable in the negative context. None of the residualized P3 amplitudes were significantly correlated with expectation bias, prior expectations, or rejection sensitivity (all $r < .25$, all $p > .1$). Thus, the P3 did not reflect any of the interindividual differences in behavior.

Time-Frequency Differences in Feedback Processing between the Two Context Groups

When comparing the feedback valence differences between the two context groups, we observed a significant difference at 7–9 Hz from 100 to 348 msec at frontocentral electrodes (largest effect at FCz; see Figure 4A, top left).

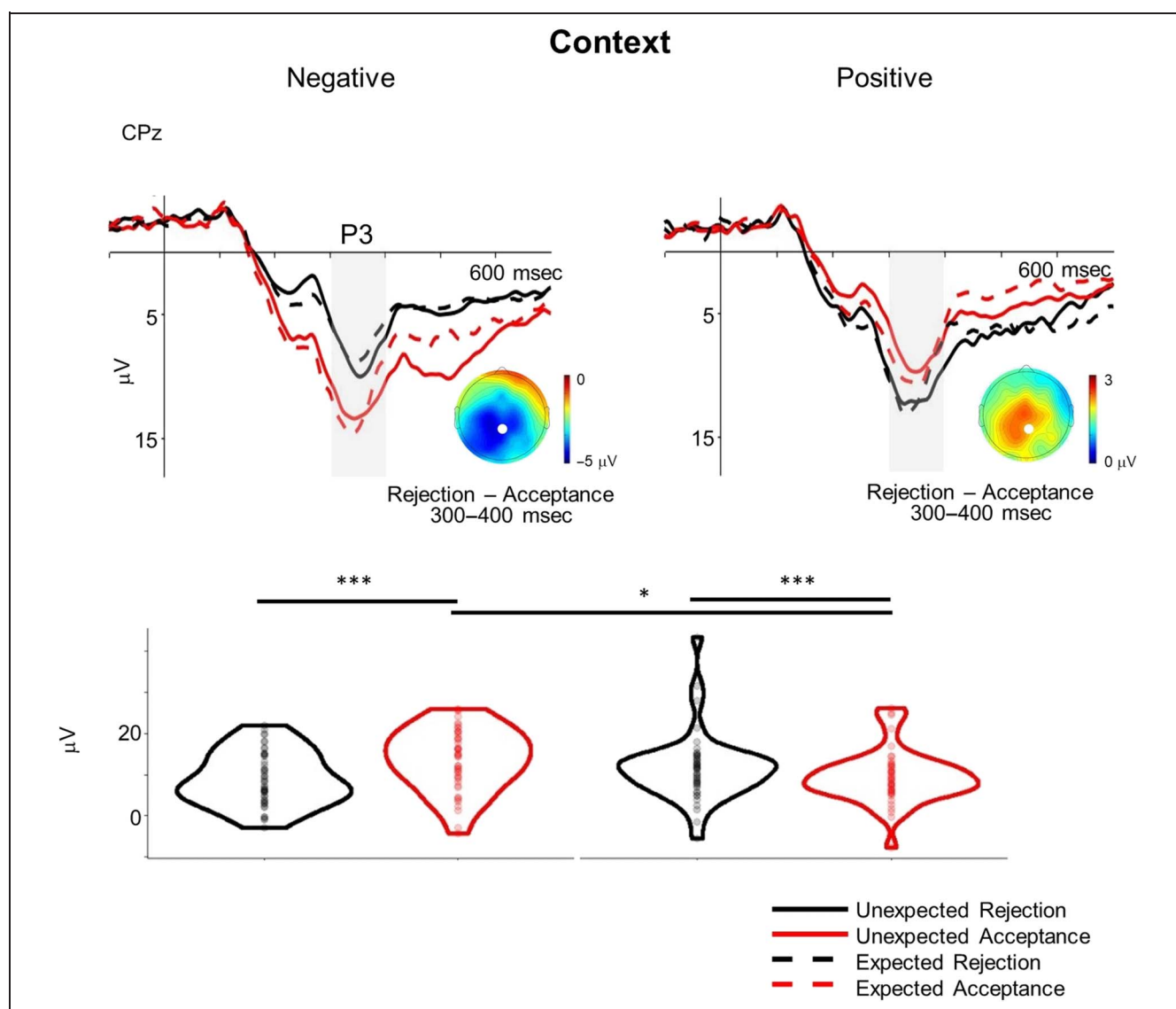


Figure 3. Grand averages of the P3 for rejection and acceptance in the two contexts. * $p < .05$, ** $p < .01$, *** $p < .001$. The gray-shaded area depicts the analyzed time window. Analyzed electrodes are marked in white.

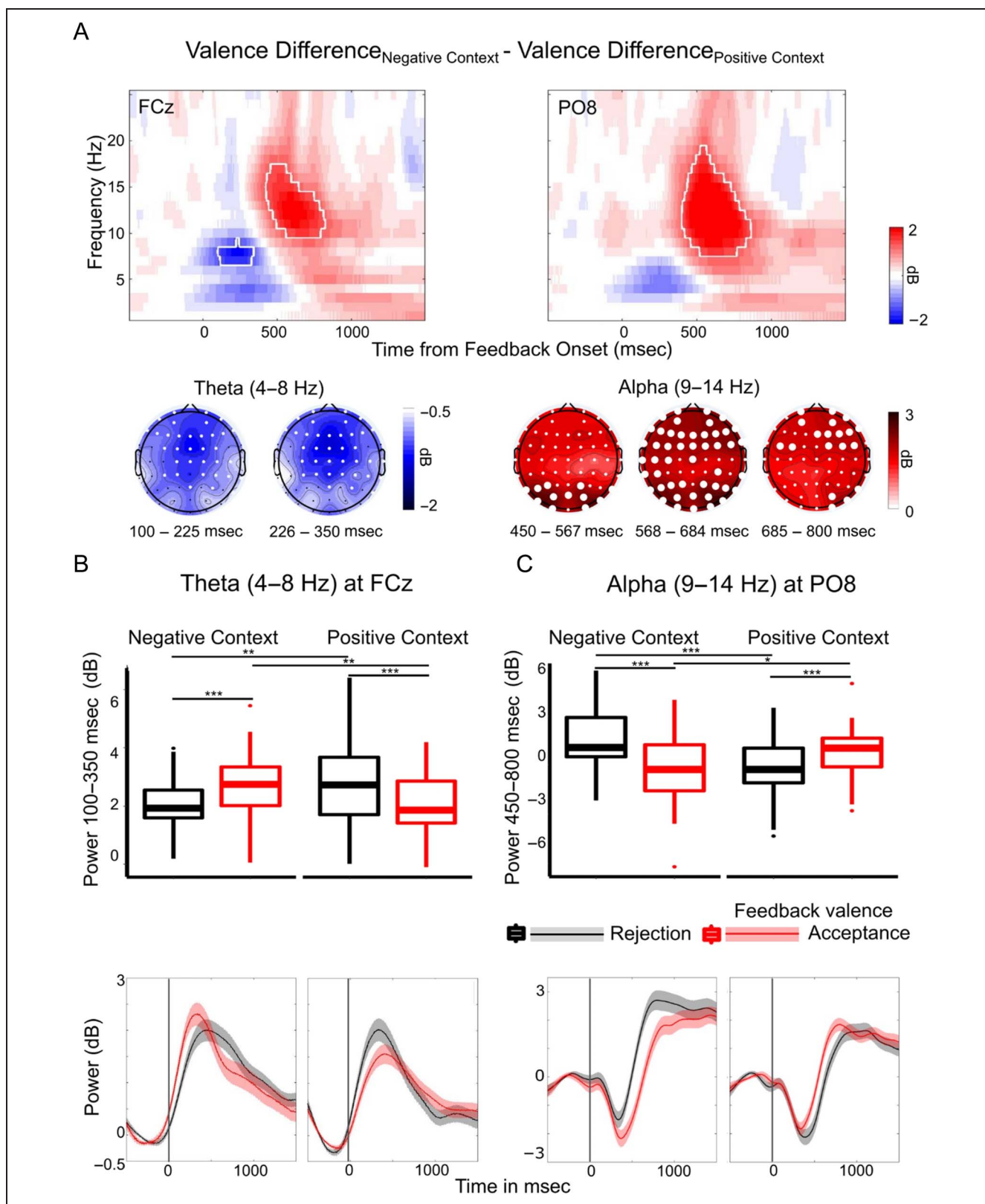


Figure 4. (A) Results of the permutation test for the time–frequency analyses for FCz and PO8. Time–frequency clusters surviving the maximum value correction are outlined in white. Bottom: Topography of context differences (negative – positive) in the power difference (dB) between rejection and acceptance for theta (left) and alpha (right) activity. Electrodes with significant differences in the whole time window are depicted in white (lenient correction, significant differences in any of the frequencies) or in white and fat (strict correction, average over the frequencies is significant in the whole time window). (B and C) Box plots of the average theta (B) and alpha (C) activity in the two contexts. Time course of theta (B) and alpha (C) power for rejection and acceptance. Shaded areas depict standard errors.

Theta power was higher for acceptance than rejection in the negative context, with the opposite pattern in the positive context. Between 372 and 1044 msec, rejection yielded higher power than acceptance in the 6–21 Hz range in the negative context, with an opposite pattern in the positive context. This effect was spread over all electrodes with the largest effect at PO8 (see Figure 4A, top right). We also observed a significant difference at 4 Hz from 136 to 480 msec that was maximal at C4, with higher values for acceptance than rejection in the negative context, with the opposite pattern in the positive context.

Based on these results, our hypotheses, and on canonical definitions of frequency bands, two time–frequency clusters were chosen for further post hoc and single-trial analyses (see Methods): a theta cluster (4–8 Hz, 100–350 msec; see Figure 4A, bottom left) at FCz and an alpha cluster (9–14 Hz, 450–800 msec; see Figure 4A, bottom right) at PO8. *t* Tests on the average theta and alpha activity in these clusters (Figure 4B) showed significantly less theta power for rejection than for acceptance in the negative context, $t(46) = -5.10, p < .001$, Cohen's $d = -0.66$ (Figure 4B, left) and significantly more for rejection than for acceptance in the positive context, $t(44) = 5.58, p < .001$, Cohen's $d = 0.5$ (Figure 4B, right). Comparing each valence between the contexts showed that in response to rejection, participants in the negative context showed significantly less theta power than participants in positive context, $t(68.9) = -2.80, p = .007$, Cohen's $d = -0.60$, whereas in response to acceptance, they showed significantly more, $t(89.7) = 2.94, p = .004$, Cohen's $d = 0.62$ (Figure 4B, top). In summary, in the negative context, rejection trials elicited weaker early frontal theta activity than acceptance trials, whereas participants in the positive context showed the opposite pattern.

The residualized average theta power assessed separately for acceptance, and rejection was not correlated with the expectation bias or rejection sensitivity (all $r < .28$, all $p > .069$). Prior expectations were also uncorrelated with theta activity in acceptance trials in the negative context ($r = .09, p = .554$), but in the positive context there was a negative correlation of theta power to acceptance trials with prior expectations ($r = -.46, p = .002$), meaning that participants who had expected more acceptance from the beginning showed weaker theta responses to that feedback. A Fisher's z test showed that these correlations differed significantly from each other ($z = 2.72, p = .003$). The theta response to rejection did not correlate with prior expectations in either context (negative context: $r = -.06, p = .67$; positive context: $r = .010, p = .94$). The correlation between prior expectations and theta to acceptance in the positive context was also significantly higher than the correlation between prior expectations and theta to rejection ($z = -2.04, p = .041$). Summarizing, theta power was largest for the infrequent, unexpected feedback condition in both contexts, and interindividual

differences in prior expectations led to differential theta responses to acceptance in the positive context.

Looking at alpha power (Figure 4C), rejection elicited significantly stronger responses than acceptance in the negative context, $t(46) = 7.71, p < .001$, Cohen's $d = 0.88$ (Figure 4C, left), but the opposite was true in the positive context, $t(44) = -4.55, p < .001$, Cohen's $d = -0.55$ (Figure 4C, right). Alpha power in response to rejection was significantly higher in the negative than in the positive context, $t(89.3) = 4.60, p < .001$, Cohen's $d = 0.97$, whereas in response to acceptance, it was significantly lower, $t(89.4) = -2.39, p = .019$, Cohen's $d = -0.5$ (Figure 4C, top). Looking at the alpha time course (Figure 4C, bottom) revealed that both groups showed first a power decrease and then an increase, which differed between rejection and acceptance. In summary, the negative context showed a stronger alpha increase and less theta activity for rejection compared with acceptance, whereas the positive context showed the opposite effect.

Models for the Relationship between PE, Theta Power, and Changes in Expectations

We had hypothesized that theta would be increased for PEs and would relate to changes in expectations elicited by PEs. The first theta results were consistent with this, as theta was differentially modulated by feedback in the two contexts. We thus proceeded to formally test our hypothesis using (generalized) linear mixed models.

Averaged Model for Pathway 1

We first tested on the behavioral level, if PEs triggered changes in expectations, possibly moderated by context, prior expectations, and rejection sensitivity. The optimization of random effects showed that a model with random slopes for the PE was better than one without (AIC difference = -90); therefore, random slopes for PE were included in all models (note that these are not calculated for the averaged model though). Two models had an AIC difference smaller than 2 to the best model, and these three models were averaged (see Table 1, Rows 1–3).

The model average is summarized in Table 2 (Rows 1–14). The averaged model contained a significant threefold interaction between prior expectations, negative PE, and context (Table 2, Row 13). To understand this interaction better, the model was calculated again for the two contexts separately (expectation change depending on prior expectations, PE, and context plotted in Figure 5A). In the negative context, positive ($b = 0.33, SE = 0.14, z = 2.41, p = .016$) as well as negative ($b = 0.76, SE = 0.15, z = 4.94, p < .001$) PEs predicted more changes in expectations than zero PEs. However, participants with more negative prior expectations changed their expectations more often after negative PEs ($b = -0.43, SE = 0.09, z = -4.83, p < .001$; Figure 5A, left). In addition, participants with more

Table 1. Overview over Averaged Models for the Three Pathways

<i>Parameters Included in Model</i>	<i>df</i>	<i>AIC</i>	<i>Weight (in Average)</i>	<i>Weight (in Full Model Set)</i>
<i>Model Pathway 1</i>				
Context, PE, prior, context × PE, context × prior, PE × prior, context × PE × prior	18	10080.85	0.49	0.39
Context, PE, RS, prior, context × PE, context × RS, context × prior, PE × prior, context × PE × prior	20	10081.69	0.32	0.26
Context, PE, RS, prior, context × PE, context × prior, PE × prior, context × PE × prior	19	10082.66	0.20	0.16
<i>Model Pathway 2</i>				
Context, PE, RS, context × PE, context × RS	15	39609.16	0.72	0.42
Context, PE, prior, RS, context × PE, context × RS	16	39611.07	0.28	0.16
<i>Model Pathway 3</i>				
Context, theta, context × theta	5	10436.85	0.70	0.42
Context, theta, context × theta, prior	6	10438.52	0.30	0.18

PE = PE in trial $t - 1$; prior = prior expectations; RS = rejection sensitivity.

positive prior expectations changed their expectations generally more often ($b = 0.12$, $SE = 0.05$, $z = 2.64$, $p = .008$). In contrast, participants in the positive context changed their expectations only more often after positive than after zero PEs ($b = 1.02$, $SE = 0.13$, $z = 8.11$, $p < .001$), whereas all other factors did not predict changes in expectations (see Figure 5A, right).

In summary, participants in the two contexts used PEs, that is, subjectively unexpected feedback, differently to change their expectations. Specifically, participants in the negative context used negative as well as positive PEs to guide their expectation changes, whereas in the positive context, they used only positive PEs. Importantly, although interindividual differences had less of an impact in the positive context, the usage of PEs depended on interindividual differences in expectations before the task in the negative context (see Figure 5A). In the Pathway 2 model, we therefore explored possible neural underpinnings of these differences by testing whether theta responses to PEs were modulated in a similar way as changes in expectations by context and prior expectations.

Averaged Model Pathway 2

As detailed above, we hypothesized that the same predictors that predicted expectation changes should also predict theta responses. Therefore, the predictors from the averaged model from Pathway 1 were entered in the full model to predict theta activity (Table 1, Rows 4–5). Again, a full model containing random slopes for PE was better

than one without (AIC difference = -26); therefore, random slopes for PEs were included. The averaged model is summarized in Table 2 (Rows 15–23), which included a positive effect of positive PEs on theta responses and an interaction of context with both types of PE.

To test if this interaction between PE and condition (see Figure 5B) was solely attributable to the different objective probability of the feedback (as this influenced also the probability of specific PEs) in the two contexts, we conducted a single-trial control analysis with objective probability (see Methods), PE, and the interaction of both variables with context as predictors for theta activity. Objective probability was group mean centered.

Again, optimization of random effects showed that the best random effects structure contained random slopes for PE. The averaged model showed that negative PEs and objectively less probable feedback predicted stronger theta activity (see Table 3 for summary of models included in the average and Table 4 for parameters of the averaged model). As the effect of positive PEs and the interaction between PE and context disappeared, the control analysis suggests that these effects were only due to the different objective probability of PEs in the two contexts. However, the results also show that negative PEs predicted theta responses in addition to objective probability (as the effect became significant when including this predictor).

The influence of PEs on theta was different than their influence on behavioral changes in expectations: In both contexts, only negative PEs triggered a theta increase (see Table 4, Rows 3–4). In contrast, positive PEs

Table 2. Model Terms of the Averaged Models (Full Average)

<i>Fixed Effects</i>	<i>b</i>	<i>SE</i>	<i>z</i>	<i>p</i>
<i>Model Pathway 1</i>				
Intercept	−0.24	0.18	1.32	.19
Context	0.01	0.11	0.12	.90
PE-negative _{<i>t</i>−1}	1.33	0.32	4.22	<.001
PE-positive _{<i>t</i>−1}	−0.39	0.29	1.32	.18
Prior	0.24	0.11	2.20	.028
RS	−0.15	0.30	0.49	.62
Context × PE-negative _{<i>t</i>−1}	−0.59	0.20	2.90	.004
Context × PE-positive _{<i>t</i>−1}	0.71	0.18	3.87	<.001
Context × prior	−0.11	0.07	1.60	.11
Context × RS	0.12	0.21	0.56	.58
Prior × PE-negative _{<i>t</i>−1}	−0.81	0.19	4.29	<.001
Prior × PE-positive _{<i>t</i>−1}	0.17	0.18	0.94	.35
Context × prior × PE-negative _{<i>t</i>−1}	0.38	0.12	3.10	.002
Context × prior × PE-positive _{<i>t</i>−1}	−0.09	0.11	0.81	.42
<i>Model Pathway 2</i>				
Intercept	0	0	–	–
Context	0.08	0.07	1.23	.22
PE-negative _{<i>t</i>}	0.00	0.04	0.013	.98
PE-positive _{<i>t</i>}	0.18	0.04	4.16	<.001
Prior	−0.03	0.06	0.47	.64
Context × PE-negative _{<i>t</i>}	0.08	0.04	2.00	.046
Context × PE-positive _{<i>t</i>}	−0.15	0.05	3.29	<.001
RS	−0.23	0.21	1.10	.27
Context × RS	0.30	0.20	1.49	.14
<i>Model Pathway 3</i>				
Intercept	−0.02	0.06	0.41	.68
Context	0.11	0.08	1.33	.18
Theta	0.03	0.01	3.22	.001
Context × theta	−0.03	0.01	2.58	.01
Prior	0.04	0.15	0.29	.78

PE-negative_{*t*−1} = effect of negative versus zero PE in previous trial; PE-positive_{*t*−1} = effect of positive versus zero PE in trial *t*−1; prior = prior expectations; RS = rejection sensitivity.

triggered changes in expectations in both contexts (see Figure 5A). Negative PEs lead to expectation changes only in the negative context. Hence, effects of PE on theta and expectation changes were comparable only in

the negative context. Therefore, in the Pathway 3 model, we tested the hypothesis that the theta response to negative PEs predicts changes in expectations only in the negative context.

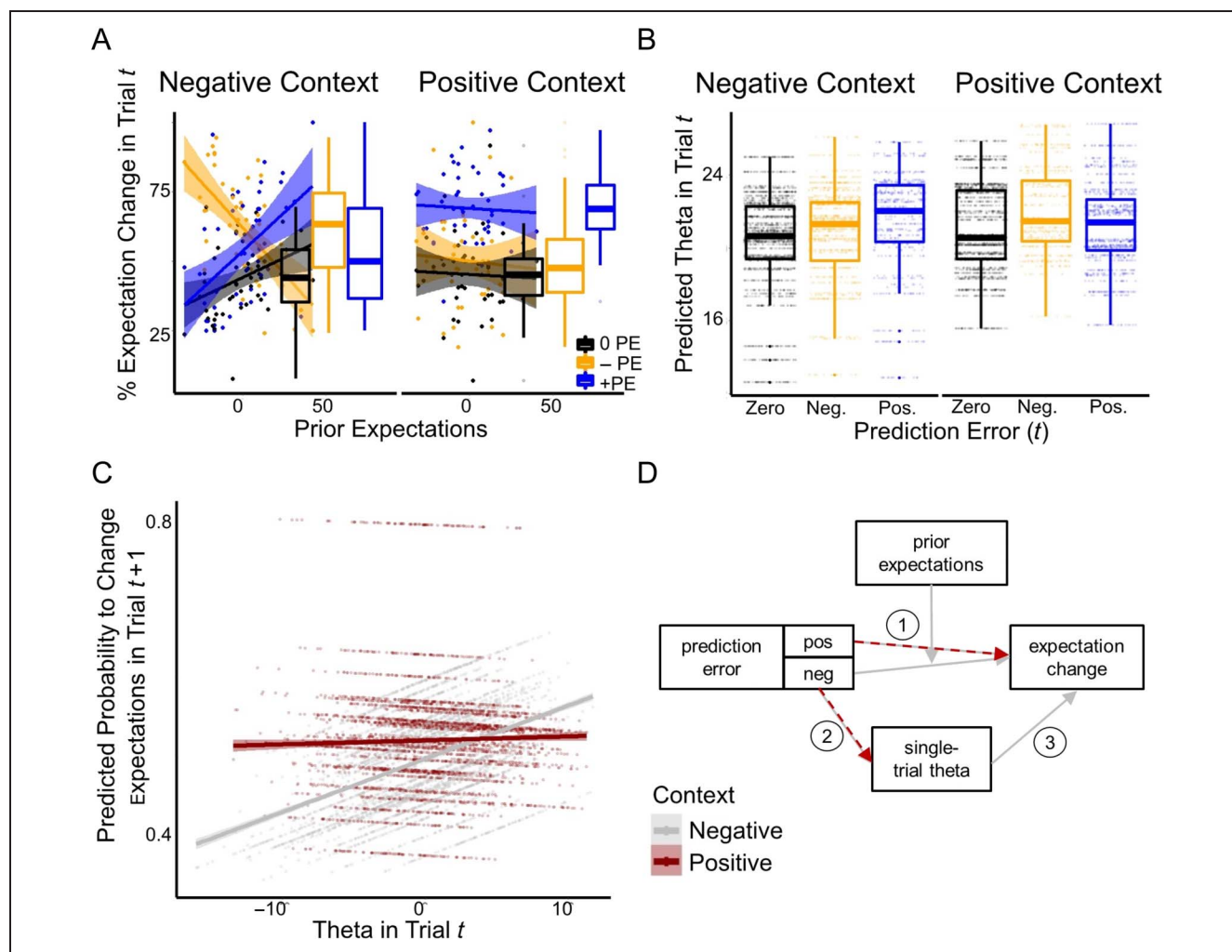


Figure 5. (A) Pathway 1 results: Percentage of trials t in which expectations were changed, depending on prior expectations, PE in trial $t - 1$, and context. (B) Pathway 2 results: Model predicted theta activity in trial t . (C) Pathway 3 results: Model predicted probability to update in trial $t + 1$ for the two contexts. The lines depict the mean in the two contexts. (A–C) Dots display values of single participants and trials. (D) Summary of the results of the three pathway models: Arrows from one variable to the other depict significant positive effects in the final model. Arrows pointing toward another arrow depict a significant moderating effect of the variable at the arrow's origin.

Averaged Model Pathway 3

For Pathway 3, we thus used a model with a threefold interaction between prior expectations, theta, and context as the full model, as it was hypothesized that the extent to which theta is relevant for changes in expectations depends on context and prior expectations. No random

slopes for theta were included, as the full model with random slopes had a worse model fit than the one without (AIC difference = 1.7). The averaged model was calculated from two models (overview in Table 1, Rows 6–7 and results from averaged model in Table 2, Rows 24–28). This averaged model showed that theta predicted expectation changes in the negative context only ($b =$

Table 3. Overview over Averaged Models for Theta Control Analyses

Parameters Included in Model	df	AIC	Weight (in Average)	Weight (in Full Model Set)
Context, objective probability, PE, context \times objective probability	13	39571.50	0.53	0.49
Objective probability, PE	11	39573.04	0.24	0.23
Context, objective probability, PE	12	39573.14	0.23	0.22

Table 4. Parameters of the Averaged Linear Mixed Model Predicting Theta Activity

Predictor	<i>b</i>	<i>SE</i>	<i>z</i>	<i>p</i>
Intercept	20.59	0.37	55.13	< .001
Objective probability	-2.33	0.45	5.31	< .001
PE-negative _{<i>t</i>}	0.63	0.14	4.66	< .001
PE-positive _{<i>t</i>}	0.24	0.14	1.70	.090
Context	0.47	0.52	0.88	.38
Context × objective probability	0.56	0.72	0.79	.43

PE-negative_{*t*} = effect of negative versus zero PE; PE-positive_{*t*} = effect of positive versus zero PE.

0.029, *SE* = 0.009, *z* = 3.31, *p* < .001; positive context: *b* = -0.006, *SE* = 0.010, *z* = -0.57, *p* = .57) but that prior expectations did not moderate this effect (Figure 5C). This suggests that theta power indeed drives changes in expectations, but only in a rejecting context and unrelated to prior expectations. Lastly, we calculated the predictive model fit for all the averaged models (see Methods), which was found to be better than chance for all models (see Table 5).

Taken together, the models provide evidence that PEs are relevant for changes in expectations, but differently so in a rejecting versus an accepting context, and that the use of PEs is modulated by interindividual differences only in a rejecting context. Also, theta responses predicted changes in expectations only in the negative context, but unrelated to interindividual differences in prior expectations. The theta results can thus not fully explain the behavioral effects (Figure 5D).

DISCUSSION

People have different expectations about being rejected or accepted by others and thus likely respond differently to social situations that meet or do not meet these expectations. How the social situation and interindividual differences in rejection sensitivity interact in shaping the response to social rejection has not been researched in detail before. Therefore, in a between-subject design, we presented 92 participants with either mostly rejection or mostly acceptance from peers and measured changes in their rejection expectations and neural response to feedback with EEG. Interindividual differences in trial-by-trial expectation changes emerged only in the context with high rejection rates. Neural responses to feedback (P3, theta) showed an inverse pattern to acceptance and rejection in the two groups, indicating that they signal feedback probabilities and deviation from expectations rather than

Table 5. Predictive Model Fit for the Averaged Models

	Pathway		
	1	3	2
% Responses (change vs. no change) predicted correctly	63.5%	56.7%	
95th percentile of null hypothesis distribution	55%	54.2%	
Maximum value of null hypothesis distribution	56.4%	55.3%	
Average residuals of model predictions	0.455	0.488	
5th percentile of null hypothesis distribution	0.496	0.498	
Minimum value of null hypothesis distribution	0.495	0.496	
Accuracy (<i>r</i>)			0.64
95th percentile of null hypothesis distribution			0.082
Maximum value of null hypothesis distribution			0.16

Average residual = average absolute difference between actual expectation change ("0" or "1") and model predicted probability to change; *r* = Pearson's *r* for correlation between predicted and measured theta values.

being merely a social threat detection signal. Interestingly, neural feedback processing differentiated more between rejection and acceptance when the probability to be rejected was high. Moreover, theta power to feedback predicted expectation changes in the next trial only in the negative context. However, this finding was unrelated to the interindividual differences in learning from social feedback. We conclude that the behavioral and neural responses to social rejection strongly depend on the context and that people's expectations to be rejected matter mainly in threatening social situations. Our results clarify how social rejection expectation biases are maintained or discarded in different social situations. Moreover, we show that these expectation biases are not linked to a biased neural representation of rejection probability. These findings add to comprehending difficulties of people with high rejection sensitivity.

Effects of Social Context and Prior Expectations on the Behavioral Response to Social Feedback

Manipulation checks showed that our paradigm created a rejecting social situation, as participants in the negative context clearly stated that they had been mostly rejected after the paradigm. Interestingly, despite realizing the rejection, participants stayed optimistic during the experiment, as indicated by the positive expectation bias scores. In contrast, participants in the positive context already adapted their expectations during the experiment toward the high rate of acceptance, although expectations here remained slightly less positive than the actual acceptance rate. However, this underestimation of acceptance was significantly smaller than the overestimation in the negative context. This shows that, on average, people maintain an optimistic bias even in an overall threatening context, whereas in a friendly situation, they adjust their expectations toward the situation but stay a little pessimistic. These results critically extend prior work, as former studies did not manipulate the overall probability of social rejection (van der Veen et al., 2019; Kortink et al., 2018; van der Molen et al., 2017, 2018; Koban et al., 2017; Cao et al., 2015; Dekkers et al., 2015). It has to be noted though that, as participants were told to rate at least 50% of the profiles with "Yes, I want to meet this person," when preparing for the task, a pragmatic, unbiased expectation before and during the task would have been to expect acceptance in at least half of the trials. However, only 45% of participants followed this reasoning in their expectations before the task.

Importantly, the higher people's rejection sensitivity was, the lower was their optimistic bias in the negative context, which was mediated by prior expectations. Rejection sensitivity thus drives expectations before encountering a social situation, and if these negative expectations are met, they are maintained throughout the situation. Contrary to prior findings (Olsson, Carmona, Downey, Bolger, & Ochsner, 2013), even people with high rejection

expectations were able to change their negative expectations when they were clearly disconfirmed in our study. This is reflected in their high acceptance expectations in the positive context. This advances the understanding of rejection sensitivity, as to our knowledge no study so far has examined changes in rejection expectations in people with high rejection sensitivity after experiencing acceptance as compared with rejection.

Effects of Social Feedback on P3

The P3 results confirmed our hypothesis of higher P3 amplitudes for objectively less probable feedback, which is in line with former studies showing a higher P3 for unexpected or infrequent (also nonsocial) stimuli (Dekkers et al., 2015; Gutz et al., 2015; Duncan-Johnson & Donchin, 1977). In contrast, the smaller difference in subjective probability between rejection and acceptance in the negative context was not reflected in a weaker differentiation between acceptance and rejection in the P3. On the contrary, the P3 difference was much stronger in the negative than in the positive context, with the effect size of the P3 amplitude being nearly four times higher compared with the positive context. Importantly, if the differences found in EEG parameters were solely attributable to stimulus frequency, they should be exactly reversed in the two contexts. The greater differentiation in the negative context suggests that valence is processed more in a threatening social context, where it is more relevant to accurately monitor social feedback (Bennett, Sasmita, Maloney, Murawski, & Bode, 2019; Syrjämäki & Hietanen, 2019; Sacco, Wirth, Hugenberg, Chen, & Williams, 2011; Bernstein, Young, Brown, Sacco, & Claypool, 2008; Pickett et al., 2004).

Some former studies that analyzed averaged EEG responses to social feedback found no modulation of P3 in healthy samples with a broad range of rejection sensitivity (Kortink et al., 2018; van der Molen et al., 2018; Leitner et al., 2014). In contrast to these former studies, our study directly linked single-trial behavioral and EEG responses to social feedback in a large sample. Thereby, it revealed behavioral differences between people with high and low initial social expectations that were not detected in former studies. At the same time, our results confirmed that these differences are decoupled from P3 responses to social feedback. This is surprising, given the sensitivity of this neural marker to subjective expectations in other contexts (Polich, 2007; Duncan-Johnson & Donchin, 1977) and suggests that neural processes other than those linked to the representation of feedback probability might be changed in people with strong expectation biases. These biases might occur when expressing expectations, rather than in the expectations themselves (Will et al., 2020). RT data from one social feedback study suggest that people with high fear of negative evaluation take longer to state their expectations of being rejected or accepted, hinting to differences in decision-making processes when explicitly stating expectations (van der Molen et al., 2014).

Alternatively, feedback probabilities might first be represented accurately, as indexed by the P3, but later reappraisal processes might change subjective feedback probabilities toward the initial expectations, which then determine the expectations in the next trial. Similar processes have been observed in people with depression, namely, the negatively biased reappraisal of performance feedback that disconfirmed their negative self-view (Kube & Rozenkrantz, 2021; Kube et al., 2019). Future studies need to examine these possibilities, for example, by manipulating or explicitly assessing decision-making or reappraisal processes and their neural markers like the late positive potential (Hajcak et al., 2010).

Effects of Social Feedback on Theta and Alpha Activity

Another prominent neural response to social feedback was a theta power increase around 100–350 msec over frontal areas. This replicates results from former studies using social feedback paradigms (Harrewijn et al., 2018; Kortink et al., 2018; van der Molen et al., 2017, 2018; van Noordt, White, Wu, Mayes, & Crowley, 2015; Leitner et al., 2014). Extending these former results, we show how the social context modulates theta responses to rejection and acceptance. In the positive context, rejection elicited stronger theta responses than acceptance, which is similar to former studies using equal probabilities for rejection and acceptance (Harrewijn et al., 2018; Kortink et al., 2018; van der Molen et al., 2017, 2018; Leitner et al., 2014). However, this pattern was reversed in the negative context, which implies that social rejection does not always evoke stronger theta responses than acceptance, as other social feedback studies have concluded, but that theta responses depend on the social context (van der Molen et al., 2017; van Noordt et al., 2015). Specifically, theta responses in our study were stronger for the more improbable feedback in each context. Moreover, theta responses seem to be driven by the subjective experience of the probability as we found correlations with interindividual differences in subjective probability of acceptance: Participants with higher prior expectations of acceptance showed weaker theta responses to that feedback. We can thus conclude that theta to social feedback does not primarily signal social threat but subjective expectation violations. Therefore, it is unlikely that it directly relates to the dACC activity found in response to social pain (van der Molen et al., 2017; Eisenberger, 2015a).

Apart from the hypothesized theta effects, we found a striking valence effect on alpha power in the two context groups. Alpha increased after rejection feedback in the negative context and after acceptance feedback in the positive context. The first finding might reflect faster disengagement of cortical areas from the processing of rejection (Palva & Palva, 2011; Klimesch, Sauseng, & Hanslmayr, 2007). Occipital alpha has also been linked to visual attention and more generally to vigilance effects

(van Driel, Ridderinkhof, & Cohen, 2012; Macdonald, Mathan, & Yeung, 2011; Sadaghiani et al., 2010). Stronger alpha responses to rejection in the negative context and vice versa might thus be due to less attention directed to the more probable feedback. Future studies should examine the specificity of these effects for social feedback processing.

Changes in Expectations and Single-trial Theta Response to Social Feedback

We had hypothesized that rare rejections in the positive context would be largely ignored in learning, whereas rare acceptance in the negative context would be weighed more to maintain positive expectations. Indeed, we found that the overall probability of rejection in a social situation modulated how people changed their expectations from trial to trial. Specifically, participants used negative as well as positive PEs to change their expectations in the negative context, whereas only positive PEs were used for learning in the positive context. As hypothesized, how people changed their expectations from trial to trial was also influenced by prior expectations. Interestingly, these interindividual differences were only relevant when probability of rejection was high, analogue to the results from the moderated mediation model on the averaged expectations discussed above. Only in the negative context, the more rejection people expected from the beginning, the more often they changed their expectations after negative PEs.

This extends previous work showing that people with social anxiety disorder show greater learning rates for negative social feedback when changing their self-evaluation, but their learning rates for positive feedback do not differ from those of healthy controls (Koban et al., 2017). Maintaining an optimistic bias even when being continuously rejected might be a coping mechanism that people with high rejection sensitivity lack (Tanaka & Ikegami, 2015; Maner, DeWall, Baumeister, & Schaller, 2007). Our results show that this is achieved by disregarding negative PEs in learning, even if they are very frequent. This implies that to change their negative expectations, people with high rejection sensitivity might have to learn to ignore social rejection signals, rather than to attend more to social acceptance signals.

To examine what predicted these differential changes in expectations on the neural level, we analyzed the theta response to PEs. Stronger theta responses were found to predict greater expectation updates in former studies and were thus considered a neural substrate of using PEs for learning (Luft, 2014; Cavanagh, Frank, Klein, & Allen, 2010). Although the second linear mixed model suggested that PEs elicited theta responses differentially in the two contexts, control analyses showed that the interaction effects between PE and context on theta could be better explained by a negative effect of the objective probability of the feedback and an additional positive effect of negative PEs, which were unmodulated by context. This

suggests that the theta response encodes unsigned PEs (“surprise” or stimulus improbability) as well as specifically negative PEs, which has been previously suggested for social feedback (Kortink et al., 2018; van der Molen et al., 2017). Former studies on nonsocial feedback have either shown a similar effect (Cavanagh, Figueroa, Cohen, & Frank, 2012) or found that midfrontal theta responds only to negative PEs (Cohen, Elger, & Ranganath, 2007). The reasons for these diverging results are not fully understood yet. One reason might be a different operationalization of PEs. Often, PEs are inferred from objective probabilities or the selection history of the participant. In contrast, we defined PE as deviation from participants’ explicit subjective expectations in a specific trial. We can therefore extend former results on the link between theta and single-trial PEs by showing that, in addition to the objective probability of the feedback, theta encodes negative deviations from subjective expectations (Paul & Pourtois, 2017; Janssen, Poljac, & Bekkering, 2016; Mas-Herrero & Marco-Pallarés, 2014; Cavanagh et al., 2010).

Finally, we showed that theta responses indeed predicted changes in expectations, but only in the negative context. This implies that theta responses reflect the relevance of PEs for learning also for social feedback, as has been shown for nonsocial feedback before (Luft, 2014; Cavanagh et al., 2010; Marco-Pallarés et al., 2008). However, contrary to our hypothesis, theta responses and their relevance for expectation changes were not modulated by interindividual differences in prior expectations. Instead, people with low rejection expectations show the same theta response to PEs as people with strong rejection expectations, but it does not lead to expectation changes (see Unger, Heintz, & Kray, 2012, for similar results on punishment sensitivity, error negativity, and learning accuracy). Our findings critically advance our understanding of the role of midfrontal theta in learning from social feedback. Specifically, they challenge the view of theta as a social threat detection marker. In contrast to the sensitivity for the subjective probability of acceptance found in the positive context, theta shows no sensitivity for interindividual differences in responses to rejection, thereby confirming other studies on social feedback and theta (Kortink et al., 2018). Similar to the P3, theta responses might reflect early, automatic feedback processing, which contributes only partly to the putatively more complex development of expectation biases.

Limitations

Our paradigm did not include a nonsocial control condition, as the task was embedded in a larger study and an additional condition was not feasible. Therefore, we cannot claim specificity of the above detailed differences between people with high and low rejection sensitivity to responses to social feedback. However, our behavioral results align with other studies that did include nonsocial

control conditions and showed the specificity of interindividual differences for social stimuli (Will et al., 2020; Koban et al., 2017). In addition, one nonsocial reward learning study with a roughly comparable context manipulation showed partly different ERP results, although it focused on the feedback-related negativity (Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003). In conclusion, although we show robust effects for social feedback, this specificity has to be confirmed in future studies.

Furthermore, recently, methodological concerns have been raised regarding the causal interpretation of traditional mediation analyses (Lange, Hansen, Sørensen, & Galatius, 2017; Fritz, Kenny, & MacKinnon, 2016). We therefore want to stress that we do not claim that the mediation of the effect of rejection sensitivity over prior expectations on the expectation bias is necessarily a (or even the only) causal path. Self-reported general rejection sensitivity and social expectations in a specific situation are rather both manifestations of the underlying personality than cause and effect.

Moreover, although we enhanced the personal relevance of the feedback by informing participants that they would receive contact details of their “matches,” it remains unknown how comparable abstract social feedback cues are to the complex social signals people receive in real life. Especially, natural social signals are often more ambiguous, which probably leads to greater differences in their interpretation between people with different rejection expectations (Zimmer-Gembeck & Nesdale, 2013).

Conclusion

Our findings show how social context and personality influence responses to social rejection. Trial-by-trial dynamics reveal how people with low rejection expectations tend to ignore experienced rejection to uphold their optimistic bias. Interestingly, P3 and theta responses to social feedback do not relate to these expectation biases. Instead, the sensitivity of the P3 to valence is enhanced in a context with high rejection probability, indicating a greater need to monitor feedback in a threatening social situation (Pickett et al., 2004). Similar to nonsocial learning, midfrontal theta plays a role in learning from social PEs (Luft, 2014; Cavanagh et al., 2010). Our study highlights how the consideration of social context, personality, and development of neural and behavioral responses to feedback over time can help to understand and potentially counteract detrimental and persistent rejection expectations.

Acknowledgments

We thank Leah Angel and Moritz Borchering for their help with data acquisition.

Reprint requests should be sent to Ulrike M. Krämer, Department of Neurology, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany, or via e-mail: ulrike.kraemer@uni-luebeck.de.

Author Contributions

Pauline Petereit: Conceptualization; Formal analysis; Investigation; Writing—Original draft. Sarah Jessen: Writing—Review & editing. Tatiana Goregliad Fjaellingsdal: Writing—Review & editing. Ulrike M. Krämer: Conceptualization; Funding acquisition; Supervision; Writing—Review & editing.

Funding Information

Ulrike M. Krämer, Deutsche Forschungsgemeinschaft (<https://dx.doi.org/10.13039/501100001659>), grant number: KR3691/8-1.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience (JoCN)* during this period were $M(\text{an})/M = .407$, $W(\text{oman})/M = .32$, $M/W = .115$, and $W/W = .159$, the comparable proportions for the articles that these authorship teams cited were $M/M = .549$, $W/M = .257$, $M/W = .109$, and $W/W = .085$ (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance. The authors of this article report its proportions of citations by gender category to be as follows: $M/M = 0.5$; $W/M = .265$; $M/W = .059$; $W/W = .176$.

REFERENCES

- Barton, K. (2020). *Multi-model inference*. R package version 1.43.17. <https://CRAN.R-project.org/package=MumIn>
- Bates, D., Maechler, M., Bolker, B., Walker, S., Bojensen Christensen, R. H., Singmann, H., et al. (2020). *Linear mixed-effects models using Eigen and S4* (1.1-25) [R]. <https://github.com/lme4/lme4/>
- Bennett, D., Murawski, C., & Bode, S. (2015). Single-trial event-related potential correlates of belief updating. *eNeuro*, 2, ENEURO.0076-15.2015. <https://doi.org/10.1523/ENEURO.0076-15.2015>, PubMed: 26473170
- Bennett, D., Sasmita, K., Maloney, R. T., Murawski, C., & Bode, S. (2019). Monetary feedback modulates performance and electrophysiological indices of belief updating in reward learning. *Psychophysiology*, 56, e13431. <https://doi.org/10.1111/psyp.13431>, PubMed: 31274199
- Bernstein, M. J., Young, S. G., Brown, C. M., Sacco, D. F., & Claypool, H. M. (2008). Adaptive responses to social exclusion: Social rejection improves detection of real and fake smiles. *Psychological Science*, 19, 981–983. <https://doi.org/10.1111/j.1467-9280.2008.02187.x>, PubMed: 19000206
- Beyer, F., Buades-Rotger, M., Claes, M., & Krämer, U. M. (2017). Hit or run: Exploring aggressive and avoidant reactions to interpersonal provocation using a novel fight-or-escape paradigm (FOE). *Frontiers in Behavioral Neuroscience*, 11, 190. <https://doi.org/10.3389/fnbeh.2017.00190>, PubMed: 29089875
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33, 261–304. <https://doi.org/10.1177/0049124104268644>
- Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, 63, 452–459. <https://doi.org/10.1037/0022-3514.63.3.452>
- Cao, J., Gu, R., Bi, X., Zhu, X., & Wu, H. (2015). Unexpected acceptance? Patients with social anxiety disorder manifest their social expectancy in ERPs during social feedback processing. *Frontiers in Psychology*, 6, 1745. <https://doi.org/10.3389/fpsyg.2015.01745>, PubMed: 26635659
- 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. <https://doi.org/10.1093/cercor/bhr332>, PubMed: 22120491
- 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. <https://doi.org/10.1016/j.neuroimage.2009.11.080>, PubMed: 19969093
- Cohen, M. X. (2014). *Analyzing neural time series data: Theory and practice* (1st ed.). Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9609.001.0001>
- Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *NeuroImage*, 35, 968–978. <https://doi.org/10.1016/j.neuroimage.2006.11.056>, PubMed: 17257860
- De Pascalis, V., Varriale, V., & Rotonda, M. (2012). EEG oscillatory activity associated to monetary gain and loss signals in a learning task: Effects of attentional impulsivity and learning ability. *International Journal of Psychophysiology*, 85, 68–78. <https://doi.org/10.1016/j.ijpsycho.2011.06.005>, PubMed: 21704660
- Dekkers, L. M. S., van der Molen, M. J. W., Gunther Moor, B., van der Veen, F. M., & van der Molen, M. W. (2015). Cardiac and electro-cortical concomitants of social feedback processing in women. *Social Cognitive and Affective Neuroscience*, 10, 1506–1514. <https://doi.org/10.1093/scan/nsv039>, PubMed: 25870439
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134, 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>, PubMed: 15102499
- Denissen, J. J. A., Penke, L., Schmitt, D. P., & van Aken, M. A. G. (2008). Self-esteem reactions to social interactions: Evidence for sociometer mechanisms across days, people, and nations. *Journal of Personality and Social Psychology*, 95, 181–196. <https://doi.org/10.1037/0022-3514.95.1.181>, PubMed: 18605859
- Downey, G., & Feldman, S. I. (1996). Implications of rejection sensitivity for intimate relationships. *Journal of Personality and Social Psychology*, 70, 1327–1343. <https://doi.org/10.1037/0022-3514.70.6.1327>, PubMed: 8667172
- Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology*, 14, 456–467. <https://doi.org/10.1111/j.1469-8986.1977.tb01312.x>, PubMed: 905483
- Eisenberger, N. I. (2015a). Meta-analytic evidence for the role of the anterior cingulate cortex in social pain. *Social Cognitive and Affective Neuroscience*, 10, 1–2. <https://doi.org/10.1093/scan/nsu120>, PubMed: 25210052
- Eisenberger, N. I. (2015b). Social pain and the brain: Controversies, questions, and where to go from here. *Annual Review of Psychology*, 66, 601–629. <https://doi.org/10.1146/annurev-psych-010213-115146>, PubMed: 25251482

- Feldman, S., & Downey, G. (1994). Rejection sensitivity as a mediator of the impact of childhood exposure to family violence on adult attachment behavior. *Development and Psychopathology*, *6*, 231–247. <https://doi.org/10.1017/S0954579400005976>, PubMed: 23438329
- Fritz, M. S., Kenny, D. A., & MacKinnon, D. P. (2016). The combined effects of measurement error and omitting confounders in the single-mediator model. *Multivariate Behavioral Research*, *51*, 681–697. <https://doi.org/10.1080/00273171.2016.1224154>, PubMed: 27739903
- Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*, 2279–2282. <https://doi.org/10.1126/science.1066893>, PubMed: 11910116
- Glazer, J. E., Kelley, N. J., Pornpattananangkul, N., Mittal, V. A., & Nusslock, R. (2018). Beyond the FRN: Broadening the time-course of EEG and ERP components implicated in reward processing. *International Journal of Psychophysiology*, *132*, 184–202. <https://doi.org/10.1016/j.ijpsycho.2018.02.002>, PubMed: 29454641
- Gruenenfelder-Steiger, A. E., Harris, M. A., & Fend, H. A. (2016). Subjective and objective peer approval evaluations and self-esteem development: A test of reciprocal, prospective, and long-term effects. *Developmental Psychology*, *52*, 1563–1577. <https://doi.org/10.1037/dev0000147>, PubMed: 27690495
- Gutz, L., Renneberg, B., Roepke, S., & Niedeggen, M. (2015). Neural processing of social participation in borderline personality disorder and social anxiety disorder. *Journal of Abnormal Psychology*, *124*, 421–431. <https://doi.org/10.1037/a0038614>, PubMed: 25603358
- Gutz, L., Roepke, S., & Renneberg, B. (2016). Cognitive and affective processing of social exclusion in borderline personality disorder and social anxiety disorder. *Behaviour Research and Therapy*, *87*, 70–75. <https://doi.org/10.1016/j.brat.2016.08.020>, PubMed: 27616717
- Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: An integrative review. *Developmental Neuropsychology*, *35*, 129–155. <https://doi.org/10.1080/87565640903526504>, PubMed: 20390599
- Harrewijn, A., van der Molen, M. J. W., van Vliet, I. M., Tissier, R. L. M., & Westenberg, P. M. (2018). Behavioral and EEG responses to social evaluation: A two-generation family study on social anxiety. *NeuroImage: Clinical*, *17*, 549–562. <https://doi.org/10.1016/j.nicl.2017.11.010>, PubMed: 29527481
- Hepper, E. G., Hart, C. M., Gregg, A. P., & Sedikides, C. (2011). Motivated expectations of positive feedback in social interactions. *Journal of Social Psychology*, *151*, 455–477. <https://doi.org/10.1080/00224545.2010.503722>, PubMed: 21755655
- Herzberg, P. Y. (2003). Faktorstruktur, gütekriterien und konstruktvalidität der deutschen übersetzung des aggressionsfragebogens von buss und perry. [Factor structure, quality criteria and construct validity of the German translation of the aggression questionnaire by Buss and Perry]. *Journal of Differential and Diagnostic Psychology/ Zeitschrift Für Differentielle Und Diagnostische Psychologie*, *24*, 311–323. <https://doi.org/10.1024/0170-1789.24.4.311>
- Holroyd, C. B., Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport*, *14*, 2481–2484. <https://doi.org/10.1097/00001756-200312190-00037>, PubMed: 14663214
- Janssen, D. J. C., Poljac, E., & Bekkering, H. (2016). Binary sensitivity of theta activity for gain and loss when monitoring parametric prediction errors. *Social Cognitive and Affective Neuroscience*, *11*, 1280–1289. <https://doi.org/10.1093/scan/nsw033>, PubMed: 26969862
- Jones, R. M., Somerville, L. H., Li, J., Ruberry, E. J., Libby, V., Glover, G., et al. (2011). Behavioral and neural properties of social reinforcement learning. *Journal of Neuroscience*, *31*, 13039–13045. <https://doi.org/10.1523/JNEUROSCI.2972-11.2011>, PubMed: 21917787
- Kleiner, M., Brainard, D., & Pelli, D. (2007). *What's new in Psychtoolbox-3?* https://pure.mpg.de/rest/items/item_1790332/component/file_3136265/content
- Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Research Reviews*, *53*, 63–88. <https://doi.org/10.1016/j.brainresrev.2006.06.003>, PubMed: 16887192
- Koban, L., Schneider, R., Ashar, Y. K., Andrews-Hanna, J. R., Landy, L., Moscovitch, D. A., et al. (2017). Social anxiety is characterized by biased learning about performance and the self. *Emotion*, *17*, 1144–1155. <https://doi.org/10.1037/emo0000296>, PubMed: 28358557
- Korn, C. W., Prehn, K., Park, S. Q., Walter, H., & Heekeren, H. R. (2012). Positively biased processing of self-relevant social feedback. *Journal of Neuroscience*, *32*, 16832–16844. <https://doi.org/10.1523/JNEUROSCI.3016-12.2012>, PubMed: 23175836
- Kortink, E. D., Weeda, W. D., Crowley, M. J., Gunther Moor, B., & van der Molen, M. J. W. (2018). Community structure analysis of rejection sensitive personality profiles: A common neural response to social evaluative threat? *Cognitive, Affective, & Behavioral Neuroscience*, *18*, 581–595. <https://doi.org/10.3758/s13415-018-0589-1>, PubMed: 29651690
- Kube, T., Glombiewski, J. A., Gall, J., Touissant, L., Gärtner, T., & Rief, W. (2019). How to modify persisting negative expectations in major depression? An experimental study comparing three strategies to inhibit cognitive immunization against novel positive experiences. *Journal of Affective Disorders*, *250*, 231–240. <https://doi.org/10.1016/j.jad.2019.03.027>, PubMed: 30870773
- Kube, T., & Rozenkrantz, L. (2021). When beliefs face reality: An integrative review of belief updating in mental health and illness. *Perspectives on Psychological Science*, *16*, 247–274. <https://doi.org/10.1177/1745691620931496>, PubMed: 32818386
- Kvålseth, T. O. (1985). Cautionary note about R^2 . *American Statistician*, *39*, 279–285. <https://doi.org/10.1080/00031305.1985.10479448>
- Lange, T., Hansen, K. W., Sørensen, R., & Galatius, S. (2017). Applied mediation analyses: A review and tutorial. *Epidemiology and Health*, *39*, e2017035. <https://doi.org/10.4178/epih.e2017035>, PubMed: 29121709
- Leiner, D. J. (2019). *SoSci Survey* (3.2.24) [Computer software]. <https://www.sosicurvey.de>
- Leitner, J. B., Hehman, E., Jones, J. M., & Forbes, C. E. (2014). Self-enhancement influences medial frontal cortex alpha power to social rejection feedback. *Journal of Cognitive Neuroscience*, *26*, 2330–2341. https://doi.org/10.1162/jocn_a_00645, PubMed: 24738770
- Leng, Y., Qian, X., & Zhu, Y. (2018). Modulation of brain response to peer rejection by rejection sensitivity: An exploratory study. *Neuropsychologia*, *117*, 389–397. <https://doi.org/10.1016/j.neuropsychologia.2018.07.003>, PubMed: 29981785
- Loeb, E. L., Hessel, E. T., & Allen, J. P. (2016). The self-fulfilling prophecy of adolescent social expectations. *International Journal of Behavioral Development*, *40*, 555–564. <https://doi.org/10.1177/0165025415618274>, PubMed: 28082755
- Loeb, E. L., Tan, J. S., Hessel, E. T., & Allen, J. P. (2018). Getting what you expect: Negative social expectations in early adolescence predict hostile romantic partnerships and friendships into adulthood. *Journal of Early Adolescence*,

- 38, 475–496. <https://doi.org/10.1177/0272431616675971>, PubMed: 29628605
- London, B., Downey, G., Bonica, C., & Paltin, I. (2007). Social causes and consequences of rejection sensitivity. *Journal of Research on Adolescence*, *17*, 481–506. <https://doi.org/10.1111/j.1532-7795.2007.00531.x>
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213. <https://doi.org/10.3389/fnhum.2014.00213>, PubMed: 24782741
- Luft, C. D. B. (2014). Learning from feedback: The neural mechanisms of feedback processing facilitating better performance. *Behavioural Brain Research*, *261*, 356–368. <https://doi.org/10.1016/j.bbr.2013.12.043>, PubMed: 24406725
- Luft, C. D. B., Nolte, G., & Bhattacharya, J. (2013). High-learners present larger mid-frontal theta power and connectivity in response to incorrect performance feedback. *Journal of Neuroscience*, *33*, 2029–2038. <https://doi.org/10.1523/JNEUROSCI.2565-12.2013>, PubMed: 23365240
- Luu, P., Tucker, D. M., & Makeig, S. (2004). Frontal midline theta and the error-related negativity: Neurophysiological mechanisms of action regulation. *Clinical Neurophysiology*, *115*, 1821–1835. <https://doi.org/10.1016/j.clinph.2004.03.031>, PubMed: 15261861
- Macdonald, J. S. P., Mathan, S., & Yeung, N. (2011). Trial-by-trial variations in subjective attentional state are reflected in ongoing prestimulus EEG alpha oscillations. *Frontiers in Psychology*, *2*, 82. <https://doi.org/10.3389/fpsyg.2011.00082>, PubMed: 21687452
- Maner, J. K., DeWall, C. N., Baumeister, R. F., & Schaller, M. (2007). Does social exclusion motivate interpersonal reconnection? Resolving the ‘porcupine problem’. *Journal of Personality and Social Psychology*, *92*, 42–55. <https://doi.org/10.1037/0022-3514.92.1.42>, PubMed: 17201541
- Marco-Pallares, J., Cucurell, D., Cunillera, T., García, R., Andrés-Pueyo, A., Münte, T. F., et al. (2008). Human oscillatory activity associated to reward processing in a gambling task. *Neuropsychologia*, *46*, 241–248. <https://doi.org/10.1016/j.neuropsychologia.2007.07.016>, PubMed: 17804025
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*, 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>, PubMed: 17517438
- 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. https://doi.org/10.1162/jocn_a_00516, PubMed: 24188368
- Meyer, A., Lerner, M. D., Reyes, A. D. L., Laird, R. D., & Hajcak, G. (2017). Considering ERP difference scores as individual difference measures: Issues with subtraction and alternative approaches. *Psychophysiology*, *54*, 114–122. <https://doi.org/10.1111/psyp.12664>, PubMed: 28000251
- Olsson, A., Carmona, S., Downey, G., Bolger, N., & Ochsner, K. N. (2013). Learning biases underlying individual differences in sensitivity to social rejection. *Emotion*, *13*, 616–621. <https://doi.org/10.1037/a0033150>, PubMed: 23914767
- Palva, S., & Palva, J. M. (2011). Functional roles of alpha-band phase synchronization in local and large-scale cortical networks. *Frontiers in Psychology*, *2*, 204. <https://doi.org/10.3389/fpsyg.2011.00204>, PubMed: 21922012
- Paul, K., & Pourtois, G. (2017). Mood congruent tuning of reward expectation in positive mood: Evidence from FRN and theta modulations. *Social Cognitive and Affective Neuroscience*, *12*, 765–774. <https://doi.org/10.1093/scan/nsw010>, PubMed: 28199707
- Pickett, C. L., Gardner, W. L., & Knowles, M. (2004). Getting a cue: The need to belong and enhanced sensitivity to social cues. *Personality and Social Psychology Bulletin*, *30*, 1095–1107. <https://doi.org/10.1177/0146167203262085>, PubMed: 15359014
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*, 2128–2148. <https://doi.org/10.1016/j.clinph.2007.04.019>, PubMed: 17573239
- Rana, S., Midi, H., & Sarkar, S. K. (2010). Validation and performance analysis of binary logistic regression model. In N. Mastorakis, V. Mladenov, A. Zaharim, & C. A. Bulucea (Eds.), *Proceedings of the WSEAS International Conference on Environmental, Medicine and Health Sciences* (pp. 23–25). WSEAS Press. <https://www.wseas.us/e-library/conferences/2010/Penang/EMEH/EMEH-09.pdf>.
- Reichenberger, J., Schwarz, M., König, D., Wilhelm, F. H., Voderholzer, U., Hillert, A., et al. (2016). Angst vor negativer sozialer bewertung: Übersetzung und validierung der furcht vor negativer evaluation–Kurzskala (FNE-K). *Diagnostica*, *62*, 169–181. <https://doi.org/10.1026/0012-1924/a000148>
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning. II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.
- Sacco, D. F., Wirth, J. H., Hugenberg, K., Chen, Z., & Williams, K. D. (2011). The world in black and white: Ostracism enhances the categorical perception of social information. *Journal of Experimental Social Psychology*, *47*, 836–842. <https://doi.org/10.1016/j.jesp.2011.03.001>
- Sadaghiani, S., Scheeringa, R., Lehongre, K., Morillon, B., Giraud, A.-L., & Kleinschmidt, A. (2010). Intrinsic connectivity networks, alpha oscillations, and tonic alertness: A simultaneous electroencephalography/functional magnetic resonance imaging study. *Journal of Neuroscience*, *30*, 10243–10250. <https://doi.org/10.1523/JNEUROSCI.1004-10.2010>, PubMed: 20668207
- Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, *14*, 1475–1479. <https://doi.org/10.1038/nn.2949>, PubMed: 21983684
- Spicer, S. G., Mitchell, C. J., Wills, A. J., & Jones, P. M. (2020). Theory protection in associative learning: Humans maintain certain beliefs in a manner that violates prediction error. *Journal of Experimental Psychology: Animal Learning and Cognition*, *46*, 151–161. <https://doi.org/10.1037/xan0000225>, PubMed: 31556642
- Staebler, K., Helbing, E., Rosenbach, C., & Renneberg, B. (2011). Rejection sensitivity and borderline personality disorder. *Clinical Psychology & Psychotherapy*, *18*, 275–283. <https://doi.org/10.1002/cpp.705>, PubMed: 21110407
- Sun, S., & Yu, R. (2014). The feedback related negativity encodes both social rejection and explicit social expectancy violation. *Frontiers in Human Neuroscience*, *8*, 556. <https://doi.org/10.3389/fnhum.2014.00556>, PubMed: 25120457
- Syrjämäki, A. H., & Hietanen, J. K. (2019). The effects of social exclusion on processing of social information—A cognitive psychology perspective. *British Journal of Social Psychology*, *58*, 730–748. <https://doi.org/10.1111/bjso.12299>, PubMed: 30480823
- Tanaka, H., & Ikegami, T. (2015). Fear of negative evaluation moderates effects of social exclusion on selective attention to social signs. *Cognition and Emotion*, *29*, 1306–1313. <https://doi.org/10.1080/02699931.2014.977848>, PubMed: 25397922
- Teolis, A. (1998). *Computational signal processing with wavelets*. Springer Science & Business Media. <https://doi.org/10.1007/978-1-4612-4142-3>

- Unger, K., Heintz, S., & Kray, J. (2012). Punishment sensitivity modulates the processing of negative feedback but not error-induced learning. *Frontiers in Human Neuroscience*, *6*, 186. <https://doi.org/10.3389/fnhum.2012.00186>, PubMed: 22754518
- 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. https://doi.org/10.1162/jocn_a_00110, PubMed: 21812570
- van der Molen, M. J. W., Dekkers, L. M. S., Westenberg, P. M., van der Veen, F. M., & van der Molen, M. W. (2017). Why don't you like me? Midfrontal theta power in response to unexpected peer rejection feedback. *Neuroimage*, *146*, 474–483. <https://doi.org/10.1016/j.neuroimage.2016.08.045>, PubMed: 27566260
- van der Molen, M. J. W., Harrewijn, A., & Westenberg, P. M. (2018). Will they like me? Neural and behavioral responses to social-evaluative peer feedback in socially and non-socially anxious females. *Biological Psychology*, *135*, 18–28. <https://doi.org/10.1016/j.biopsycho.2018.02.016>, PubMed: 29524466
- van der Molen, M. J. W., Poppelaars, E. S., Van Hartingsveldt, C. T. A., Harrewijn, A., Gunther Moor, B., & Westenberg, P. M. (2014). Fear of negative evaluation modulates electrocortical and behavioral responses when anticipating social evaluative feedback. *Frontiers in Human Neuroscience*, *7*, 936. <https://doi.org/10.3389/fnhum.2013.00936>, PubMed: 24478667
- van der Veen, F. M., Burdzina, A., & Langeslag, S. J. E. (2019). Don't you want me, baby? Cardiac and electrocortical concomitants of romantic interest and rejection. *Biological Psychology*, *146*, 107707. <https://doi.org/10.1016/j.biopsycho.2019.05.007>, PubMed: 31136776
- van Driel, J., Ridderinkhof, K. R., & Cohen, M. X. (2012). Not all errors are alike: Theta and alpha EEG dynamics relate to differences in error-processing dynamics. *Journal of Neuroscience*, *32*, 16795–16806. <https://doi.org/10.1523/JNEUROSCI.0802-12.2012>, PubMed: 23175833
- van Noordt, S. J. R., White, L. O., Wu, J., Mayes, L. C., & Crowley, M. J. (2015). Social exclusion modulates event-related frontal theta and tracks ostracism distress in children. *Neuroimage*, *118*, 248–255. <https://doi.org/10.1016/j.neuroimage.2015.05.085>, PubMed: 26048623
- Weiss, A. R., Gillies, M. J., Philiastides, M. G., Apps, M. A., Whittington, M. A., FitzGerald, J. J., et al. (2018). Dorsal anterior cingulate cortices differentially lateralize prediction errors and outcome valence in a decision-making task. *Frontiers in Human Neuroscience*, *12*, 203. <https://doi.org/10.3389/fnhum.2018.00203>, PubMed: 29872384
- Will, G.-J., Moutoussis, M., Womack, P. M., Bullmore, E. T., Goodyer, I. M., Fonagy, P., et al. (2020). Neurocomputational mechanisms underpinning aberrant social learning in young adults with low self-esteem. *Translational Psychiatry*, *10*, 96. <https://doi.org/10.1038/s41398-020-0702-4>, PubMed: 32184384
- Will, G.-J., Rutledge, R. B., Moutoussis, M., & Dolan, R. J. (2017). Neural and computational processes underlying dynamic changes in self-esteem. *eLife*, *6*, e28098. <https://doi.org/10.7554/eLife.28098>, PubMed: 29061228
- Williams, K. D. (2007). Ostracism. *Annual Review of Psychology*, *58*, 425–452. <https://doi.org/10.1146/annurev.psych.58.110405.085641>, PubMed: 16968209
- Zimmer-Gembeck, M. J., & Nesdale, D. (2013). Anxious and angry rejection sensitivity, social withdrawal, and attribution in high and low ambiguous situations: Rejection sensitivity and reactions. *Journal of Personality*, *81*, 29–38. <https://doi.org/10.1111/j.1467-6494.2012.00792.x>, PubMed: 22583003
- Zioga, I., Hassan, R., & Luft, C. D. B. (2019). Success, but not failure feedback guides learning during neurofeedback: An ERP study. *Neuroimage*, *200*, 26–37. <https://doi.org/10.1016/j.neuroimage.2019.06.002>, PubMed: 31200067
- Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. New York: Springer. <https://doi.org/10.1007/978-0-387-87458-6>