INTRODUCTION

Fast and efficient outcome evaluation plays an eminent role in the monitoring and adjustment of human behavior. In many situations, making optimal choices relies on the brain's ability to accurately predict the incentive costs and benefits of particular options and, based on that capacity, learn to adjust future decision strategies accordingly. The neurocognitive system underlying the monitoring and evaluation of outcomes has been investigated using various methods, including event-related potentials (ERP) in the human electroencephalogram (EEG). Numerous prior studies have focused on an ERP amplitude variation, which emerges at fronto-medial scalp positions about 200 to 350 ms following the onset of an external decision feedback (Miltner et al., 1997; for an overview see Sambrook &
critical that it receives very enriched and nuanced information about action-outcome values as reflected by fine-grained variations in RewP amplitudes. Contrary to that, early studies on the ∆RewP effect indicated that it corresponds to a simple dichotomous classification into good and bad outcomes (e.g., Gehring & Willoughby, 2002; Hajcak et al., 2006; Yeung & Sanfey, 2004). However, with an increase in complexity of experimental tasks later research repeatedly demonstrated that RewP amplitudes can mirror a more finely graded scaling of outcome values (e.g., Bellebaum et al., 2010; Frömer et al., 2016; Kreussel et al., 2012; Meadows et al., 2016; Osinsky et al., 2018; Osinsky et al., 2012; also see Sambrook & Goslin, 2015). For instance, hierarchical reinforcement learning has primarily been investigated with pseudo-reward tasks where an overall goal depends on the success of various subgoals, and hence, involves multiple decision steps (e.g., Diuk et al., 2013; Ribas-Fernandes et al., 2011, 2019). These tasks revealed independent reward prediction errors for subgoal- and overall goal-related actions in the aMCC, providing support for the assumption that the aMCC responds to information from multiple levels of abstraction. In this regard, an increasingly recognized question is the modulation of the RewP by the level of feedback information content as well as the subjective weighing of different information features in single- and multi-step contexts (e.g., Cockburn & Holroyd, 2018). In particular, we recently introduced an adaptation of the so-called doors task, in which each decision led to a singular outcome event with two interlaced layers of independent temporal consequences, that is, an immediate monetary consequence and a more delayed monetary consequence (Osinsky et al., 2018). Thus, the two consequences of each outcome could either converge (e.g., a positive immediate consequence and a positive delayed consequence) or diverge (e.g., a positive immediate consequence but a negative delayed consequence) with regard to their valence. Our results indicated that RewP amplitude is proportional to the additively integrated value of both consequences. Moreover, ∆RewP sensitivity to the two consequence types correlated with choice-adjustments in consequence-specific manner and were relatively stable over a 1-week interval. Hence, signals sent to the aMCC, as reflected by the RewP, appear to convey integrated information about multiple temporal consequences of a single outcome event, with individuals stably differing in the weightings of delayed and instant consequences.

In the study presented here, we aimed to replicate and extend our earlier findings by addressing the following issues. First, we were interested whether RewP amplitude can reflect even more complexly interlaced outcome information. We therefore increased the number of temporal consequence layers from two to three. That is, each outcome now has an immediate monetary consequence that is paid out directly after the experiment, an intermediate consequence paid out 2 weeks after the experiment and a delayed consequence paid out 4 months after the experiment. Based on our prior report (Osinsky et al., 2018), we expected that RewP amplitudes are particularly proportional to the additive value of the three payouts, resulting in a roughly continuous and linear scaling with increasing additive outcome value. However, we also expected that the size of the ∆RewP effect should decrease with an increase in the payout delay, reflecting the general devaluation of delayed compared to immediate rewards (cf. Cherniakowski & Holroyd, 2013; Qu et al., 2013).

Second, we investigated the link between differences in ∆RewP and trial-to-trial behavioral adjustments. The evidence for such link is somewhat mixed with some studies reporting a substantive relation (e.g., Arbel et al., 2013; Cohen & Ranganath, 2007; Hewig et al., 2011; Holroyd & Krigolson, 2007; Santesso et al., 2008; Yasuda et al., 2004) whereas others do not (e.g., Chase et al., 2011; Luft...
et al., 2013, 2014; Walsh & Anderson, 2011). In our previous study (Osinsky et al., 2018), we observed rather small but specific relations between the ∆RewP responsivities to different temporal consequences and consequence-driven adjustments in choice behavior. We therefore expected to replicate this pattern in the current study.

Third, we investigated whether individual differences in RewP sensitivities to the three consequence layers can be measured reliably in terms of classical test theory. In the face of a growing interest in how inter-individual variance in RewP is related to personality and/or psychopathology (e.g., Bakic et al., 2016; Bress & Hajcak, 2013; Cherniakowski & Holroyd, 2013; Mothes et al., 2016; Mussel et al., 2015; Proudfit, 2015; Riepl et al., 2016; Schmidt, Holroyd, et al., 2017; Schmidt, Mussel, et al., 2017; Takács et al., 2015; Tsypes et al., 2019), researchers have recently begun to examine the psychometric properties of RewP absolute amplitudes and ∆RewP effects (e.g., Bress et al., 2015; Ethridge & Weinberg, 2018; Levinson et al., 2017; Luking et al., 2017). These studies have already indicated that the RewP is a reliable biomarker of reward processing. However, since psychometric properties are task- and sample-specific it remains unclear whether the results obtained for more simple reward stimuli also generalize for more complex and multilayered rewards. We therefore analyzed internal consistency reliability of absolute RewP amplitudes across conditions and condition-wise as well as ∆RewP scores in our task.

Fourth, we were interested in whether inter-individual differences in RewP responsivity to the three temporal consequence layers are relatively stable over a time period of a few months. In personality research, rank-order consistency is one important prerequisite to consider a certain variable as trait-like (e.g., Roberts & DelVecchio, 2000). To check for such relative stability of RewP amplitudes, we retested a large subgroup of our participants for a second time 3 months after their first participation. Although this study uses more complex forms of feedback, we expect RewP reliability and stability to be about as high as with simple feedback to assume trait-like properties.

In addition, we were also interested whether oscillatory activity in the theta range (4–8 Hz) supports complex feedback processing. Fronto-medial theta (FMθ) activity has been observed in response to a range of processes that require the need for top-down control (Bernat et al., 2015; for a review see Cavanagh & Frank, 2014). Reinforcement learning theory links FMθ to control processes of the aMCC over the actor module when a successful execution of the entire task is under threat (Holroyd & Umemoto, 2016). Thus, RewP and FMθ are likely related but not functionally redundant EEG phenomena, with the former reflecting an input signal to the aMCC and the later representing an output. To further uncover potential similarities and differences between these two signatures in the processing of multicomponent outcomes, we also conducted all analyses mentioned above for FMθ power in an exploratory manner.

## 2  METHOD

### 2.1  Participants and procedure

As we were also interested in individual differences, we collected data from a relatively large sample of 102 students from the University of Osnabrueck (i.e., all participants who responded to the announcement of the study and fit the inclusion criteria). For three of these students, less than 20 trials per condition were available for averaging (see below). Thus, the final sample consisted of 99 participants (71 women, mean age = 22.36, SD = 2.94). All had normal or corrected-to-normal vision and reported to be free of any psychiatric or neurological condition. They received partial course credit as well as the money they won in the task for their participation (immediate wins: $M = 1.97, SD = 0.27$; intermediate wins: $M = 3.92, SD = 0.60$; delayed wins: $M = 10.42, SD = 1.40$). The participants returned to the laboratory 2 weeks and 4 months after the experiment to collect their monetary rewards. All participants were asked to be contacted for participating in a retest session 3 months later. The final retest sample consisted of 58 individuals (41 women, mean age = 21.52, $SD = 2.67$) who gave permission, returned to the laboratory, and had a sufficient number of trials available for averaging EEG data (i.e., 20 trials per condition; 1 participant was excluded). The date for the retest session was not pushed behind the final reimbursement of the main session to ensure that enough participants were still available for testing. The task and procedure were the same as for the main testing session. After arrival, participants received general information about the study and gave written informed consent. Afterward, EEG recordings were prepared before participants received instructions and completed the behavioral task. Following the task, individual account significance was assessed with a written question. The entire experiment took approximately 2 hr and was in accordance with the declaration of Helsinki.

### 2.2  Task and behavioral analyses

Participants completed a complex reinforcement-learning task during which they could win or lose money for three monetary accounts. The immediate account would be cashed out directly after the experiment, the intermediate account after 2 weeks and the delayed account 4 months after the experiment. The combination of possible consequences for
each account (win or loss) was independently manipulated, and hence, resulted in eight possible outcomes that were probabilistically linked to eight different doors so that participants could learn optimal choice strategies by trial-and-error. The participants were told that within each block, a certain outcome preferably hides behind a certain door and that these preferences changed after each block. They were not informed about specific probabilistic ratios. In fact, each outcome was linked to its preferred door 11 times and to every other door three times in each block. In previous studies, we observed a high level of difficulty to learn the links based on these probabilistic ratios (Osinsky et al., 2018). This ensures almost equal trial numbers for every outcome for analysis. The option-outcome links were randomized before each block.

Each trial started with a blank screen for 250 ms followed by the eight doors arranged in a 3 × 3 matrix with a blank square in the middle. Participants choose one door by pressing the corresponding key on a numerical pad. During this choice period, the current balances of all three monetary accounts were presented at the bottom of the screen. After the choice, a fixation cross appeared for 1,000 ms until the feedback was presented for 1,200 ms. The feedback stimulus consisted of a colored letter within a shape that signaled the outcome for all three consequences simultaneously. The stimuli-account mapping was counterbalanced across participants (also see statistical analyses). All participants that took part in the retest received the same stimuli-account mapping as in the main testing session. A single uppercase letter represented a mythical creature behind the door, either F for fairy [germ. Fee] that added money to one monetary account or O for orc [germ. Ork], that stole money from this account. The color of the letter was either green or red, where a green letter indicated a win to another monetary account and a red letter a loss to that account. The letter was surrounded by either a square or a triangle that represented a win and loss to the third account, respectively. Since perceived reward value decreases with increasing distance to payout time (Green & Myerson, 2004) and losses have a higher subjective value than wins (Tom et al., 2007; Tversky & Kahneman, 1992), the following outcome values were used: Immediate: −1 Cents, +2 Cents; intermediate: −2 Cents, +4 Cents; delayed: −5 Cents, +10 Cents. The task consisted of 12 blocks with 32 trials each. All stimuli were presented on a 32-inch monitor with a screen distance of 130 cm. Timings were controlled using Presentation software (Neurobehavioral Systems, Inc., Albany, CA).

To examine adjustments in choice behavior for each of the eight outcome conditions in trial N, we calculated the frequency of choosing a different door versus choosing the same door again in trial N + 1. Ratios range from 0 (always chooses the same door again after a certain outcome) to 1 (always switches the door after a certain outcome).

2.3 | Electrophysiological recording and preprocessing

Continuous EEG was recorded using a 64-channel actiCAP electrode cap according to the international 10–10 system. All electrodes were referenced to FCz and Afz served as ground. Electrode impedances were kept below 20kΩ using ECI electrode gel. The signal was amplified and recorded with a BrainAmp System and BrainVision Recorder software (Brain Products GmbH, Gilching, Germany), bandpass filtered at 0.016–250 Hz and digitized with a sampling rate of 500 Hz.

Offline preprocessing of the EEG signal was conducted using BrainVision Analyzer software (Brain Products GmbH, Gilching, Germany). For the ERP analysis, the signal was re-referenced to the mean of all electrodes and ocular artifacts were corrected using independent component analysis as implemented in BrainVision Analyzer in semi-automatic mode. The signal was bandpass-filtered with a Butterworth zero phase filter (0.1–30 Hz, roll-off: 48 dB/oct, time constant: 3.18) and segmented into 1,000 ms epochs around feedback onset (−200 to 800 ms). Segments containing voltage steps greater than 100 μV/ms or maximum-minimum differences of 150 μV or larger within a sliding time-window of 600 ms were rejected from further analyses. All remaining segments were averaged separately for each of the eight feedback conditions and baseline corrected (−200 to 0 ms). A minimum number of 20 segments per condition was required for statistical analysis (artifact free trial numbers per condition for ERP and wavelet analyses are shown in Supporting Information Table S1). Because the visual inspection of the grand average waveforms revealed differences in ∆RewP onsets between conditions, we analyzed ∆RewP onset latencies as the 50% fractional peak latency (Kiesel et al., 2008; Luck, 2014). These onset differences were not statistically significant ($F(1.27,173.72) = 1.23, p = .29, η_p^2 = .01$), nor was there a significant interaction of condition and stimuli-mapping group ($F(3.56,170.89) = .58, p = .66, η_p^2 = .01$). Therefore, the RewP was quantified as mean amplitude at FCz between 200 and 300 ms post feedback-onset. For time-frequency analysis, the EEG signal was re-segmented into 2,000 ms epochs around feedback onset (−1,000 to 1,000 ms). Segments containing voltage steps greater than 100 μV/ms or maximum-minimum differences of 150 μV or larger were rejected. Remaining segments were transformed using a family of complex Morlet wavelets from 1 to 30 Hz in 30 logarithmic steps. These modulated Gaussian sine functions are defined as: $t = Ae^{-t^2/2\omega^2}$, with parameter $c = 6.7$. Due to unit energy normalization, all frequency levels have the same total energy of 1.0. Resulting power levels for each frequency layer were then baseline-corrected using the −800 to −200 ms time window and transformed to a decibel (dB) scale. Segments were averaged for each participant and condition. FM0 was quantified as mean power from 4.09 to 7.34 Hz within 200 to 400 ms post feedback onset at FCz.
2.4 Statistical analysis

All statistical analysis were performed using the Statistical Package for the Social Sciences (SPSS for Windows, version 25; Chicago, IL, USA). For all inference statistics the alpha level was set at .05. Switching frequencies, RewP amplitude and FMθ power were analyzed, using 2 × 2 × 2 × 3 mixed design analyses of variance (ANOVAs) with immediate, intermediate and delayed consequences (positive/negative) as within-subject factors and stimuli-account mapping as a between subject factor (3 groups; see below). For resulting interactions between within-subject factors, aggregated scores across non-impacting variables and difference scores (i.e., ΔRewP) of effective variables were computed. Whenever the assumption of sphericity was violated, Greenhouse–Geisser correction was used to adjust the degrees of freedom. Two-tailed t tests for paired samples were performed to check whether the effect of individual outcomes differs depending on other outcomes.

Because the association between outcome valence and letters and colors are relatively over-learned compared to geometric shapes and might, therefore, be more easily processed, the stimulus-account mapping was considered as a between-subject factor (3 groups) as suggested by an anonymous reviewer. For group 1, the immediate consequence was indicated by the letter, the intermediate consequence by the color of the letter and the delayed consequence by the shape surrounding the letter. For group 2, the shape signaled the immediate wins and losses, the letter the intermediate consequence and the color of the letter the delayed consequence. For the third group of participants, the color and shape indicated immediate and intermediate consequences, respectively, whereas the letter signaled the money won or lost for the delayed account. The sample was not split into groups for stability and reliability analyses because the stimulus mapping was constant between the main session and the retest.

In order to extract individual responsivity of the RewP to the different consequence types, we calculated ΔRewP difference-scores by subtracting aggregated RewP amplitudes for negative consequences from those of positive consequences separately for each of the three temporal consequence types. To assess internal consistency reliability, Spearman–Brown corrected correlations (r_{sb}) were calculated between absolute RewP amplitudes and ΔRewP scores resulting from odd and even trials (i.e., split-half reliability). Moreover, we calculated Cronbach’s alpha across all eight conditions, indicating the general interrelation between condition-averaged RewP amplitudes (cf. Thigpen et al., 2017). Relative temporal stability of absolute RewP amplitudes and ΔRewP scores from T1 to T2 was assessed by Pearson correlation coefficients. To test whether the RewP has a behavioral effect, multiple regressions for switching frequencies on the ΔRewP scores were performed. Effect sizes were calculated in terms of partial eta squared for ANOVAs and in terms of Cohen’s d for t tests. The same statistical analyses were applied for FMθ.

3 RESULTS

3.1 Behavioral data

An ANOVA on switching frequencies (SF-scores) revealed significant main effects for immediate consequences (F(1,96) = 47.41, p < .001, η^2_p = .33), intermediate consequences (F(1,96) = 45.35, p < .001, η^2_p = .32) and delayed consequences (F(1,96) = 90.40, p < .001, η^2_p = .49), all reflecting greater tendencies for switched choices after negative compared to positive consequences (Figure 1). Furthermore, we found a significant interaction of immediate consequences and intermediate consequences’ (F(1,96) = 9.29, p = .003, η^2_p = .09), showing that the tendency of switching upon negative intermediate consequences increased when immediate consequences were negative as well (t(98) = 3.07, p = .003, d = .44). The same applies to the interaction of immediate consequences and delayed consequences (F(1,96) = 9.29, p = .003, η^2_p = .09) and an increased tendency of switching upon negative delayed consequences when the immediate consequences were negative as well (t(98) = 3.08, p = .003, d = .44). Altogether, the tendency to stay with a certain choice increased with approximating the optimal outcome. We found no significant main effect or interactions of stimuli-account mappings (all F(2,96) < 1.34, all p > .27).

3.2 Time-domain RewP

Mean RewP amplitudes are shown in Figure 2. Grand average ERP waveforms and topographical maps are shown in Figure 3a. The ANOVA on RewP amplitudes revealed significant main effects of immediate consequences (F(1,96) = 84.13, p < .001, η^2_p = .47), intermediate consequences (F(1,96) = 64.97, p < .001, η^2_p = .40) and delayed
consequences \( (F(1,96) = 20.05, p < .001, \eta^2_p = .17) \) with greater amplitudes for positive outcomes for each consequence type. There were significant interactions for immediate consequences and intermediate consequences \( (F(1,96) = 22.64, p < .001, \eta^2_p = .19) \) as well as immediate consequences and delayed consequences \( (F(1,96) = 6.30, p = .014, \eta^2_p = .06) \). These interactions reflected a modulation of the ∆RewP effect size for intermediate consequences and delayed consequences by the immediate factor. In particular, the ∆RewP effect was larger for intermediate \((r(98) = 4.78, p < .001, d = 0.68)\) and delayed consequences \((r(98) = 2.54, p = .013, d = .36)\) when the immediate consequence was positive compared to negative. Further, we found a main effect of stimuli-account mappings \( (F(2,96) = 5.76, p = .004, \eta^2_p = .11) \) and a significant interaction of stimuli-account mappings and immediate consequences \( (F(2,96) = 5.94, p = .004, \eta^2_p = .11) \). Post hoc tests showed greater amplitudes for positive immediate outcomes than after negative immediate outcomes within all stimuli-account mappings \( (all t(98) > 4.421, all p < .001) \). However, this effect was significantly larger when the immediate consequence was signaled by the color as compared to the letter \((t(65) = 3.74, p < .001, d = .91)\). There were no other significant effects of

**FIGURE 2** Mean RewP amplitudes per condition at FCz across groups for the main testing session and the retest. Error bars represent the standard error of the mean (SEM)

**FIGURE 3** Grand averaged ERP waves for each outcome (immediate, intermediate, delayed consequences) are shown in the upper row, whereas the lower row of graphs depicts the difference waves for positive minus negative consequences per condition and topographical maps for difference waves (200–300 ms). Graphs on the left side (A) refer to the main testing session (T1). Graphs on the right side (B) are based on the retest data. All waveforms are collapsed across stimuli-account mapping groups (group-wise results are shown in Supporting Information Figure S1)
stimuli-account mapping on in RewP amplitude for the immediate consequence.

Altogether, all three consequence types induced a significant ∆RewP effect, which, however, occurred more pronounced for the immediate compared to the intermediate and delayed consequences. Moreover, the general sensitivity of the RewP to intermediate and delayed consequences was stronger when the immediate consequences of a given outcome were positive compared to negative. The strongest RewP was evoked by the optimum outcome, which entailed positive consequences at all three payout timepoints. The more the deviation from this optimum, the smaller the RewP. Overall, a similar result pattern was observed for the retest data (see Figures 2 and 3b). A detailed description of the retest results can be found in the Supporting Information.

Considering all eight condition-wise RewP amplitudes as items, we observed a very high Cronbach’s alpha of .99. Moreover, when considering each condition separately, we also found very high split-half reliabilities ($r_{SB} = .92$ to .95). With regard to ∆RewP scores, split-half reliabilities were only in a low to moderate range ($r_{SB} = .51$ to .61) (see Table 1). For the condition-wise absolute RewP amplitudes, we observed high test-retest correlation coefficients (see Table 1). Concerning the ΔRewP scores, we observed substantive stability for the immediate and delayed consequences. The ΔRewP for intermediate consequences showed a very low and non-significant test-retest stability. Given the rather low split-half reliabilities of the ΔRewP scores, we applied double correction for attenuation ($\rho_{tt2} = \frac{r_{SB}}{\sqrt{r_{SB} \cdot r_{SB}}}$) to estimate test-retest stabilities under the assumption of perfectly reliable ΔRewP scores (Muchinsky, 1996). These estimates pointed to a strong temporal stability of ∆RewP scores for immediate consequences and moderate temporal stability for delayed consequences. Interestingly, we observed no significant interrelations between the three ΔRewP scores ($r = -.10$ to .13, all $p > .21$), indicating that the individual responsibilities to the three consequence types were independent of each other. Again, this might be caused by low reliabilities of these difference scores. However, even when we applied a double correction for attenuation the correlations between ΔRewP remained in a low range ($r = -.18$ to .23). To analyze the potential link between individual differences in RewP and consequence-driven decision-making, we conducted multiple regressions of ΔRewP scores on ΔSF-Scores for each consequence type. With adjusted $R^2 < .01$ for all of the regression models, none of these allowed statistically significant predictions (all $F(3,95) < 1.47$, all $p > .23$).

### 3.3 | FMθ

Grand average time-frequency plot for the difference of positive minus negative consequences per condition at FCz are shown in Figure 4. The ANOVA on mean FMθ power revealed a significant effect of immediate consequences ($F(1,96) = 18.59$, $p < .001$, $\eta^2_p = .16$) with stronger FMθ when the immediate consequences were negative. There was a significant interaction of intermediate consequences and delayed consequences ($F(1,96) = 9.15$, $p = .003$, $\eta^2_p = .087$) with stronger FMθ when the intermediate and delayed

---

**TABLE 1** Spearman–Brown split-half reliability and temporal stability of ΔRewP and ΔFMθ

|                | RewP | FMθ |
|----------------|------|-----|
|                | Reliability | Temporal stability | Reliability | Temporal stability |
| Immediate consequences | $r_{SB}$ | $r$ | $p$ | $r_{SB}$ | $r$ | $p$ |
| ∆RewP/ΔFMθ | .61 | .65*** | 1 | .66 | .51*** | .74 |
| Positive | .98 | .81*** | .83 | .92 | .64*** | .70 |
| Negative | .98 | .81*** | .83 | .91 | .63*** | .68 |
| Intermediate consequences | $r_{SB}$ | $r$ | $p$ | $r_{SB}$ | $r$ | $p$ |
| ∆RewP/ΔFMθ | .61 | .12 | .22 | .43 | .10 | .24 |
| Positive | .98 | .80*** | .82 | .92 | .60*** | .67 |
| Negative | .99 | .80*** | .82 | .92 | .62*** | .69 |
| Delayed consequences | $r_{SB}$ | $r$ | $p$ | $r_{SB}$ | $r$ | $p$ |
| ∆RewP/ΔFMθ | .51 | .39** | .93 | .48 | .15 | .30 |
| Positive | .98 | .80*** | .82 | .92 | .64*** | .72 |
| Negative | .98 | .82*** | .84 | .91 | .59*** | .65 |

**Note:** $r_{SB} =$ Spearman–Brown split half-reliability; $r =$ Pearson correlation coefficient; $\rho =$ disattenuated correlation coefficient.

*** $p < .001$; ** $p < .005$;
consequences were both negative ($t(98) = 3.05, p = .003, d = .43$). Furthermore, there was a significant interaction of delayed consequences and stimuli-account mappings ($F(2,96) = 5.0, p = .009, \eta^2_p = .094$) as well as immediate consequences, delayed consequences and stimuli-account mappings ($F(2,96) = 3.32, p = .040, \eta^2_p = .065$). Post hoc tests showed that only within group 3, where the delayed consequence was signaled by the letter, was FMθ power greater when the delayed consequences were negative compared to positive ($t_1(32) = 1.30, p = .202, d = .32; t_2(31) = 1.36, p = .183, d = .34; t_3(33) = 2.77, p = .009, d = .67$). Further, only within group two (immediate consequences signaled by the shape; delayed consequences signaled by the color) there was larger ΔFMθ in response to delayed consequences when the immediate consequence was negative compared to positive ($t_1(32) = .87, p = .390, d = .21; t_2(31) = 2.22, p = .034, d = .55; t_3(33) = .69, p = .493, d = .17$).

Split-half internal consistency of ΔFMθ scores (Table 1) generally fell below what can be considered as acceptable ($r_{SB} = .43$ to .66) and showed no strong interrelations between conditions ($r = -.04$ to .11, all $p > .27$). In contrast, for absolute FMθ power per condition, we observed good to excellent split-half reliabilities ($r_{SB} = .91$ to .92). For ΔFMθ scores, we observed medium test-retest correlations for immediate consequences ($r = .51$), but only negligible and insignificant coefficients for intermediate and delayed consequences ($r = .10$ to .15). However, temporal stability for the constituent power per consequence condition was moderate across all three consequence types ($r = .59$ to .64).

Multiple regression analyses of ΔFMθ on ΔSF-Scores for each consequence type did not yield any statistically significant regression models with all adjusted $R^2 < .01$ (all $F(3,95) < 1.32, all p > .27$). Accordingly, we found no behavioral effect of FMθ.

### 4 | DISCUSSION

Human behavior is fundamentally driven by promoting actions that maximize reward through feedback-based learning. However, the outcomes of one's action can be complex. This is especially the case when a single action involves multiple consequences, which differ in the time they will occur. The aim of this study was to shed new light on the relationship between such complex forms of feedback and electrophysiological indices of reward processing (cf. Osinsky et al., 2018), by increasing the feedback information content by a third temporal dimension. The results showed gradations in feedback-related activity relative to valence of the three temporal outcome dimensions. That is, for each consequence, the RewP showed larger amplitudes for wins than for losses and these differences were greatest for immediate consequences and decreased with increasing delay in payout time. These effects were considerably smaller for FMθ and...
did not relate to adjustments in trial-to-trial choice. With that, our results are in line with recent studies that used similar levels of feedback information content following single or multi-step choices (e.g., Bellebaum et al., 2010; Cockburn & Holroyd, 2018; Frömer et al., 2016; Luft et al., 2014; Osinsky et al., 2018). They also underscore current theoretical assumptions about the role of the aMCC within a hierarchical reinforcement learning system (Holroyd & Umemoto, 2016; Holroyd & Yeung, 2012), which is able to produce and learn from prediction errors that stem from the processing of complex outcome information. However, regarding the use of more complex forms of feedback, it is not yet conclusive whether the RewP reflects multiple prediction errors that project to the aMCC from subcortical critic regions or one overall valuation signal.

Recent studies using multi-step decision paradigms indicated simultaneous dissociable prediction errors at multiple levels of the hierarchy within critic module areas only as long as the reward information of different levels is presented separately (Daw et al., 2011; Diuk et al., 2013; Ribas-Fernandes et al., 2011, 2019). Functional imaging results in human and non-human primates show that the subjective value of immediate and delayed rewards is processed in distinguishable neuron populations along an anterior-posterior gradient within the medial prefrontal cortex (Roesch et al., 2006, 2012; Wang et al., 2014). In addition, it has been shown that the firing pattern of neurons in the OFC and VS represents individual reward preferences (O’Doherty et al., 2006; Tremblay & Schultz, 1999). In this light, the lack of substantive interrelations between the ∆RewP effects may indicate that the individual responsivity to immediate and delayed consequences reflect two distinct trait-like properties of the brain, which represent the subjective weighting of outcome aspects. This could be realized by anatomical differences in neuron population (e.g., size, effective connectivity) that transfer to stable differences in the signals sent to the aMCC or by functional differences like attentional orienting to different subinformation of complex feedback stimuli (Leong et al., 2017; Niv et al., 2015; Oemisch et al., 2019). Thereby, unique motivational values based on stable preferences to immediate over delayed rewards can be placed on different reward quantities. Hence, stable differences in delay discounting may give rise to differences in item-specific utility by assigning different attentional weights to different stimulus features.

While the results of this study suggest that the value of different outcome dimensions are processed by an integrated system and reflected in the RewP, this assumption must not necessarily hold true for unsigned prediction errors or salience signals. Since it has been shown that the time windows before and after the RewP are sensitive to salience (e.g., Nieuwenhuis et al., 2004; Sambrook & Goslin, 2014), the current use of three different visual stimuli (letters, colors and geometrical shapes) to indicate the three different consequence types is a potential confound in this study because they may perceptually be easier (color and letter) or harder (shape) to decode, and thus, more or less salient. Although the stimulus-account mapping was randomized, small effects of were observed in the RewP time window. However, as can be seen in Supporting Information Figure S1, the stimulus-consequence interaction is most prominent at the N1 and does not seem to reflect perceptual differences between the stimuli themselves. While the role of salience is beyond the scope of this study, this issue deserves further attention in future studies.

If the RewP reflects stable differences in outcome evaluation, one could expect a relation between RewP amplitude and habitual decision-making. However, in the present study, we did not observe any substantive covariance between RewP responsivities to immediate and delayed outcome consequences and choice behavior. This is clearly in contrast to our previous and very similar study (Osinsky et al., 2018), where we observed such a relationship. However, the task itself is with its eight response options and three outcome dimensions rather difficult to learn and might promote exploration more than exploitation. This is also demonstrated when looking at the switching frequency of the optimal outcome which is, at about 47%, rather high. It should be noted though that the literature on the RewP-behavior link is generally highly mixed (for an overview, see Holroyd & Umemoto, 2016). What are the reasons for these inconsistencies? The system might learn simple response-outcome associations in a model-free low-level fashion that would be linked to trial-to-trial adjustments in behavior. The aMCC, however, chooses and maintains timely extended goal-directed policies that are not directly linked to trial-by-trial behavioral adjustments by following model-based principles (Holroyd & Umemoto, 2016; Sambrook & Goslin, 2015). The link between the RewP and behavioral choice should, therefore, be weaker in contexts that favor goal-directed control (Walsh & Anderson, 2012). Despite the fact that delay discounting causes people to place a higher subjective value on instantaneous outcomes, our participants reported that delayed outcomes (offering the greatest objective value) were most important to them. Hence, model-free prediction errors in this task might be based on subjective value, consistent with research that found that dopaminergic regions were only sensitive to subjective value in the evaluation phase but not during choice (Liu et al., 2011) and our finding that immediate positive reward has the largest effect on the RewP. In contrast, a later goal directed system may continue to calculate prediction errors in a model-based fashion, which eventually drove behavioral choice in our task (cf. Bayer & Glimcher, 2005). There is also the possibility that the weighting of immediate and delayed rewards might shift during the course of the experiment depending on how much money has been collected for each account.

To extend our previous work (Osinsky et al., 2018), we also analyzed oscillatory correlates of multidimensional...
feedback processing. FM0 power has consistently been interpreted as a within network communication mechanism indicating increased need for cognitive control (for a review see Cavanagh & Frank, 2014). Although flexible recruitment of different brain areas seems especially crucial following multiple external demands, our results indicated that, contrary to the time-domain RewP, FM0 responses to the different delay times is much smaller and depend on the quantity of negative outcomes rather than positive ones. Still, making fine graded distinctions between multiple favorable options to learn, which one is the most rewarding enables the tradeoff between reward maximization and the minimization of resource expenditure. Consistent with this idea and previous findings, the RewP is primarily driven by reward-related variance (Foti et al., 2011; Proudfit, 2015; for a review see Sambrook & Goslin, 2015). When it comes to adjusting control, however, a fine graded dissociation may be not as crucial as it is with rewards (i.e., multiple unfavorable outcomes require similar adaptations in licensed control). Therefore, the results of this study are consistent with the idea that the time domain RewP and FM0 reflect dissociable components of reward and control processing, respectively (e.g., Holroyd & Umemoto, 2016; Luft et al., 2014; Osinsky et al., 2016; Rawls et al., 2020).

Besides the general relation between the RewP and multicomponent outcome processing, we were interested in psychometric properties of the RewP, an issue which has already attracted some attention in more recent years (e.g., Bress et al., 2015; Ethridge & Weinberg, 2018; Levinson et al., 2017; Luking et al., 2017). Our analyses pointed to excellent reliabilities of condition-wise absolute RewP amplitudes and also showed that a very high proportion of RewP variance is shared between conditions. Since common true variance is largely canceled-out when subtracting two highly correlated variables from each other, it is not surprising that reliabilities of the ∆RewP (and also ∆FMT) scores dropped to a rather unsatisfying level. This problem of difference scores is long known and some authors, therefore, questioned the use of such indices (e.g., Cronbach & Furby, 1970). However, because ERPs are no uni-dimensional measures but rather reflect a mixture of several contributing variables, the application of between-condition contrasts and difference scores is broadly accepted for isolating mechanisms of interest (e.g., Hardy, 2005; Keil et al., 2014; Luck, 2014). Accordingly, the high interrelation between condition wise RewP amplitudes in our study might be caused by individual differences in a few broad and stable neurocognitive, neurophysiological and/or neuroanatomical variables that contribute large proportions of systematic variance to each and every RewP measurement under different outcome conditions. Using absolute RewP amplitudes as indices of reward processing may, therefore, even obscure individual differences in the neurocognitive mechanism of interest. Moreover, it has been argued that the concept of reliability in terms of classical test theory is unsuitable when it comes to the measurement precision of difference scores (e.g., Fischer, 2003). Future studies on the psychometry of ERP difference scores may, therefore, apply other approaches on this issue (e.g., item response theory).

5 | CONCLUSION

In sum, our results show that the RewP is sensitive to the graded value of action outcomes, resulting from the integration of three temporally distinct consequences of a singular outcome event. From a reinforcement learning perspective (cf. Holroyd & Coles, 2002; Holroyd & Umemoto, 2016; Holroyd & Yeung, 2012), our data indicate that the aMCC may receive finely nuanced and complex teaching signals about action outcome values. Moreover, these signals may be subject to relatively stable interindividual differences, which might reflect personal reward preferences. However, because of the low spatial resolution and inverse problem of EEG, future studies on this issue may employ other techniques (e.g., combined EEG-fMRI recording) to further elucidate where exactly in the brain the integration of immediate and delayed outcome consequences into a single value takes place.

ACKNOWLEDGMENTS

Open access funding enabled and organized by ProjektDEAL.

AUTHOR CONTRIBUTION

Lena Rommerskirchen: Formal analysis; Methodology; Visualization; Writing-original draft; Writing-review & editing. Leon Lange: Formal analysis; Methodology; Visualization; Writing-original draft; Writing-review & editing. Roman Osinsky: Conceptualization; Project administration; Supervision; Writing-original draft; Writing-review & editing.

ORCID

Lena Rommerskirchen https://orcid.org/0000-0001-7715-1649

REFERENCES

Arbel, Y., Goforth, K., & Donchin, E. (2013). The good, the bad, or the useful? The examination of the relationship between the feedback related negativity (FRN) and long-term learning outcomes. Journal of Cognitive Neuroscience, 25, 1249–1260. https://doi.org/10.1162/jocn_a_00385
Bakic, J., Pourtois, G., Jepma, M., Duprat, R., De Raedt, R., & Baeken, C. (2016). Spared internal but impaired external reward prediction error signals in major depressive disorder during reinforcement learning.
learning, *Depression and Anxiety*, 34(1), 89–96. https://doi.org/10.1002/da.22576

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, 47(1), 129–141. https://doi.org/10.1016/j.neuron.2005.05.020

Bellebaum, C., Polezzi, D., & Daum, I. (2010). It is less than you expected: The feedback-related negativity reflects violations of reward magnitude expectations. *Neuropsychologia*, 48, 3334–3350. https://doi.org/10.1016/j.neuropsychologia.2010.07.023

Bernat, E. M., Nelson, L. D., & Baskin-Sommers, A. R. (2015). Time-frequency theta and delta measures index separable components of feedback processing in a gambling task. *Psychophysiology*, 52(5), 626–637. https://doi.org/10.1111/psyp.12390

Bress, J. N., & Hajcak, G. (2013). Self-report and behavioral measures of reward sensitivity predict the feedback negativity. *Psychophysiology*, 50(7), 610–616. https://doi.org/10.1111/psyp.12053

Bress, J. N., Meyer, A., & Hajcak Proudfit, G. (2015). The stability of the feedback negativity and its relationship with depression during childhood and adolescence. *Development and Psychopathology*, 27, 1285–1294. https://doi.org/10.1017/S0954579414001400

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, 18, 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Chase, H. W., Swainson, R., Durham, L., Benham, L., & Cools, R. (2011). Feedback-related negativity codes prediction error but not behavioral adjustment during probabilistic reversal learning. *Journal of Cognitive Neuroscience*, 23(4), 936–946. https://doi.org/10.1162/jocn.2010.21456

Cherniawski, A. S., & Holroyd, C. B. (2013). High temporal discounters overvalue immediate rewards rather than undervalue future rewards: An event-related brain potential study. *Cognitive Affective and Behavioral Neuroscience*, 13, 36–45. https://doi.org/10.3758/s13415-012-0122-x

Cockburn, J., & Holroyd, C. B. (2018). Feedback information and the reward positivity. *International Journal of Psychophysiology*, 132, 243–251. https://doi.org/10.1111/psyp.13508

Cohen, M. X., & Ranganath, C. (2007). Reinforcement learning signals predict future decisions. *The Journal of Neuroscience*, 27, 371–378. https://doi.org/10.1523/jneurosci.4421-06.2007

Cronbach, L. J., & Furby, L. (1970). How we should measure “change”: Or should we? *Psychological Bulletin*, 74(1), 68–80. https://doi.org/10.1037.h0029382

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans’ choices and striatal prediction errors. *Neuron*, 69(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Diuk, C., Tsai, K., Wallis, J., Botvinick, M., & Niv, Y. (2013). Hierarchical learning induces two simultaneous, but separable, prediction errors in human basal ganglia. *Journal of Neuroscience*, 33(13), 5797–5805. https://doi.org/10.1523/JNEUROSCI.5445-12.2013

Ethridge, P., & Weinberg, A. (2018). Psychometric properties of neural responses to monetary and social rewards across development. *International Journal of Psychophysiology*, 132, 311–322. https://doi.org/10.1016/j.ijpsycho.2018.01.011

Fischer, G. H. (2003). The precision of gain scores under an Item Response Theory perspective: A comparison of asymptotic and exact conditional inference about change. *Applied Psychological Measurement*, 27, 3–26. https://doi.org/10.1177/0146621602239474

Foti, D., Weinberg, A., Dien, J., & Hajcak, G. (2011). Event-related potential activity in the basal ganglia differentiates rewards from non-rewards: Temporospatial principal components analysis and source localization of the feedback negativity. *Human Brain Mapping*, 32(12), 2207–2216. https://doi.org/10.1002/hbm.21182

Frömer, R., Stürmer, B., & Sommer, W. (2016). The better, the bigger: The effect of graded positive performance feedback on the reward positivity. *Biological Psychology*, 114, 61–68. https://doi.org/10.1016/j.biopsycho.2015.12.011

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, 295(5563), 2279–2282. https://doi.org/10.1126/science.1066893

Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, 130, 769–792. https://doi.org/10.1037/0033-2909.130.5.769

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, 71(2), 148–154. https://doi.org/10.1016/j.biopsycho.2005.04.001

Hardy, T. C. (2005). Event related potentials: A methods handbook. MIT Press.

Hewig, J., Kretschmer, N., Trippe, R. H., Hecht, H., Coles, M. G. H., Holroyd, C. B., & Miltnner, W. H. R. (2011). Why humans deviate from rational choice. *Psychophysiology*, 48, 507–514. https://doi.org/10.1111/j.1469-8886.2010.01081.x

Holroyd, C. B., & Miltnner, W. H. R. (2011). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109, 679–709. https://doi.org/10.1037/0033-295X.109.4.679

Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, 44(6), 913–917. https://doi.org/10.1111/j.1469-8886.2007.00561.x

Holroyd, C. B., & Unemoto, A. (2016). The research domain criteria framework: The case for anterior cingulate cortex. *Neuroscience and Biobehavioral Reviews*, 71, 418–443. https://doi.org/10.1016/j.neubiorev.2016.09.021

Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences*, 16, 122–128. https://doi.org/10.1016/j.tics.2011.12.008

Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenmann, E. S., Luck, S. J., Luu, P., Miller, G. A., & Yee, C. M. (2014). Committee report: Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology*, 51, 1–21. https://doi.org/10.1111/psyp.12147

Kiesel, A., Miller, J., Jolicœur, P., & Jolicœur, B. (2008). Measurement of ERP latency differences: A comparison of single-participant and jackknife-based scoring methods. *Psychophysiology*, 45(2), 250–274. https://doi.org/10.1111/j.1469-8886.2007.00618.x

Kreussel, L., Hewig, J., Kretschmer, N., Hecht, H., Coles, M. G., & Miltnner, W. H. R. (2012). The influence of the magnitude, probability, and valence of potential wins and losses on the amplitude of the feedback negativity. *Psychophysiology*, 49(2), 207–219. https://doi.org/10.1111/j.1469-8886.2011.02911.x

Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V., & Niv, Y. (2017). Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron*, 93(2), 451–463. https://doi.org/10.1016/j.neuron.2016.12.040

Levinson, A. R., Speed, B. C., Infantolino, Z. P., & Hajcak, G. (2017). Reliability of the electrocortical response to gains and losses in the...
doors task. *Psychophysiology,* 54, 601–607. https://doi.org/10.1111/psych.12813

Liu, X., Hairston, J., Schrier, M., & Fan, J. (2011). Common and distinct networks underlying reward valence and processing stages: A meta-analysis of functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews,* 35(5), 1219–1236. https://doi.org/10.1016/j.neubiorev.2010.12.012

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). MIT Press.

Luft, C. D. B., Norle, G., & Bhattacharya, J. (2013). High-learners present larger mid-frontal theta power and connectivity in response to incorrect performance feedback. *The Journal of Neuroscience,* 33, 2029–2038. https://doi.org/10.1523/jneurosci.2565-12.2013

Luft, C. D. B., Takase, E., & Bhattacharya, J. (2014). Processing graded feedback: Electrophysiological correlates of learning from small and large errors. *Journal of Cognitive Neuroscience,* 26(5), 1180–1193. https://doi.org/10.1162/jocn_a_00543

Luking, K. R., Nelson, B. D., Infantolino, Z. P., Sauder, C. L., & Hajcak, G. (2017). Internal consistency of functional magnetic resonance imaging and electroencephalography measures of reward in late childhood and early adolescence. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging,* 2(3), 289–297. https://doi.org/10.1016/j.bpsc.2016.12.004

Meadows, C. C., Gable, P. A., Lohse, K. R., & Miller, M. W. (2016). The effects of reward magnitude on reward processing: An averaged and single trial event-related potential study. *Biological Psychology,* 118, 154–160. https://doi.org/10.1016/j.biopsycho.2016.06.002

Mittner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a generic neural system for error detection. *Journal of Cognitive Neuroscience,* 9, 788–798. https://doi.org/10.1162/jocn.1997.9.6.788

Mothes, H., Enge, S., & Strobel, A. (2016). The interplay between feedback-related negativity and individual differences in altruistic punishment: An EEG study. *Cognitive Affective and Behavioral Neuroscience,* 16, 276–288. https://doi.org/10.3758/s13415-015-0388-x

Muchinsky, P. M. (1996). The correction for attenuation. *Educational and Psychological Measurement,* 56(1), 63–75. https://doi.org/10.1177/001316449605601004

Mussel, P., Reiter, A., Osinsky, R., & Hewig, J. (2015). State and trait greed, its impact on risky decision making and underlying neural mechanism. *Social Neuroscience,* 10, 126–134. https://doi.org/10.1080/17470919.2014.965340

Nieuwenhuis, S., Yeung, N., Holroyd, C. B., & Cohen, J. D. (2004). Sensitivity of electrophysiological activity from medial frontal cortex to utilitarian and performance feedback. *Cerebral Cortex,* 14(7), 741–747. https://doi.org/10.1093/cercor/bhh034

Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. *Journal of Neuroscience,* 35(21), 8145–8157. https://doi.org/10.1523/JNEURScis.2978-14.2015

O’Doherty, J. P., Buchanan, T. W., Seymour, B., & Dolan, R. J. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron,* 49, 157–166. https://doi.org/10.1016/j.neuron.2005.11.014

Oermisch, M., Westendorf, S., Azimi, M., Hassan, S. A., Ardid, S., Tiesinga, P., & Womelsdorf, T. (2019). Feature-specific prediction errors and surprise across macaque fronto-striatal circuits. *Nature Communications,* 10(1), 1–15. https://doi.org/10.1038/s41467-018-08184-9

Osinsky, R., Holst, K., & Ulrich, N. (2018). When two become one: Electro cortical correlates of the integration of multiple action consequences. *International Journal of Psychophysiology,* 132, 252–261. https://doi.org/10.1016/j.ijpsycho.2017.11.014

Osinsky, R., Mussel, P., & Hewig, J. (2012). Feedback-related potentials are sensitive to sequential order of decision outcomes in a gambling task. *Psychophysiology,* 49(12), 1579–1589. https://doi.org/10.1111/j.1469-8986.2012.01473.x

Osinsky, R., Seeger, J., Mussel, P., & Hewig, J. (2016). Face-induced expectancies influence neural mechanisms of performance monitoring. *Cognitive Affective and Behavioral Neuroscience,* 16, 261–275. https://doi.org/10.3758/s13415-015-0387-y

Proudfit, G. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology,* 52(4), 449–459. https://doi.org/10.1111/psyp.12370

Qu, C., Huang, Y., Wang, Y., & Huang, Y.-X. (2013). The delay effect on outcome evaluation: Results from an event-related potential study. *Frontiers in Human Neuroscience,* 7, 748. https://doi.org/10.3389/fnhum.2013.00748

Rawls, E., Miskovic, V., Moody, S. N., Lee, Y., Shirtcliff, E. A., & Lamm, C. (2020). Feedback-related negativity and frontal midline theta reflect dissociable processing of reinforcement. *Frontiers in Human Neuroscience,* 13, 452. https://doi.org/10.3389/fnhum.2019.00452

Ribas-Fernandes, J. J. F., Shahnazian, D., Holroyd, C. B., & Botvinick, M. M. (2019). Subgoal- and goal-related reward prediction errors in medial prefrontal cortex. *Journal of Cognitive Neuroscience,* 31(1), 8–23. https://doi.org/10.1162/jocn_a_01341

Ribas-Fernandes, J. J. F., Solway, A., Diuk, C., McGuire, J. T., Barto, A. G., Niv, Y., & Botvinick, M. M. (2011). A neural signature of hierarchical reinforcement learning. *Neuron,* 71(2), 370–379. https://doi.org/10.1016/j.neuron.2011.05.042

Riepl, K., Mussel, P., Osinsky, R., & Hewig, J. (2016). Influences of state and trait affect on behavior, feedback-related negativity, and P3b in the ultimatum game. *PLoS ONE,* 11(1), e0146358. https://doi.org/10.1371/journal.pone.0146358

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin,* 126, 3–25. https://doi.org/10.1037/0033-2909.126.1.3

Roesch, M. R., Bryden, D. W., Cerri, D. H., Haney, Z. R., & Schoenbaum, G. (2012). Willingness to wait and altered encoding of time-discounted reward in the orbitofrontal cortex with normal aging. *Journal of Neuroscience,* 32, 5525–5533. https://doi.org/10.1523/JNEUROSCI.0586-12.2012

Roesch, M. R., Taylor, A. R., & Schoenbaum, G. (2006). Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron,* 51, 509–520. https://doi.org/10.1016/j.neuron.2006.06.027

Sambrook, T. D., & Goslin, G. (2015). A neural reward prediction error revealed by a metaanalysis of ERPs using great grand averages. *Psychological Bulletin,* 141(1), 213–235. https://doi.org/10.1037/bul0000006

Sambrook, T. D., & Goslin, J. (2014). Mediofrontal event-related potentials in response to positive, negative and unsigned prediction errors. *Neuropsychologia,* 61, 1–10. https://doi.org/10.1016/j.neuropsychologia.2014.06.004
Sambrook, T. D., Hardwick, B., Wills, A. J., & Goslin, J. (2018). Model-free and model-based reward prediction errors in EEG. *NeuroImage, 178*, 162–171. https://doi.org/10.1016/j.neuroimage.2018.05.023

Santesso, D. L., Dillon, D. G., Birk, J. L., Holmes, A. J., Goetz, E., Bogdan, R., & Pizzagalli, D. A. (2008). Individual differences in reinforcement learning: Behavioral, electrophysiological, and neuroimaging correlates. *NeuroImage, 42*, 807–816. https://doi.org/10.1016/j.neuroimage.2008.05.032

Schmidt, B., Holroyd, C. B., Debener, S., & Hewig, J. (2017). I can’t wait! Neural reward signals in impulsive individuals exaggerate the difference between immediate and future rewards. *Psychophysiology, 54*, 409–415. https://doi.org/10.1111/psyp.12796

Schmidt, B., Mussel, P., Osinsky, R., Rasch, B., Debener, S., & Hewig, J. (2017). Work first then play: Prior task difficulty increases motivation-related brain responses in a risk game. *Biological Psychology, 126*, 82–88. https://doi.org/10.1016/j.biopsych.2017.04.010

Takács, Á., Kóbor, A., Janacsek, K., Honbolygó, F., Csépe, V., & Németh, D. (2015). High trait anxiety is associated with attenuated feedback-related negativity in risky decision making. *Neuroscience Letters, 600*, 188–192. https://doi.org/10.1016/j.neulet.2015.08.012

Thigpen, N. N., Kappenman, E. S., & Keil, A. (2017). Assessing the internal consistency of the event-related potential: An example analysis. *Psychophysiology, 54*(1), 123–138. https://doi.org/10.1111/psyp.12629

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision making under risk. *Science, 315*(5811), 515–518. https://doi.org/10.1126/science.1134239

Tremblay, L., & Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature, 398*, 704–708. https://doi.org/10.1038/19525

Tsypes, A., Angus, D. J., Martin, S., Kemkes, K., & Harmon-Jones, E. (2019). Trait anger and the reward positivity. *Personality and Individual Differences, 144*, 24–30. https://doi.org/10.1016/j.paid.2019.02.030

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297–323. https://doi.org/10.1007/BF00122574

Walsh, M. M., & Anderson, J. R. (2011). Modulation of the feedback-related negativity by instruction and experience. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 19048–19053. https://doi.org/10.1073/pnas.1117189108

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience & Biobehavioral Reviews, 36*(8), 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

Wang, Q., Luo, S., Monterosso, J., Zhang, J., Fang, X., Dong, Q., & Xue, G. (2014). Distributed value representation in the medial prefrontal cortex during intertemporal choices. *Journal of Neuroscience, 34*(22), 7522–7530. https://doi.org/10.1523/JNEUROSCI.0351-14.2014

Yasuda, A., Sato, A., Miyawaki, K., Kuman, H., & Kuboki, T. (2004). Error-related negativity reflects detection of negative reward prediction error. *NeuroReport, 15*, 2561–2565. https://doi.org/10.1097/00001756-200411150-00027

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience, 24*(28), 6258–6264. https://doi.org/10.1523/JNEUROSCI.4537-03.2004

**SUPPORTING INFORMATION**

Additional Supporting Information may be found online in the Supporting Information section.

**Supplementary Material**

**TABLE S1** Numbers of artifact-free trials used for time-domain RewP and FMθ analyses

**FIGURE S1** Grand average ∆RewP (200–300 ms) depending on stimulus-account mapping for the main testing session

**FIGURE S2** Grand average ∆FMθ (200–300 ms) depending on stimulus-account mapping for the main testing session

**How to cite this article:** Rommerskirchen L, Lange L, Osinsky R. The reward positivity reflects the integrated value of temporally threefold-layered decision outcomes. *Psychophysiology*. 2021;58:e13789. https://doi.org/10.1111/psyp.13789