# I knew that! Response-based Outcome Predictions and Confidence Regulate Feedback Processing and Learning

Frömer, R.[1,2], Nassar, M.R.[2], Bruckner, R.,[3,4,5], Stürmer, B.[6], Sommer, W.[1], & Yeung, N.[7]

[1] Humboldt-Universität zu Berlin, [2] Brown University, [3] Freie Universität Berlin, [4] Max Planck School of Cognition, Leipzig, [5] International Max Planck Research School LIFE, Berlin, [6] International Psychoanalytic University, [7] University of Oxford

Romy Frömer

Brown University

Cognitive, Linguistic and Psychological Sciences

190 Thayer Street Providence, RI, USA

romy_fromer@brown.edu

**Abstract**

Influential theories emphasize the importance of predictions in learning: we learn from feedback to the extent that it is surprising, and thus conveys new information. Here we explore the hypothesis that surprise depends not only on comparing current events to past experience, but also on online evaluation of performance via internal monitoring. Specifically, we propose that people leverage insights from response-based performance monitoring – outcome predictions and confidence – to control learning from feedback. In line with predictions from a Bayesian inference model, we find that people who are better at calibrating their confidence to the precision of their outcome predictions learn more quickly. Further in line with our proposal, EEG signatures of feedback processing are sensitive to the accuracy of, and confidence in, post-response outcome predictions. Taken together, our results suggest that online predictions and confidence serve to calibrate neural error signals to improve the efficiency of learning.

*Keywords:* confidence, reinforcement learning, graded feedback

Feedback is crucial to learning and adaptation. Across domains it is thought that feedback drives learning to the degree that it is unexpected and, hence, provides new information, e.g., in the form of prediction errors that express the discrepancy between actual and expected outcomes [1-9]. Yet, the same feedback can be caused by multiple sources: we may be wrong about what is the correct thing to do, or we may know what to do but accidentally still do the wrong thing [10]. When we know we did the latter, we should discount learning about the former [11,12]. Imagine for instance learning to throw darts. You know the goal you want to achieve – hit the bullseye - and you might envision yourself performing the perfect throw to do so. However, you find that the throw you performed as intended missed the target entirely and did not yield the desired outcome: In this case, you should adjust what you believe to be the right angle to hit the bullseye, based on *how* you missed that last throw. On a different throw you might release the dart at a different angle than intended and thus anticipate the ensuing miss: In this case, you may not want to update your beliefs on what is the right angle of throw. How do people assign credit to either of these potential causes of feedback when learning how to perform a new task? How do they regulate how much to learn from a given feedback depending on how much they know about its causes?

Performance monitoring, i.e., the internal evaluation of one's own actions, could reduce surprise about feedback and uncertainty about its causes by providing information about execution errors. For instance in the second dart throw example, missing the target may be unsurprising if performance monitoring detected that, for example, the dart was released differently than desired (cf. Fig. 1A). In simple categorical choices, people are often robustly aware of their response errors [13-16] and this awareness is reflected in neural markers of error detection [17]. Although errors are often studied in simple categorization tasks in which responses

are either correct or incorrect, in many tasks, errors occur on a graded scale (e.g.,, a dart can miss the target narrowly or by a large margin), and both error detection, as well as feedback processing are sensitive to error magnitude [18-21]. People are even able to report gradual errors reasonably accurately [22-24].

This ability may be afforded by reliance on internal models to predict the outcome of movements [25], e.g., based on an efference copy of a motor command. These predictions could help discount execution errors in learning from feedback. In fact, if these predictions perfectly matched the execution error that occurred, the remaining mismatch between predicted and obtained feedback (sensory prediction error) could serve as a reliable basis for adaptation and render feedback maximally informative about the mapping from actions to outcomes (Fig. 1B).

Although participants are able to evaluate their own performance reasonably well, error detection is far less certain than outlined in the ideal scenario above, and the true cause of feedback often remains uncertain to some extent. People are critically sensitive to uncertainty, and learn more from feedback when they expect it to be more informative [1,26-29]. Uncertainty about what caused a given feedback inevitably renders it less informative, similar to decreases in reliability, and this uncertainty should be taken into account when learning from it. Confidence could support such adaptive learning from feedback by providing a read-out of the subjective precision of predicted outcomes [28,30-32], possibly relying on shared neural correlates of confidence with error detection[33,34]. Similar to its role in regulating learning of transition probabilities [31,35], information seeking/exploration in decision making [36,37], and hierarchical reasoning[38], people could leverage confidence to calibrate their use of online predictions. In line with this suggestion, people learn more about advice givers when they are more confident in the choices that advice is about [39]. In the throwing example above, the more confident you are about

the exact landing position of the dart, the more surprised you should be when you find that landing position to be different: The more confident you are, the more evidence you have that your internal model linking angles to landing positions is wrong, and the more information you get about how this model is wrong. Thus, you should learn more when you are more confident. However, this reasoning assumes that your predictions are in fact more precise when you are more confident, i.e., that your confidence is well calibrated (Fig. 1B).

In the present study, we tested the hypothesis that performance monitoring – error detection and confidence [40] – adaptively regulates learning from feedback. This hypothesis predicts that error detection and confidence afford better learning, with confidence mediating the relationship between outcome predictions and feedback, and that learning is compromised when confidence is mis-calibrated (Fig. 1C). It further predicts that established neural correlates of feedback processing, such as the feedback-related negativity (FRN) and the P3a[41], should integrate information about post-response outcome predictions and confidence. That is to say, an error that could be predicted based on internal knowledge of how an action was executed should *not* yield a large surprise (P3a) or reward prediction error (FRN) signal in response to an external indicator of the error (feedback). However, any prediction error should be *more* surprising when predictions were made with higher confidence. We formalize our predictions using a Bayesian model of learning and test them using behavioral and EEG data in a modified time-estimation task.

## Results

### Rationale and Approach

Our hypothesis that performance monitoring regulates adaptive learning from feedback makes two key behavioral predictions (Fig. 1C): 1) The precision of outcome predictions (i.e., the correlation between predicted and actual outcomes) should increase with confidence. 2) Learners with superior calibration of confidence to the precision of their outcome predictions should learn more quickly. Our hypothesis further predicts that feedback processing will be critically modulated by an agent's outcome prediction and confidence. We tested these predictions mechanistically using computational modeling and empirically based on behavioral and EEG data from 40 participants performing a modified time-estimation task (Fig. 1D). In comparison to darts throwing as used in our example, the time estimation task requires a simple response – a button press – such that errors map onto a single axis that defines whether the response was provided too early, timely, or too late and by how much. These errors can be mapped onto a feedback scale and, just as in the darts example where one learns the correct angle and acceleration to hit the bullseye, participants here can learn the target timing interval. In addition to requiring participants to learn and produce a precisely-timed action on each trial, our task also included two key measurements that allowed us to better understand how performance monitoring affects feedback processing: 1) Participants were required to predict the feedback they would receive on each trial and indicate it on a scale visually identical to the feedback scale (Fig. 1D, Prediction) and 2) Participants indicated their degree of confidence in this prediction (Fig. 1D, Confidence). Only following these judgments would they receive feedback about their time estimation performance.

**A Mechanism for Performance Monitoring-Augmented Learning**

As a demonstration of proof of the hypothesized learning principles, we implemented a computational model that uses performance monitoring to optimize learning from feedback in that same task (Fig. 1E). The agent's goal is to learn the mapping between its actions and their outcomes (sensory consequences) in the time-estimation task, wherein feedback on an initially unknown scale must be used to learn accurately-timed actions. Learning in this task is challenged in two ways: First, errors signaled by feedback include contributions of response noise, e.g., through variability in the motor system or in the representations of time [24,42]. Second, the efference copy of the executed response (or the estimate of what was done) varies in its precision. To overcome these challenges, the agent leverages performance monitoring: It infers the contribution of response noise to a given outcome based on an outcome prediction derived from the efference copy, and the degree of confidence in its prediction based on an estimate of the current efference copy noise. The agent then weighs Prediction and Intended Response as a function of Confidence and Response Noise when updating beliefs about the Target and the Scale based on Feedback.

We compare this model to one that has no insights into its trial-by-trial performance, but updates based on feedback and its fidelity due to response noise alone (Feedback), and another model that has insights into its trial-by-trial performance allowing it to generate predictions, and into the average precision of its predictions, but not the precision of its current prediction (Feedback + Prediction). We find that performance improves as the amount of insight into the agent's performance increases (Fig. 1F): The optimally-calibrated Bayesian learner with performance monitoring outperforms both other models. Further, in line with our behavioral predictions, we find in this model that confidence varies with the precision of predictions

(Fig.2A, Fig. S1) and, when varying the fidelity of confidence as a read-out of precision (Confidence Calibration), agents with superior Confidence Calibration learn better (Fig. 2B, Fig. S1). We next sought to test whether participants' behavior likewise displays these hallmarks of our hypothesis.

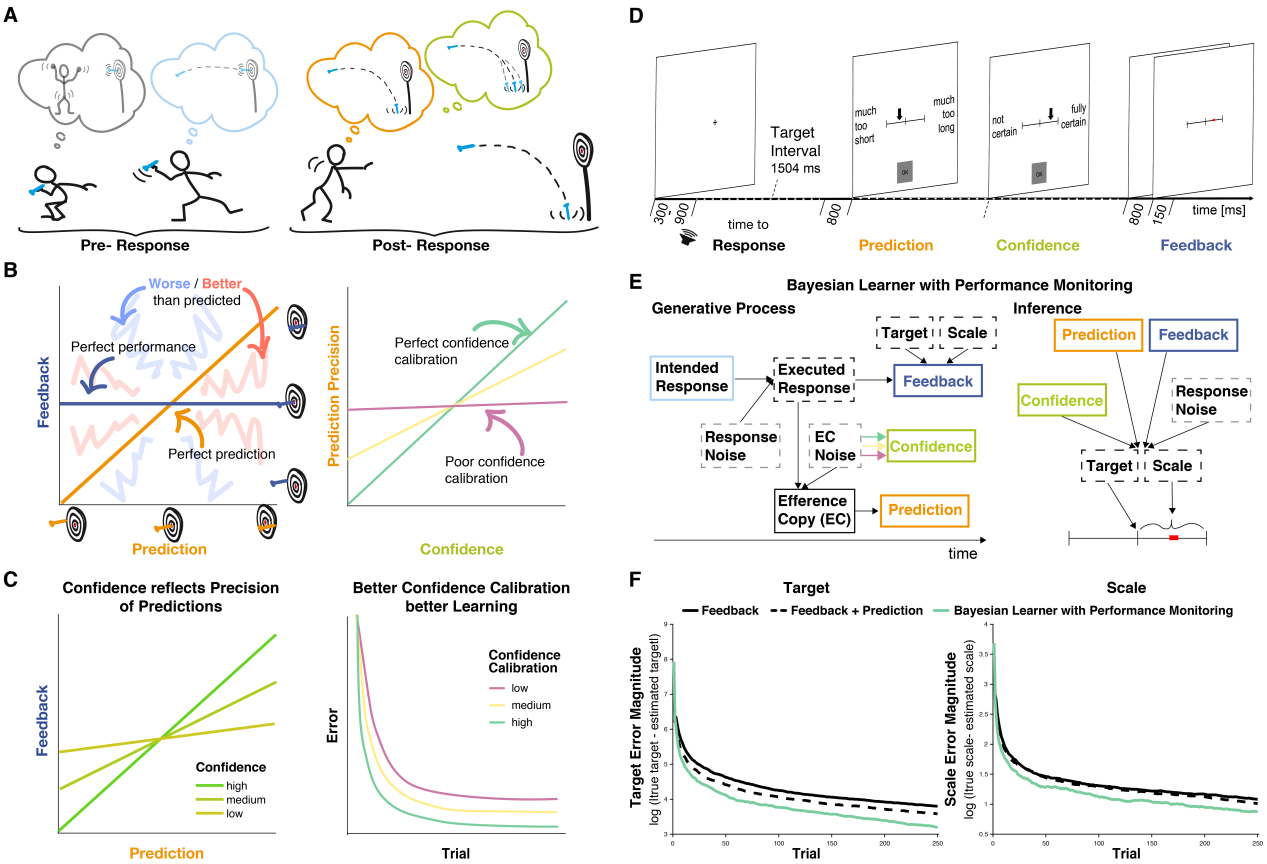

**Figure 1. Interactions between performance monitoring and feedback processing. A** Illustration of dynamic updating of predicted outcomes based on response information. Pre-response the agent aims to hit the bullseye and selects the action he believes achieves this goal. Post-response the agent realizes that he made a mistake and predicts to miss the target entirely, being reasonably confident in his prediction. In line with his prediction and thus unsurprisingly the darts hits the floor. **B** Illustration of key concepts. Left: The feedback received is plotted against the prediction. Performance and prediction can vary in their accuracy independently. Perfect performance (zero deviation from the target, dark blue line) can occur for accurate or inaccurate predictions and any performance, including errors, can be predicted perfectly (predicted error is identical to performance, orange line). When predictions and feedback diverge, outcomes (feedback) can be better (closer to the target, light red shading) or worse (farther from the target, light blue shading) than predicted. The more they diverge the less precise the predictions are. Right: The precision of the prediction is plotted against confidence in that prediction. If confidence closely tracks the precision of the predictions, i.e., if agents know when their predictions are probably right and when they're not, confidence calibration is high (green). If confidence is independent of the precision of the predictions, then confidence calibration is low. **C** Illustration of theoretical hypotheses. Left: We expect the correspondence between predictions and Feedback to be stronger when confidence is high and to be weaker when confidence is low. Right: We expect that agents with better confidence calibration learn better. **D** Trial schema. Participants learned to produce a time interval by pressing a button following a tone with their left index finger. Following each response, they indicated on a visual analog scale in sequence the estimate of their accuracy (anchors: "much too short" = "viel zu kurz" to "much too long" = "viel zu lang") and their confidence in that estimate (anchors: "not certain" = "nicht sicher" to "fully certain" = "völlig sicher") by moving an arrow slider. Finally, feedback was provided on a visual analog scale for 150 ms. The current error was displayed as a red square on the feedback scale relative to the target interval indicated by a tick mark at the center (Target, t) with undershoots shown to the left of the center and overshoots to the right, and scaled relative to the feedback anchors of -/+ 1s (Scale, s; cf. **E**). Participants are told neither Target nor Scale and instead need to learn them based on the feedback. **E** Bayesian Learner with Performance Monitoring. The learner selects an intended response (i) based on the current estimate of the Target. The Intended Response and independent Response Noise produce the Executed Response (r). The Efference Copy (c) of this response varies in its precision as a function of Efference Copy Noise. It is used to generate a Prediction as the deviation from the estimate of Target scaled by the estimate of Scale. The Efference Copy Noise is estimated and expressed as Confidence (co), approximating the precision of the Prediction. Learners vary in their Confidence Calibration (cc), i.e., the precision of their predictions, and higher Confidence Calibration (arrows: green > yellow > magenta) leads to more reliable translation from Efference Copy precision to Confidence. Feedback is provided according to the Executed Response and depends on the Target and Scale, which are unknown to the learner. Target and Scale are inferred based on Feedback (f), Response Noise, Prediction, and Confidence. Variables that are observable to the learner are displayed in solid boxes, whereas variables that are only partially observable are displayed in dashed boxes. **F** Target and scale error (absolute deviation of the current estimates from the true values) for the Bayesian learner with Performance monitoring (green, optimal calibration), a Feedback-only Bayesian Learner (solid black), and a Bayesian Learner with Outcome Prediction (dashed black).

**Confidence reflects precision of outcome predictions**

To test the predictions of our model empirically, we examined behavior of 40 human participants performing the modified time-estimation task. To test whether the precision of outcome predictions increases with confidence, we regressed participants' *signed* timing production errors (signed error magnitude; scale: undershoot (negative) to overshoot (positive)) on their *signed* outcome predictions (Predicted Outcome; same scale as for signed error magnitude), Confidence, Block, as well as their interactions. Our results support our first behavioral prediction (Table 1): As expected, predicted outcomes and actual outcomes were positively correlated, indicating that participants could broadly indicate the direction and magnitude of their errors. Crucially, this relationship between predicted and actual outcomes was stronger for predictions made with higher confidence (Fig. 2C).

In addition to this expected pattern, we found that both outcome predictions, as well as confidence calibration, improved across blocks, suggestive of learning at the level of performance monitoring (Fig. S2). Note however that participants tended to bias their predictions towards the center of the scale in early blocks, when they had little knowledge about the target interval and could thus determine neither over- vs. undershoots nor their magnitude. This strategic behavior may give rise to the apparent improvements in performance monitoring.

**Table 1.** *Relations between actual performance outcome (signed error magnitude), predicted outcome, confidence in predictions and their modulations due to learning across blocks of trials*

| | | | **Signed error magnitude** | | |
|---|---|---|---|---|---|
| *Predictors* | *Estimates* | *SE* | *CI* | *t* | *p* |
| Intercept | 4.63 | 9.99 | -14.94 – 24.20 | 0.46 | 6.427e-01 |
| Predicted Outcome | 523.99 | 29.66 | 465.86 – 582.12 | 17.67 | **7.438e-70** |
| Block | 29.47 | 8.12 | 13.56 – 45.37 | 3.63 | **2.832e-04** |
| Confidence | -27.07 | 11.05 | -48.73 – -5.42 | -2.45 | **1.428e-02** |
| Predicted Outcome: Block | -149.70 | 21.90 | -192.62 – -106.78 | -6.84 | **8.145e-12** |
| Predicted Outcome: Confidence | 322.56 | 27.31 | 269.03 – 376.09 | 11.81 | **3.477e-32** |
| Block: Confidence | -25.52 | 9.15 | -43.46 – -7.58 | -2.79 | **5.297e-03** |
| Predicted Outcome: Block: Confidence | 90.68 | 33.65 | 24.73 – 156.64 | 2.69 | **7.043e-03** |
| **Random Effects** | | | **Model Parameters** | | |
| Residuals | 54478.69 | | N | 40 | |
| Intercept | 3539.21 | | Observations | 9996 | |
| Confidence | 2813.79 | | log-Likelihood | -68816.092 | |
| Predicted Outcome | 22357.33 | | Deviance | 137632.185 | |

*Formula: Signed error magnitude ~ Predicted Outcome*Block*Confidence+(Confidence+ Predicted Outcome +Block|participant); Note: ":" indicates interactions between predictors*

To test more directly our assumption that Confidence tracks the precision of predictions, we followed up on these findings with a complementary analysis of Confidence as the dependent variable and tested how it relates to the precision of predictions (absolute discrepancy between predicted and actual outcome, see sensory prediction error, SPE below), the precision of performance (error magnitude), and how those change across blocks (Table 2). Consistent with our assumption that Confidence tracks the precision of predictions, we find that it increases as the discrepancy between predicted and actual outcome decreases. Confidence was also higher for larger errors, presumably because their direction (i.e., overshoot or undershoot) is easier to judge. The relationships with both the precision of the prediction and error magnitude changed across blocks, and confidence increased across blocks as well.

**Table 2.** *Relations of Confidence with the Precision of Prediction and the Precision of Performance and changes across Blocks*

| | | | **Confidence** | | |
|---|---|---|---|---|---|
| *Predictors* | *Estimates* | *SE* | *CI* | *t* | *p* |
| (Intercept) | 0.26 | 0.04 | 0.18 – 0.33 | 6.35 | **2.187e-10** |
| Block | 0.05 | 0.02 | 0.02 – 0.08 | 3.05 | **2.257e-03** |
| Sensory Prediction Error (SPE) | -0.44 | 0.04 | -0.52 – -0.36 | -10.84 | **2.289e-27** |
| Error Magnitude (EM) | 0.17 | 0.05 | 0.08 – 0.27 | 3.73 | **1.910e-04** |
| Block : SPE | -0.08 | 0.04 | -0.15 – -0.00 | -1.99 | **4.642e-02** |
| Block : EM | 0.15 | 0.05 | 0.05 – 0.25 | 3.07 | **2.167e-03** |
| **Random Effects** | | | **Model Parameters** | | |
| Residuals | 0.12 | | N | 40 | |
| Intercept | 0.06 | | Observations | 9996 | |
| SPE | 0.03 | | log-Likelihood | -3640.142 | |
| Error Magnitude | 0.06 | | Deviance | 7280.284 | |
| Block | 0.01 | | | | |
| Error Magnitude: Block | 0.04 | | | | |

*Formula: Confidence ~ (SPE+ Error Magnitude)*Block+(SPE+Error Magnitude *Block|participant);*
*Note: ":" indicates interactions between predictors*

To test whether these effects reflect monotonic increases in confidence and its relationships with prediction error and error magnitude, as expected with learning, we fit a model with block as a categorical predictor and SPE and Error Magnitude nested within blocks (Supplementary File 1). We found that confidence increased numerically from each block to the next, with significant differences between block 1 and 2, as well as block 3 and 4. Its relationship to error magnitude was reduced in the first block compared to the remaining blocks and enhanced in the final two blocks compared to the remaining blocks. These findings are thus consistent with learning effects. While the precision of predictions was more strongly related to confidence in the final block compared to the remaining blocks, it was not less robustly related in the first block, and instead somewhat weaker in the third block. This pattern is thus not consistent with learning.

Importantly, whereas error magnitude was robustly related to confidence only in the last two blocks, the precision of the prediction was robustly related to confidence throughout.

Having demonstrated that, across individuals, confidence reflects the precision of their predictions (via the correlation with SPE), we next quantified this relationship for each participant separately as an index of their confidence calibration. While quantifying the relationship, we controlled for changes in performance across blocks, and to ease interpretation, we sign-reversed the obtained correlations so that higher values correspond to better confidence calibration. We next tested our hypothesis that confidence calibration relates to learning.

**Superior calibration of confidence judgments relates to superior learning**

To empirically test our second behavioral prediction, that people with better confidence calibration learn faster, we modeled log-transformed trial-wise error magnitude as a function of Trial (linear and quadratic effects to account for non-linearity in learning, i.e., stronger improvements in the beginning), Confidence Calibration for each participant (Fig. 2D inset), and their interaction (Table 3). As expected, Confidence Calibration interacted significantly with the linear Trial component, i.e., with learning (Fig. 2D). Thus, participants with better confidence calibration showed greater performance improvements during the experiment. Importantly, Confidence Calibration did not significantly correlate with overall performance (cf. Fig. 2D inset), supporting the assumption that confidence calibration relates to learning (performance change), rather than performance per se. Confidence calibration was also not correlated with individual differences in response variance (r = - 2.07e-4, 95%CI = [-0.31, 0.31], $p$ = .999), and the interaction of confidence calibration and block was robust to controlling for running average response variance (Supplementary File 2).

**Table 3.** *Confidence calibration modulation of learning effects on performance*

| | | | **log Error Magnitude** | | |
|---|---|---|---|---|---|
| *Predictors* | *Estimates* | *SE* | *CI* | *t* | *p* |
| (Intercept) | 5.17 | 0.06 | 5.05 – 5.30 | 80.74 | **0.000e+00** |
| Confidence Calibration | 0.58 | 0.58 | -0.57 – 1.72 | 0.99 | 3.228e-01 |
| Trial (linear) | -0.59 | 0.07 | -0.72 – -0.45 | -8.82 | **1.197e-18** |
| Trial (quadratic) | 0.16 | 0.02 | 0.11 – 0.20 | 6.80 | **1.018e-11** |
| Trial (linear) : Confidence Calibration | -0.86 | 0.32 | -1.48 – -0.24 | -2.72 | **6.467e-03** |
| **Random Effects** | | | **Model Parameters** | | |
| Residuals | 1.18 | | N | 40 | |
| Intercept | 0.12 | | Observations | 9996 | |
| Trial (linear) | 0.03 | | log-Likelihood | -15106.705 | |
| | | | Deviance | 30213.411 | |

*Formula: log Error Magnitude ~ (Confidence Calibration* Trial(linear)+ Trial(quadratic) + (Trial(linear)|participant); Note: ":" indicates interactions between predictors*

Thus, taken together, our model simulations and behavioral results align with the behavioral predictions of our hypothesis: Participants' outcome predictions were better related to actual outcomes when those outcome predictions were made with higher confidence, and individuals with superior confidence calibration showed better learning.

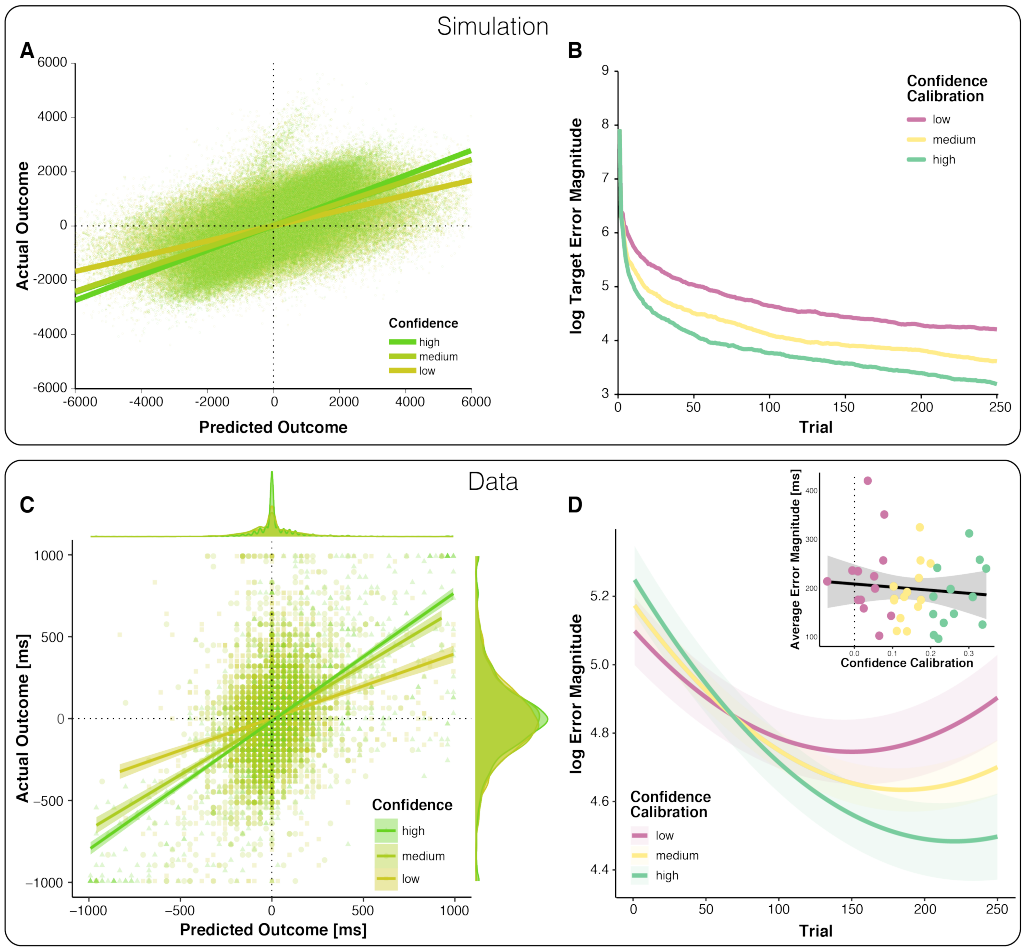


**Figure 2. Relationships between outcome predictions and actual outcomes in the model and observed data (top vs. bottom)**. **A**. Model prediction for the relationship between Prediction and actual outcome (Feedback) as a function of Confidence. The relationship between predicted and actual outcomes is stronger for higher confidence. Note that systematic errors in the model's initial estimates of target (overestimated) and scale (underestimated) give rise to systematically late responses, as well as underestimation of predicted outcomes in early trials, visible as a plume of datapoints extending above the main cloud of simulated data. **B** The model-predicted effect of Confidence Calibration on learning. Better Confidence Calibration leads to better learning. **C**. Observed relationship between predicted and actual outcomes. Each data point corresponds to one trial of one participant; all trials of all participants are plotted together. Regression lines are local linear models visualizing the relationship between predicted and actual error separately for high, medium and low confidence. At the edges of the plot, the marginal distributions of actual and predicted errors are depicted by confidence levels. **D**. Change in error magnitude across trials as a function of confidence calibration. Lines represent LMM-predicted error magnitude for low, medium and high confidence calibrations, respectively. Shaded error bars represent corresponding SEMs. Note that the combination of linear and quadratic effects approximates the shape of the learning curves, better than a linear effect alone, but predicts an exaggerated uptick in errors towards the end, cf. Fig. S3. Inset: Average Error Magnitude for every participant plotted as a function of Confidence Calibration level. The vast majority of participants show positive confidence calibration. The regression line represents a local linear model fit and the error bar represents the standard error of the mean.

**Outcome Predictions and Confidence modulate feedback signals and processing**

At the core of our hypothesis and model lies the change in feedback processing as a function of outcome predictions and confidence. It is typically assumed that learning relies on prediction errors, and signatures of prediction errors have been found in scalp-recorded EEG signals. Before testing directly how feedback is processed, as reflected in distinct feedback related ERP components, we will show how these prediction errors vary over time, and as a function of confidence.

We dissociate three signals that can be processed to evaluate feedback (Fig 3A): The objective magnitude of the error (Error Magnitude) reflects the degree to which performance needs to be adjusted regardless of whether that error was predicted or not. The reward prediction error (RPE), thought to drive reinforcement learning, indexes whether the outcome of a particular response was better or worse than expected. The sensory prediction error (SPE), thought to underlie forward model-based and direct policy learning in the motor domain [43], indexes whether the outcome of a particular response was close to or far off the predicted one. To illustrate the difference between the two prediction errors, one might expect to miss a target 20 cm to the left but find the arrow misses it 20 cm to the right instead. There is no RPE, as the actual outcome is exactly as good or bad as the predicted one, however, there is a large SPE, because the actual outcome is very different from the predicted one.

Our hypothesis holds that predictions should help discount noise in the error signal and more so for higher confidence. Prediction errors should thus be smaller than error magnitude and particularly so when confidence is higher. We find that this is true in our data (Figure 3B, Supplementary File 3 & 4, note that unlike SPE, by definition RPE cannot be larger than error magnitude and that its magnitude, but not sign varies robustly with confidence).

To examine changes in these error signals with trial-to-trial changes in confidence and learning, we regressed each of these signals onto Confidence, Block, and their interaction (Supplementary File 5, cf. Fig 3B). Consistent with our assumption that confidence tracks the precision of predictions, SPE decreased as confidence increased (b = -71.20, $p$ >.001), but there were no significant main effects on error magnitude or reward prediction error. However, Confidence significantly interacted with Block on all variables (Error Magnitude: b = 30.09, $p$ <.001, RPE: b = -64.48, $p$ <.001, SPE: b = 16.99, $p$ = .005), such that in the first block, increased Confidence is associated with smaller Error Magnitudes, less negative RPE, as well as smaller SPE. All error signals further decreased significantly across blocks (Error Magnitude: b = -37.10, $p$ <.001, RPE: b = 36.26, $p$ <.001, SPE: b = - 17.54, $p$ <.001, block wise comparisons significant only from block 1 to 2 cf. Supplementary File 6). These parallel patterns might emerge because prediction errors are derived from and thus might covary with error magnitude. To test whether changes in prediction errors were primarily driven by improvements in error magnitude rather than predictions, we reran the previous RPE and SPE models with error magnitude as a covariate (Supplementary File 7). Controlling for error magnitude notably reduced linear block effects on RPE (b = 36.26 to b = 10.3). It further eliminated block effects on SPE: b = -17.54 to b = 3.65, $p$ = .274,) as well as the interaction of confidence and Block (b = 0.10, $p$ = .984), while the hypothesized main effect of Confidence prevailed (b = -60.12, p <.001).

In summary, we find that all error signals decrease across blocks as performance improves. Although higher confidence is associated with smaller error signals in all three variables early in learning, across all blocks we find that confidence only has a consistent relationship with smaller sensory prediction errors.

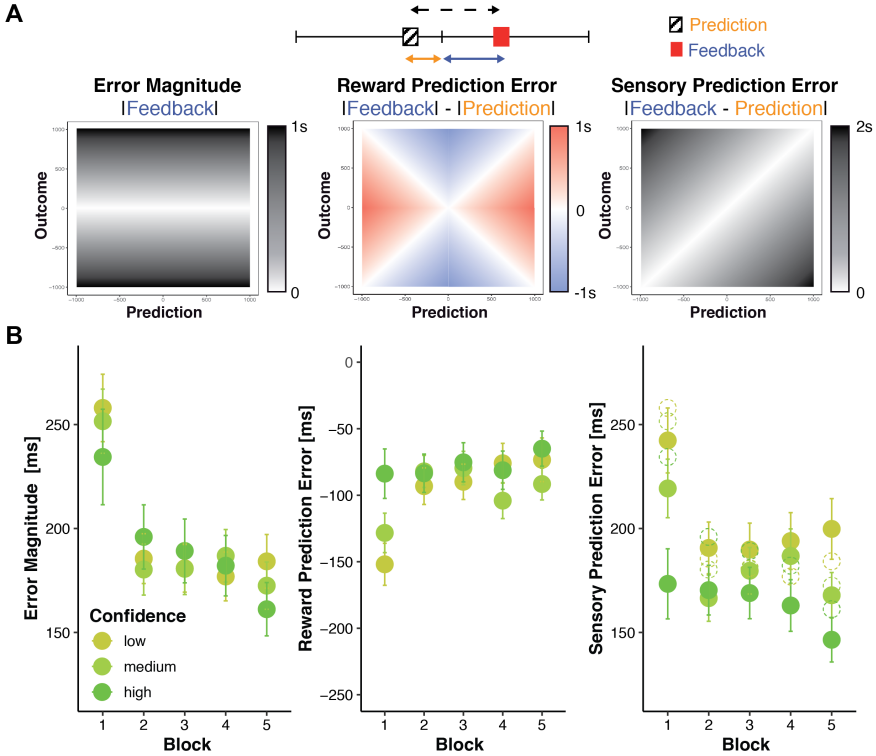


**Figure 3. Changes in objective and subjective feedback. A Dissociable information provided by feedback.** An example for a prediction (hatched box) and a subsequent feedback (red box) are shown overlaid on a rating/feedback scale. We derived three error signals that make dissociable predictions across combinations of predicted and actual outcomes. The solid blue line indicates Error Magnitude (distance from outcome to goal). As smaller errors reflect greater rewards, we computed Reward Prediction Error (RPE) as the signed difference between negative Error Magnitude and the negative predicted error magnitude (solid orange line, distance from prediction to goal). Sensory Prediction Error (SPE, dashed line) was quantified as the absolute discrepancy between feedback and prediction. Values of Error Magnitude (left), RPE (middle), and SPE (right) are plotted for all combinations of prediction (x-axis) and outcome (y-axis) location. **B Predictions and confidence associate with reduced error signals.** Average error magnitude (left), Reward Prediction Error (center) and Sensory Prediction Error (right) are shown for each block and confidence tercile. Average prediction errors are smaller than average error magnitudes (dashed circles), particularly for higher confidence.

Taken together, these results are consistent with our hypothesis that outcome predictions and confidence optimize feedback processing. Accordingly, we predicted that participants' internal evaluations would modulate feedback processing as indexed by distinct feedback-related potentials in the EEG: the feedback-related negativity (FRN), P3a and P3b. Thus, the amplitude of a canonical index of signed RPE [44], the FRN, should increase to the extent that outcomes were worse than predicted, that is, with more negative-going RPE. P3a amplitude, a neural signature

of surprise[45], should increase with the absolute difference between participants' outcome predictions and actual outcomes (i.e., with SPE) and be enhanced in trials in which participants indicated higher confidence in their outcome predictions. To further explore the possible role of performance monitoring in learning, we also tested the joint effects of our experimental variables on the P3b as a likely index of learning[46].

If participants did not take their predictions into account, ERP amplitudes should scale with the actual error magnitude reflected in the feedback (Error Magnitude). Note that both RPE and SPE are equivalent to Error Magnitude in the special case where predicted errors are zero (Fig. 3A), and thus Error Magnitude can be thought of as the default RPE and SPE that would arise if an individual predicted perfect execution on each trial. Thus, if participants did not take knowledge of their own execution errors into account, their FRN and P3a should both simply reflect Error Magnitude. A key advantage of our experimental design is that RPE, SPE and Error Magnitude vary differentially as a function of actual and predicted outcomes (Fig. 3A), which allowed us to test our predictions by examining whether ERP amplitudes are modulated by prediction errors (SPE and RPE) and Confidence, while controlling for other factors including Error Magnitude.

**Reward Prediction Error modulates the Feedback-Related Negativity.** The feedback-related negativity (FRN) is an error-sensitive ERP component with a fronto-central scalp distribution that peaks between 230 to 330 ms following feedback onset [47](Fig. 4A). It is commonly thought to index neural encoding of RPE [44]: Its amplitude increases with the degree to which an outcome is worse than expected and, conversely, decreases to the extent that outcomes are better than expected [48-52]. Its amplitude thus decreases with increasing reward magnitude [20] and reward expectancy [53]. However, it is unknown whether reward prediction errors signaled by

the FRN contrast current feedback with predictions based only on previous (external) feedback, or whether they might incorporate ongoing (internal) performance monitoring. Based on our overarching hypothesis, we predicted that FRN amplitude would scale with our estimate of RPE, which quantifies the degree to which actual feedback was "better" than the feedback predicted after action execution, causing more negative RPEs to produce larger FRN amplitudes (cf. Fig. 4B). A key alternative option is that the FRN indexes the magnitude of an error irrespective of the participant's post-response outcome prediction (e.g., with large FRN to feedback indicating a large error, even when the participant knows to have committed this error) [54-56]. Note that the prediction errors experienced by most error-driven learning models would fall into this alternative category, as they would reflect the error magnitude minus some long term expectation of that magnitude, but not update these expectations after action execution. Thus, to test whether RPE explains variation in FRN above and beyond Error Magnitude, and to control for other factors, we included Error Magnitude, SPE, Confidence, and Block in the model (Table 4).

As predicted, FRN amplitude decreased with more positive-going RPEs ($b = 1.43$, $p < .001$, Fig. 4C), extending previous work that investigated prediction errors as a function of reward magnitude and frequency [44,52]. In contrast, error magnitude and SPE did not significantly affect FRN amplitude, suggesting in the case of the error magnitude that when errors can be accounted for by faulty execution, they do not drive internal reward prediction error signals. We found no other reliable effects and when including interaction terms, they were neither significant, nor supported by model selection ($\Delta X^2(10) = 10.98$, $p = .359$, $AIC_{reduced\text{-}full} = -9$, $BIC_{reduced\text{-}full} = -81$). We conclude that FRN amplitude reflects the degree to which feedback is better than predicted, and critically, that the outcome predictions incorporate information about likely execution errors.

**Table *4.*** *LMM statistics of learning effects on FRN*

| | | | Peak-to-Peak FRN amplitude | | |
|---|---|---|---|---|---|
| *Predictors* | *Estimates* | *SE* | *CI* | *t* | *p* |
| Intercept | -12.67 | 0.49 | -13.62 – -11.71 | -26.03 | **2.322e-149** |
| Confidence | -0.19 | 0.15 | -0.49 – 0.11 | -1.25 | 2.126e-01 |
| Reward Prediction Error | 1.43 | 0.41 | 0.62 – 2.24 | 3.47 | **5.302e-04** |
| Sensory Prediction Error | -0.67 | 0.42 | -1.49 – 0.15 | -1.61 | 1.078e-01 |
| Error Magnitude | 0.51 | 0.55 | -0.57 – 1.58 | 0.92 | 3.553e-01 |
| Block | -0.15 | 0.11 | -0.36 – 0.06 | -1.43 | 1.513e-01 |
| **Random Effects** | | | **Model Parameters** | | |
| Residuals | 27.69 | | $N_{vpn}$ | 40 | |
| Intercept | 9.23 | | Observations | 9678 | |
| Error Magnitude | 2.24 | | log-Likelihood | -29908.910 | |
| Block | 0.22 | | Deviance | 59817.821 | |

*Formula: FRN ~ Confidence +RPE +SPE+EM +Block + (EM + Block|participant)*

**Sensory Prediction Error and Confidence modulate P3a.** The frontocentral P3a is a surprise-sensitive positive-going deflection between 250 and 500 ms following stimulus onset (Fig. 3E)[45]. Its functional significance can be summarized as signaling the recruitment of attention for action to surprising and motivationally relevant stimuli [45,57]. P3a has been shown to increase with larger prediction errors in probabilistic learning tasks [46], higher goal-relevance in a go/no-go task [58], with increasing processing demands [59], and with meta-memory mismatch (feedback about incorrect responses given with high confidence [60]).

Surprise can be quantified using Shannon Information, which reflects the amount of information provided by an outcome given a probability distribution over outcomes[61]. As seen in Fig. 4F, this measure scales with increasing confidence, as well as SPE, i.e., increasing deviations between predicted and actual outcome (margins). To generate these predictions, we computed the Shannon Information for a range of outcomes given a range of predictive

distributions with varying precision, assuming that confidence reflects the precision of a distribution of predicted outcomes (Fig. 4E). Thus, P3a amplitude should scale with both SPE and Confidence. We tested our predictions by examining whether P3a was modulated by SPE, and Confidence, in a model that also included Error Magnitude, RPE and Block as control variables.

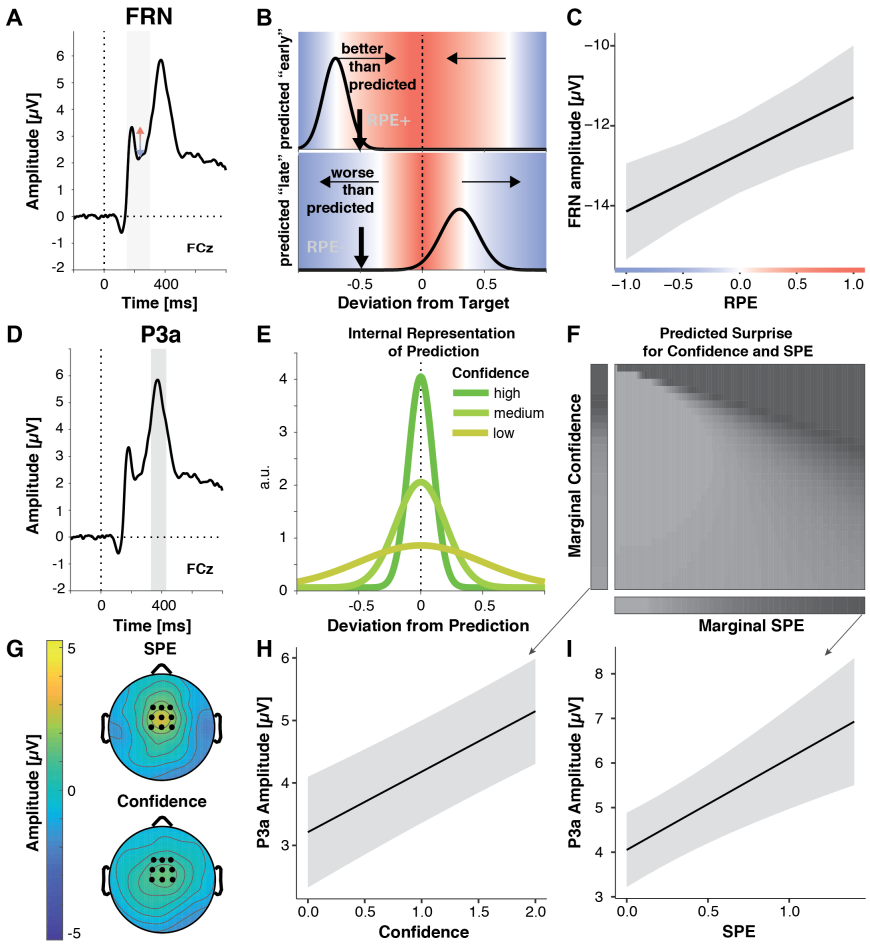

**Figure 4. Multiple Prediction Errors in Feedback Processing. A.- C. FRN amplitude is sensitive to predicted error magnitude. A**. FRN, grand mean, the shaded area marks the time interval for peak-to-peak detection of FRN. Negative peaks between 200 and 300 ms post feedback were quantified relative to positive peaks in the preceding 100 ms time window. **B**. Expected change in FRN amplitude as a function of RPE (color) for two predictions (black curves represent schematized predictive distributions around the reported prediction for a given confidence), one too early (top: high confidence in a low reward prediction) and one too late (bottom: low confidence in a higher reward prediction). Vertical black arrows mark a sample outcome (deviation from the target; abscissa) resulting in different RPE/expected changes in FRN amplitude for the two predictions, indicated by shades. Blue shades indicate negative RPEs/larger FRN, red shades indicate positive RPEs/smaller FRN and grey denotes zero. Note that these are mirrored at the goal for any predictions, and that the likelihood of the actual outcome given the prediction (y-axis) does not affect RPE. In the absence of a prediction or a predicted error of zero, FRN amplitude should increase with the deviation from the target (abscissa). **C**. LMM-estimated effects of RPE on peak-to-peak FRN amplitude visualized with the effects package; shaded error bars represent 95% confidence intervals. **D. – I. P3a amplitude is sensitive to SPE and Confidence. D**. Grand mean ERP with the time-window for quantification of P3a, 330-430 ms, highlighted. **E**. Hypothetical internal representation of predictions. Curves represent schematized predictive distributions around the reported prediction (zero on abscissa). Confidence is represented by the width of the distributions. **F**. Predictions for SPE (x-axis) and Confidence (y-axis) effects on surprise as estimated with Shannon information (darker shades signify larger surprise) for varying Confidence and SPE (center). The margins visualize the predicted main effects for Confidence (left) and SPE (bottom). **G**. P3a LMM fixed effect topographies for SPE, and Confidence. **H. – I**. LMM-estimated effects on P3a amplitude visualized with the effects package; shaded areas in **H.** (SPE) and **I.** (confidence) represent 95% confidence intervals.

As predicted, our analyses showed that P3a amplitude significantly increased with increasing SPEs, in line with the idea of stronger violations of expectations by less accurately predicted outcomes (Fig. 4G, 4I, Table 5), and with increasing Confidence (Fig. 4G, 4H). Our Shannon Information simulation also predicts a small interaction between SPE and Confidence (see slight diagonal component in Fig. 4F). However, when including the interaction term it was not significant and did not improve model fit , $\Delta X^2(10) = 10.36$, $p = .410$ (AIC $_{reduced\text{-}full}$ = -10, BIC $_{reduced\text{-}full}$ = -81), suggesting that any such effect was minimal.

In addition, P3a amplitude decreased across blocks, perhaps reflecting decreased motivational relevance of feedback as participants improved their performance and predictions[58,62]. We also observed a significant decrease of P3a with increasing Error Magnitude and larger P3a amplitudes for more negative reward prediction errors. However, these effects showed more posterior scalp distributions than those of SPE and confidence. As P3a temporally overlaps with the more posteriorly distributed P3b, these effects are likely a spillover of the P3b.

Hence, we discuss them below. Taken together our results support our hypothesis that predictions and confidence shape feedback processing at the level of the P3a.

**Table 5.** *LMM statistics of learning effects on P3a*

| | | | **P3a Amplitude** | | |
|---|---|---|---|---|---|
| *Predictors* | *Estimates* | *SE* | *CI* | *t* | *p* |
| Intercept | 4.10 | 0.42 | 3.28 – 4.93 | 9.79 | **1.293e-22** |
| Confidence | 0.97 | 0.14 | 0.70 – 1.24 | 6.96 | **3.338e-12** |
| Block | -0.91 | 0.07 | -1.05 – -0.77 | -12.93 | **3.201e-38** |
| Sensory Prediction Error | 2.06 | 0.48 | 1.11 – 3.00 | 4.27 | **1.969e-05** |
| Reward Prediction Error | -0.75 | 0.38 | -1.49 – -0.01 | -1.98 | **4.794e-02** |
| Error Magnitude | -1.95 | 0.44 | -2.81 – -1.09 | -4.43 | **9.512e-06** |
| **Random Effects** | | | **Model Parameters** | | |
| Residuals | 22.98 | | N | 40 | |
| Intercept | 6.83 | | Observations | 9678 | |
| SPE | 3.02 | | log-Likelihood | -28997.990 | |
| | | | Deviance | 57995.981 | |

*Formula: P3a ~ Confidence +Block+SPE +RPE +EM + (SPE|participant)*

**Prediction Errors, Objective Errors and Confidence converge in the P3b.** Our overarching hypothesis and model predict that outcome predictions and confidence should affect the degree to which feedback is used for future behavioral adaptation. The parietally distributed P3b scales with learning from feedback [41,46,63-65] and predicts subsequent behavioral adaptation [46,65]. P3b amplitude has been found to increase with feedback salience (reward magnitude irrespective of valence; [63]), behavioral relevance (choice vs. no choice; [66]), with more negative-going RPE [41,67], but also with better outcomes in more complex tasks [68].

Consistent with their necessity for governing behavioral adaptation, P3b was sensitive to participants' outcome predictions (cf. Table 6). P3b amplitude increased with increasing SPE (Fig. 5B, E), indicating that participants extracted more information from the feedback stimulus when outcomes were less expected. As for P3a, this SPE effect decreased across blocks, and so

did overall P3b amplitude, suggesting that participants made less use of the feedback as they improved on the task [46].

P3b amplitude also increased with negative-going RPE (Fig. 5B, C), hence, for worse-than-expected outcomes, replicating previous work [41,67]. This RPE effect was significantly modulated by Confidence and Block, indicating that the main effect needs to be interpreted with caution, and the relationship between P3b and RPE is more nuanced than previous literature suggested. As shown in Fig. 5D in the first block, P3b amplitude was highest for large negative RPE and *high* Confidence, whereas in the last block it was highest for large negative RPE and *low* Confidence (see below for follow-up analyses).

**Table 6.** *LMM statistics of learning effects on P3b*

| | | | **P3b Amplitude** | | |
|---|---|---|---|---|---|
| *Predictors* | *Estimates* | *SE* | *CI* | *t* | *p* |
| Intercept | 4.12 | 0.29 | 3.55 – 4.70 | 14.12 | **2.937e-45** |
| Block | -0.48 | 0.09 | -0.66 – -0.30 | -5.20 | **2.037e-07** |
| Confidence | 0.08 | 0.20 | -0.31 – 0.48 | 0.42 | 6.740e-01 |
| Reward Prediction Error | -1.12 | 0.46 | -2.03 – -0.22 | -2.43 | **1.493e-02** |
| Sensory Prediction Error | 1.75 | 0.47 | 0.84 – 2.66 | 3.76 | **1.691e-04** |
| Error Magnitude | -2.35 | 0.46 | -3.24 – -1.45 | -5.14 | **2.743e-07** |
| Confidence : Reward Prediction Error | -0.51 | 0.55 | -1.60 – 0.57 | -0.92 | 3.556e-01 |
| Block : Confidence | 0.07 | 0.18 | -0.28 – 0.43 | 0.41 | 6.823e-01 |
| Block : Reward Prediction Error | -0.52 | 0.44 | -1.39 – 0.34 | -1.19 | 2.359e-01 |
| Block : Sensory Prediction Error | -0.98 | 0.46 | -1.88 – -0.07 | -2.12 | **3.405e-02** |
| Block : Confidence : Reward Prediction Error | 2.22 | 0.72 | 0.81 – 3.64 | 3.08 | **2.057e-03** |
| **Random Effects** | | | | | |
| Residuals | 23.95 | | N | 40 | |
| Intercept | 3.17 | | Observations | 9678 | |
| Sensory Prediction Error | 2.16 | | log-Likelihood | -29197.980 | |
| Reward Prediction Error | 1.67 | | Deviance | 58395.960 | |
| Confidence | 0.63 | | | | |

*Formula: P3b ~ Block*(Confidence*RPE+ SPE)+Error Magnitude + (SPE+RPE+Confidence|participant);*
*Note: ":" indicates interactions*

In line with previous findings [68,69], we further observed significant increases of P3b amplitude with decreasing Error Magnitude, thus, with better outcomes (Fig. 5B, F, Table 5). We found no further significant interactions, and excluding the non-significant interaction terms from the full model did not significantly diminish goodness of fit, $\Delta X^2(5) = 10.443$, $p = .064$ ($AIC_{reduced\text{-}full} = 0$; $BIC_{reduced\text{-}full} = -35$).

Our hypothesis states that the degree to which people rely on their predictions when learning from feedback should vary with their confidence in those predictions. In the analysis above, we observed such an interaction with confidence only for RPE (and Block). RPE is derived from the contrast between Error Magnitude and Predicted Error Magnitude, and changes in either variable or their weighting could drive the observed interaction. To better understand this interaction and test explicitly whether confidence regulates the impact of predictions, we therefore ran complementary analyses where instead of RPE we included Predicted Error Magnitude (Table 7). Confirming its relevance for the earlier interaction involving RPE, Predicted Error Magnitude indeed interacted with Confidence and Block. Consistent with confidence-based regulation of learning, a follow-up analysis showed that in the first block, P3b significantly increased with higher Confidence, and importantly decreased significantly more with increasing Predicted Error Magnitude as Confidence increased (Supplementary File 8). Main effects of Predicted Error Magnitude emerged only in the late blocks when participants were overall more confident.

Hence, our P3b findings indicate that early on in learning, when little is known about the task, participants learn more and discount their predictions more when they have high confidence in those predictions. In later trials however, when confidence is higher overall, participants discount their predicted errors even when confidence is relatively lower.

*Table 7. LMM Statistics of Confidence weighted Predicted Error discounting on P3b*

| | | | **P3b Amplitude** | | |
|---|---|---|---|---|---|
| *Predictors* | *Estimates* | *SE* | *CI* | *t* | *p* |
| Intercept | 4.26 | 0.30 | 3.68 – 4.85 | 14.22 | **7.239e-46** |
| Confidence | 0.31 | 0.22 | -0.12 – 0.75 | 1.41 | 1.595e-01 |
| Predicted Error Magnitude | -0.83 | 0.46 | -1.74 – 0.07 | -1.80 | 7.133e-02 |
| Block | -0.32 | 0.11 | -0.52 – -0.11 | -2.98 | **2.860e-03** |
| Error Magnitude | -1.06 | 0.49 | -2.03 – -0.09 | -2.13 | **3.277e-02** |
| Sensory Prediction Error | 1.49 | 0.40 | 0.71 – 2.28 | 3.72 | **1.992e-04** |
| Confidence : Predicted Error Magnitude | -0.98 | 0.69 | -2.34 – 0.38 | -1.41 | 1.582e-01 |
| Confidence : Block | -0.50 | 0.20 | -0.90 – -0.11 | -2.50 | **1.249e-02** |
| Predicted Error Magnitude : Block | -1.12 | 0.56 | -2.22 – -0.02 | -2.00 | **4.540e-02** |
| Confidence : Predicted Error Magnitude : Block | 3.12 | 0.84 | 1.47 – 4.78 | 3.70 | **2.141e-04** |

| **Random Effects** | | **Model Parameters** | |
|---|---|---|---|
| Residuals | 23.98 | N | 40 |
| Intercept | 3.30 | Observations | 9678 |
| Error Magnitude | 3.43 | log-Likelihood | -29201.951 |
| Confidence | 0.72 | Deviance | 58403.902 |

*Formula: P3b ~ Block*(Confidence*Predicted Error Magnitude+ SPE)+Error Magnitude + (Error Magnitude+Confidence|participant); Note: ":" indicates interactions*

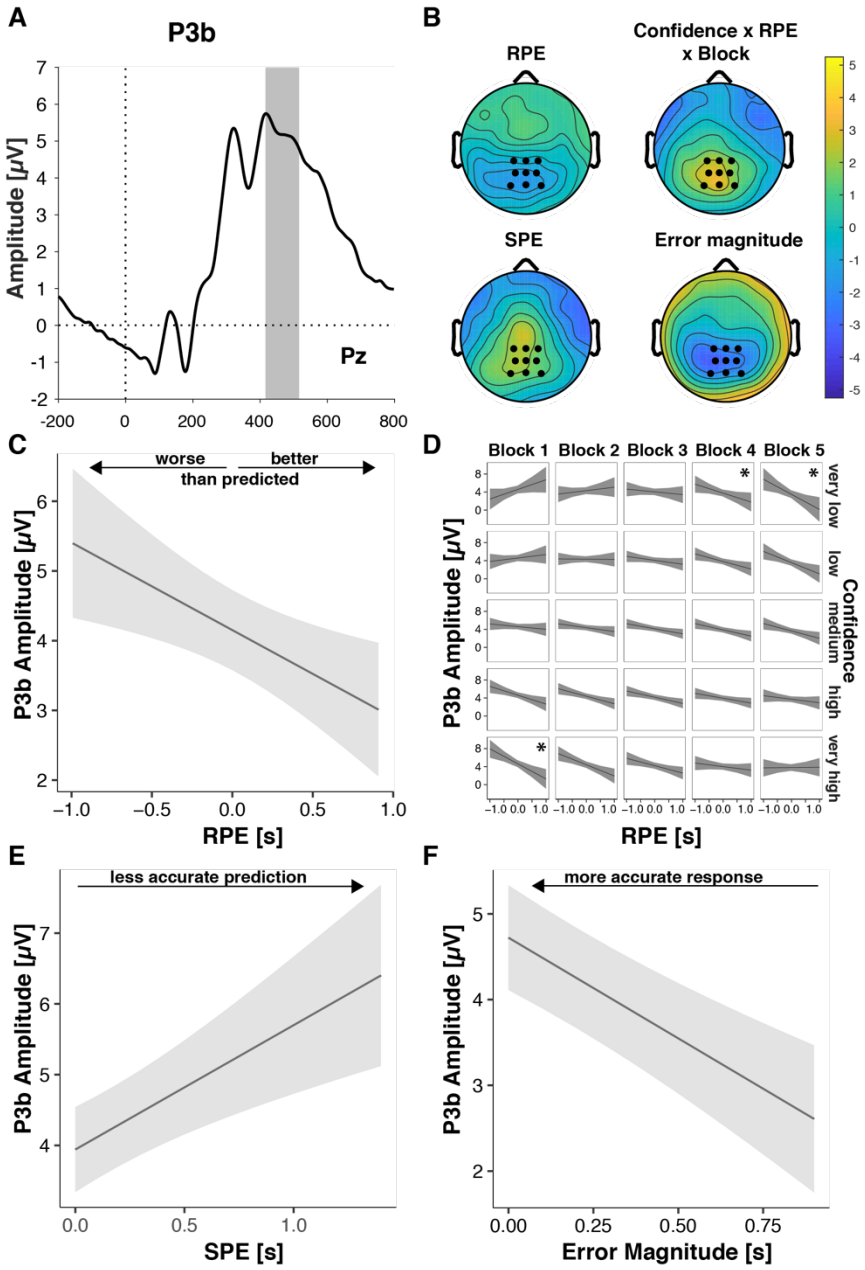


**Figure 5. Performance-relevant information converges in the P3b**. **A**. Grand average ERP waveform at Pz with the time window for quantification, 416-516 ms, highlighted. **B**. Effect topographies as predicted by LMMs for RPE, error magnitude, SPE and the RPE by Confidence by Block interaction. **C. - F**. LMM-estimated effects on P3b amplitude visualized with the effects package in R; shaded areas represent 95% confidence intervals. **C**. RPE. Note the interaction effects with Block and Confidence (D), that modulate the main effect **D**. Three-way interaction of RPE, Confidence and Block. Asterisks denote significant RPE slopes within cells. **E**. P3b amplitude as a function of SPE. **F**. P3b amplitude as a function of Error Magnitude.

We next explored whether P3b amplitude is associated with trial-by-trial adjustments. To that aim, we computed the improvement on trial n as the difference between the error on trial n and the error on trial n-1. Time-estimation responses are noisy, and thus provide only a coarse trial-by-trial indicator of learning. Consistent with regression to the mean, where larger errors are more likely followed by smaller errors, improvements increased with the magnitude of the error on the previous trial (b = 0.85, $p$ <.001, Supplementary File 9). We find, however, that this effect varies across blocks (b = 0.13, $p$ <.001), and is least pronounced in the first block when most learning takes place (Block 1: b = .66, $p$ <.001; Block 2 - 5: b >= 0.90, $p$ < .001, Supplementary File 10). We thus next tested whether P3b on trial n-1 mediates the relationship between error magnitude on trial n-1 and the improvement on the current trial, leading to stronger improvements following a given error, particularly in the first block when most learning takes place and performance is less determined by previous error alone. Indeed, we found a significant 3-way interaction between previous error magnitude, previous P3b amplitude and Block (b = -0.03, $p$ = .031, Supplementary File 9, Fig. S4) on improvement. A follow-up analysis confirmed that P3b mediated the relationship between previous error magnitude and improvement in the first block (b = 0.06, $p$ <.001, Supplementary File 10). This interaction was not significant within any of the remaining blocks. While intriguing and in line with previous work linking P3b to trial-by-trial adjustments in behavior, these results should be interpreted with a degree of caution given that the present task is not optimized to test for trial-to-trial adjustments in behavior.

Taken together, our ERP findings support our main hypothesis that individuals take their internal evaluations into account when processing feedback, such that distinct ERP components reflect different aspects of internal evaluations rather than just signaling objective error.

## Discussion

The present study explored the hypothesis that learning from feedback depends on internal performance evaluations as reflected in outcome predictions and confidence. Comparing different Bayesian agents with varying insights into their trial-by-trial performance, we show that performance monitoring provides an advantage in learning, as long as agents' confidence is accurately calibrated. To test our hypothesis empirically, we collected participants' trial-wise outcome predictions and confidence judgments in a time-estimation task prior to receiving feedback, while recording EEG. Like the simulations from the Bayesian learner with performance monitoring, our empirical results show that trial-by-trial confidence tracks the precision of outcome predictions, and individuals with better coupling between confidence and the precision of their predictions (confidence calibration) showed greater improvements in performance over the course of the experiment. Moreover, participants' subjective predictions, as well as their confidence in those predictions, influenced feedback processing as revealed by feedback-related potentials.

Our study builds on an extensive body of work on performance monitoring, proposing that deviations from performance goals are continuously monitored, and expectations are updated as soon as novel information becomes available [44,70]. Hence, performance monitoring at later stages should depend on performance monitoring at earlier stages [44]. Specifically, learning from feedback should critically depend on internal performance monitoring. Our results extend previous work demonstrating a shift from feedback-based to response-based evaluations as learning progresses [71,72]: They show that performance monitoring and learning from feedback are not mutually exclusive modes of performance evaluation; instead, they operate in concert, with

confidence in response-based outcome predictions determining the degree to which this information is relied on.

Participants' behavior displayed hallmarks of error monitoring [22-24], such that outcome predictions tracked factual errors in both direction and magnitude. Crucially, extending those previous findings, our empirical results align with unique predictions based on our hypothesis: Confidence reflected the precision of participants' outcome predictions, and participants with superior calibration of their confidence judgments to the accuracy of their predictions learned better than those with poorer calibration. This latter finding is notable given that overall confidence calibration was similar for participants with different performance quality (error magnitude, response variance). Therefore, the empirical confidence calibration effect on learning is unlikely to be a consequence of better overall ability as described in the "unskilled and unaware effect" [73] or of the dependence of confidence calibration (or metacognitive sensitivity) on performance [74]. Instead, the finding supports our hypothesis that confidence supports learning via optimized feedback processing.

Our simulations and ERP data reveal two critical mechanisms through which performance monitoring may impact learning from feedback: modulation of surprise and reduction of uncertainty via credit assignment. The main impact of error monitoring is to *reduce* the surprise about outcomes. All else being equal, a given outcome is less surprising the better it matches the predicted outcome. Consistent with discounting of predicted deviations from the goal, we found that participants' trial-by-trial subjective outcome predictions consistently modulated feedback-based evaluation reflected in ERPs as evidenced by prediction error effects. Participants' response-based outcome predictions were reflected in the amplitudes of FRN reward prediction error signals [44,50,52,75], of P3a surprise signals, as well as the P3b signals

combining information about reward prediction error and surprise. In our computational models, reducing surprise by taking response-based outcome predictions into account led to more accurate updating of internal representations supporting action selection and thus superior learning.

Learning – in our computational model and in our participants - was further supported by the adaptive regulation of uncertainty-driven updating via confidence. Specifically, as deviations from the goal were predicted with higher confidence, these more precise outcome predictions *enhanced* the surprise elicited by a given prediction error. This mechanism implemented in our model is mirrored in participants' increased P3a amplitudes for higher confidence, and further reflected in confidence-weighted impacts of predicted error magnitude on P3b, as well as larger P3b amplitudes for higher confidence in the first block when most learning took place. Thus, a notable finding revealed by our simulations and empirical data is that, counterintuitively, agents and participants learned *more* from feedback when confidence in their predictions had been high.

While FRN amplitude was not modulated by confidence, we found that P3a increased with confidence, as predicted by uncertainty-driven changes in surprise. Our results align with previous findings of larger P3a amplitude for metacognitive mismatch [60] and offer a computational mechanism underlying previous theorizing that feedback about errors committed with high confidence attracts more attention, and therefore leads to hypercorrection [76,77]. We also found that confidence modulated the degree to which predicted error magnitude reduced P3b amplitude, such that in initial blocks, where most learning took place, predicted error magnitude effects were amplified for higher confidence, whereas this effect diminished in later blocks, where predicted error magnitude effects were present also for low confidence (and performance and prediction errors were attenuated when confidence was high). This shift is intriguing and

may indicate a functional change in feedback use as certainty in the response-outcome mapping increases and less about this mapping is learned from feedback, but the effect was not directly predicted and therefore warrants further research and replication.

Confidence has typically been studied in two-alternative choice tasks, and only rarely in relation to continuous outcomes [31,35,37,78-80]. By reconceptualizing error detection as outcome prediction, our results shed new light on the well-supported claim that error monitoring and confidence are tightly intertwined [33,40,81-83] and forge valuable links between research on performance monitoring[41,44,70] and on learning under uncertainty[1,3,61,84]. In doing so, our results provide further evidence to the growing literature on the role of confidence in learning and behavioral adaptation [35-37,85].

While we captured the main effects of interest with our Bayesian model and our key behavioral results are in line with our overarching hypothesis, our behavioral findings also reveal other aspects of learning that remain to be followed up on. Unlike our Bayesian agents, participants exhibited signatures of learning not only at the level of first order performance, but also at the level of performance monitoring. The precision of their outcome predictions increased as learning progressed, as did confidence. Identifying the mechanisms that drive this metacognitive learning, i.e., whether changes in confidence follow the uncertainty in the internal model or reflect refinement of the confidence calibration to the efference copy noise, is an exciting question for future work.

Anyone who tried to learn a sport can relate to the intuition that just because you find out what you did was wrong doesn’t mean you know how to do it right. Our task also evokes this so-called distal problem, which refers to the difficulty of translating distal sensory outcomes of responses (e.g., the location of a red dot on a feedback scale) to required proximal movement

parameter changes (changes in the timing of the response). Indeed, when practicing complex motor tasks, individuals prefer and learn better from feedback following successful trials compared to error trials [86-88]. In line with the notion that in motor learning feedback about success is more informative than feedback about failure, we, like others in the time estimation task [68,69], observed increasing P3b amplitude after feedback about more accurate performance (i.e., for smaller error magnitude), in addition to prediction error effects.

In our study, the P3b component, previously shown to scale with the amount of information provided by a stimulus[89] and the amount of learning from a stimulus[46], was sensitive to both RPE and SPE, indicating that multiple learning mechanisms may act in parallel, supported by different aspects of feedback. Our findings resonate with recent work in rodents showing that prediction error signals encode multiple features of outcomes[90], and are based on distributed representations of predictions [91]. This encoding of multiple features of outcomes, like the uncertainty in predictions, may help credit assignment and support learning at multiple levels. It is still unclear to what degree different learning mechanisms – error-based, model-based, reinforcement learning – contribute to motor learning[92]. Further research is needed to identify whether the same or different learning mechanisms operate across levels, e.g., via hierarchical reinforcement learning[93,94], and how learning interacts between levels.

Taken together, our findings provide evidence that feedback evaluation is fundamentally affected by an individual's internal representations of their own performance at the time of feedback. These internal representations in turn influence how people learn and thus which beliefs they will have and which actions they will take, driving what internal and external states they will encounter in the future. The present study is a first step towards elucidating this

recursive process of performance optimization via internal performance monitoring and monitoring of external task outcomes.

## Method

### Task variables

- $t$ denotes the target interval, which was set to $t := 19$, (this simulation parameter choice was necessarily somewhat arbitrary, and choosing a different parameter does not change the model's behavior)
- $s$ denotes the feedback scale, which was set to $s := 90$,
- $r$ denotes the model's or participant's response,
- $f$ denotes the feedback in the task, which was defined as

$$f := (r - t)s \tag{1}$$

### Computational model

The Bayesian learner with performance monitoring attempted to sequentially infer the target interval $t$ and feedback scale $s$ (defining how the magnitude of a given response error translates to the magnitude of the error displayed on the visual feedback scale) over multiple trials, based on its intended response $i$, an efference copy $c$ of its executed response and feedback $f$ indicating the magnitude and direction of its timing errors. On each trial the model computed its intended response based on the inferred target interval. During learning, the model faced several obstacles including 1) the initially unknown scale of the feedback, making it difficult to judge whether feedback indicates small or large timing errors, 2) response noise, which offsets executed responses from intended ones, and 3) efference copy noise, which makes the efference copy unreliable to a degree that varies from trial to trial. Formally, the Bayesian learner with performance monitoring is represented by the following variables:

- $p(t) \coloneqq U(t; [0,100])$ denotes the model's prior distribution over the target interval $t$, as a uniform distribution (denoted by U throughout) of over possible values of t within the range 0 to 100.
- $p(s) \coloneqq U(s; [0.1, 100])$ denotes the model's prior distribution over the feedback scale $s$, as a uniform distribution over possible values of s within a range of 0.1 to 100.
- $p^{\sigma_r^2}(r|i) \coloneqq N(r; i, \sigma_r^2)$ denotes the model's response distribution (N denoting normal distributions throughout), where $i$ denotes the model's intended response, which corresponds to the expected target interval $i \coloneqq E_{p(t)} = \sum_t p(t)t$, and $\sigma_r^2$ denotes the response noise, which was set to $\sigma_r$ := 10 in terms of the standard deviation
- $p^{\sigma_c^2}(c|r) \coloneqq N(c; r, \sigma_c^2)$ denotes the model's efference-copy distribution with efference-copy noise expressed as standard deviation $\sigma_c \in \{5, 10,20\}$, where $p(\sigma_c) \coloneqq Cat(\sigma_c, 1/3)$. Here we assumed that the model was aware of its trial-by-trial efference copy noise. That is, from the perspective of the Bayesian learner with performance monitoring, efference-copy noise was not a random variable.

*Intended response*

During the response phase in the task, the model first computed its intended response $i$ and then sampled its actual response $r$ from the Gaussian response distribution. We assumed that the model's internal response monitoring system subsequently generated the noisy efference copy $c$.

*Learning*

Based on the definition of the task and the Bayesian learner with performance monitoring, the joint distribution over the variables of interest during a trial of the task is given by

$$p^{i,\sigma_c^2,\sigma_r^2}(t,s,f,r,c) := p(f|r,t,s)p^{\sigma_c^2}(c|r)p^{i,\sigma_r^2}(r)p(t,s) \tag{2}$$

To infer the target interval $t$ and the feedback scale $s$, we can evaluate the posterior distribution conditional on the efference copy $c$ and feedback $f$ and given the intended response $i$, response noise $\sigma_r^2$ and efference copy noise $\sigma_c^2$ according to Bayes' rule:

$$\begin{aligned} p^{i,\sigma_c^2,\sigma_r^2}(t,s|\, c,f) \\ &\propto \int p^{i,\sigma_c^2,\sigma_r^2}(t,s|c,f,r)\, dr \\ &\propto \int p(f|r,t,s)p^{\sigma_c^2}(c|r)p^{i,\sigma_r^2}(r)p(t,s)\, dr \end{aligned} \tag{3}$$

Note that we assumed that the Bayesian learner with performance monitoring was aware how feedback $f$ was generated in the task (cf. eq. 1), that is, conditional on the response $r$, target interval $t$ and feedback scale $s$, the model was able to exactly compute the probability of the feedback according to

$$p(f|r,t,s) = \begin{cases} 1, & if\ f = (r-t)s \\ 0, & else \end{cases} \tag{4}$$

We approximated inference using a grid over the target interval $t \in [0, 100]$ and feedback scale $s \in [0.1, 100]$. The model first computed the probability of the currently received feedback. Although it was aware how feedback was generated in the task, it suffered from uncertainty over its response due to noise in the efference copy. For each *s* and *t* on the grid, the model evaluated the Gaussian distribution

$$p^{i,\sigma_c^2,\sigma_r^2}(f|c,t,s) := N(f; (m-t)^T s, vA) \tag{5}$$

where A was a $100 \times 50$ matrix containing the grid values of *t* and *s*.

$$v = \frac{1}{\frac{1}{{\sigma_c}^2} + \frac{1}{{\sigma_r}^2}} \tag{6}$$

denotes the expected variance in feedback under consideration of both efference-copy noise ${\sigma_c}^2$ and response noise ${\sigma_r}^2$ and

$$m = v\frac{1}{{\sigma_c}^2}c + v\frac{1}{{\sigma_r}^2}i \tag{7}$$

denotes the expected feedback under the additional consideration of efference copy *c* and intended response *i.* When computing the probability of the feedback, our model thus took into account the efference copy *c* and the response it intended to produce *i*, which were weighted according to their respective reliabilities.

Second, the model multiplied the computed probability of the observed feedback with the prior over the scale and target response, i.e.,

$$p^{i,\sigma_c^2,\sigma_r^2}(t,s|c,f) \propto N(f;(m-t)^T s, vA)p(s,t) \tag{8}$$

In the grid approximation, the model started with a uniform prior on the joint distribution over $t$ and $s$ and was applied recursively, such that the posterior joint distribution on $t$ and $s$ for each trial served as the prior distribution for the subsequent one.

*Outcome prediction*

On each trial, the model reports an outcome prediction $po$ and the confidence in this prediction, which we refer to as $co$. The outcome prediction maps the discrepancy between the intended

response and the efference copy onto the feedback scale given the best guess of the current *t* and *s*.

$$po = (c - i) \sum_{s} p(s)s \tag{9}$$

Note that this outcome prediction is different from the mean of the uncertainty-weighted expectancy distribution defined in eq. 7, in that it does not take uncertainty into account, but reflects the expectation given the efference copy alone. The inverse uncertainty $\sigma_c^2$ of the efference copy, as described below, is translated into the agent's confidence report.

*Confidence Calibration*

Confidence calibration $cc \in \{0, 0.75, 1\}$ denotes the probability that the agent assumes the correct efference copy variance $\sigma_c^2$ for learning (c.f. eqs. 6-7). $cc = 1$ indicates that the subjectively assumed efference-copy precision is equal to the true precision and $cc = 0$, in contrast, indicates that the assumed precision of the efference copy is different from the true one. That is, in the case of $cc = 0.75$, the agent more likely assumes the true precision of the efference copy but sometimes fails to take it accurately into account during learning. As shown in Fig. 1, we simulated behavior of three agents that differed in their confidence calibration according to this idea.

*Confidence report*

In our model, the confidence report $\mathrm{co} \in \{3, 2, 1\}$ reflects how certain the agent thinks its efference copy is, where 3 refers to "completely certain", and 1 to "not certain". In particular,

$$co = \begin{cases} 3, \ if \ \sigma_c = 5 \\ 2, \ if \ \sigma_c = 10 \\ 1, \ if \ \sigma_c = 20 \end{cases} \tag{10}$$

That is, the confidence report is directly related to the subjective precision of the efference copy, which, as shown above, depends to the agent's level of confidence calibration.

*Model with incomplete performance monitoring*

We also applied a model that had no insight into the precision of its current predictions. The agent was thus unaware of its trial-by-trial efference-copy noise $\sigma_c$ and therefore relied on the average value of $\sigma_c$, i.e., $\sum \sigma_c p(\sigma_c) = 12$. As the model does not differentiate between precise and imprecise predictions but treats each prediction as average precise, it relies too much on imprecise predictions and too little on precise ones.

*Model without performance monitoring*

Finally, we applied a model that was aware about its response noise but, because it completely failed to consider its efference copy, it lacked insight into its trial-by-trial performance. In this version, we had

$$v = \frac{1}{\frac{1}{{\sigma_r}^2}} \tag{11}$$

and $m = i$.

This model accounts for the expected variance in the feedback, but cannot differentiate between trials in which the feedback is driven more by incorrect beliefs about the target versus incorrect execution. Therefore, it adjusts its beliefs too much following feedback primarily driven by execution errors and too little following feedback primarily driven by incorrect beliefs.

**Participants**

The experimental study included 40 participants (13 males) whose average age was 25.8 years (*SD* = 4.3) and whose mean handedness score [95] was 63.96 (*SD* = 52.09; i.e., most participants were right-handed). Participants gave informed consent to the experiment and were remunerated with course credits or 8 € per hour.

**Task and procedure**

Participants performed an adapted time-estimation task [18,47] that included subjective accuracy and confidence ratings (similar to [22-24]). Participants were instructed that their primary goal in this task is to learn to produce an initially unknown time interval. In addition, they were asked to predict the direction and magnitude of any errors they produced and their confidence in those predictions. The time-estimation task is well established for ERP analyses e.g.,[18,47], and has the advantages that it limits the degrees of freedom of the response, and precludes concurrent visual feedback that might affect performance evaluation. The task consisted of four parts on each trial, illustrated in Fig. 1B. After a fixation cross lasting for a random interval of 300-900 ms, a tone (600 Hz, 200 ms duration) was presented. Participants' task was to terminate an initially unknown target interval of 1504 ms from tone onset, by pressing a response key with their left hand. We chose a supra-second duration to make the task sufficiently difficult [18]. Following the response, a fixation cross was presented for 800 ms. Participants then estimated the accuracy of the interval they had just produced by moving an arrow on a visual analogue scale (*too short – too long*; ± 125 pixel, 3.15 ° visual angle) using a mouse cursor with their right hand. Then, on a scale of the same size, participants rated their confidence in this estimate (*not certain – fully certain*). The confidence rating was followed by a blank screen for 800 ms. Finally, participants received feedback about their performance with a red square (0.25° visual

angle) placed on a scale identical to the accuracy estimation scale but without any labels. The placement of the square on the scale visualized error magnitude in the interval produced, with undershoots shown to the left and overshoots on the right side of the center mark, indicating the correct estimate. Feedback was presented for only 150 ms to preclude eye movements. The interval until the start of the next trial was 1500 ms.

The experiment comprised five blocks of 50 trials each, with self-paced rests between blocks. We used Presentation software (Neurobs.) for stimulus presentation, event and response logging. Visual stimuli were presented on a 4/3 17'' BenQ Monitor (resolution: 1280x1024, refresh rate: 60Hz) placed at 60 cm distance from the participant. A standard computer mouse and a customized response button (accuracy 2 ms, response latency 9 ms) were used for response registration.

Prior to the experiment, participants filled in demographic and personality questionnaires: Neuroticism and Conscientiousness Scales of NEO PI-R [96] and the BIS/BAS scale [97], as well as a subset of the Raven's progressive matrices as an index for figural-spatial intelligence [98]. These measures were registered as potential control variables and for other purposes not addressed here. Participants were then seated in a shielded EEG cabin, where the experiment including EEG recording was conducted. Prior to the experiment proper, participants performed three practice trials.

**Psychophysiological recording and processing**

Using BrainVision recorder software (Brain Products, München, Germany) we recorded EEG data from 64 Ag/AgCl electrodes mounted in an electrode cap (ECI Inc.), referenced against Cz at a sampling rate of 500 Hz. Electrodes below the eyes (IO1, IO2) and at the outer canthi (LO1, LO2) recorded vertical and horizontal ocular activity. We kept electrode impedance

below 5 kΩ and applied a 100 Hz low pass filter, a time constant of 10 s, and a 50 Hz notch filter. At the beginning of the session we recorded 20 trials each of prototypical eye movements (up, down, left, right) for offline ocular artifact correction.

EEG data were processed using custom Matlab (The MathWorks Inc.) scripts [99] and EEGlab toolbox functions [100]. We re-calculated to average reference and retrieved the Cz channel. The data were band pass filtered between 0.5 and 40 Hz. Ocular artifacts were corrected using BESA [101]. We segmented the ongoing EEG from -200 to 800 ms relative to feedback onset. Segments containing artifacts were excluded from analyses, based on values exceeding ± 150 µV and gradients larger than 50 µV between two adjacent sampling points. Baselines were corrected to the 200 ms pre-stimulus interval (feedback onset).

The FRN was quantified in single-trial ERP waveforms as peak-to-peak amplitude at electrode FCz, specifically as the difference between the minimum voltage in a window from 200 to 300 ms post-feedback onset and the preceding positive maximum in a window from -100 to 0 ms relative to the detected negative peak. To define the time windows for single-trial analyses of P3a and P3b amplitudes, we first determined the average subject-wise peak latencies at FCz and Pz, respectively, and exported 100 ms time windows centered on the respective latencies. Accordingly, the P3a was quantified on single trials as the average voltage within an interval from 330 – 430 ms after feedback onset across all electrodes within a fronto-central region of interest (ROI: F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2). P3b amplitude was quantified in single trials as the average voltage within a 416 – 516 ms interval post-feedback across all electrodes within a parietally-focused region of interest (ROI: CP1, CPz, CP2, P1, Pz, P2, PO3, POz, PO4).

**Analyses**

Outlier inspection of the behavioral data identified one suspicious participant (average RT > 10 s) and one trial each in four additional participants (RTs > 6 s, 0.4% of data of remaining participants). These data were excluded from further analyses. We computed two kinds of prediction errors (cf. Fig 3A): SPE was determined as the absolute difference between predicted and actual interval length: |Prediction – Feedback|. RPE was computed as the difference between the absolute predicted error and the absolute actual error as revealed by feedback: |Prediction| – |Feedback|. We quantified confidence calibration as each participant's correlation of confidence and SPE (absolute deviation of the prediction from the actual outcome) across all trials, controlling for average error magnitude per block to account for shared changes in our confidence calibration measure with performance. To ease interpretation, we sign-reversed the correlations, such that higher values correspond to higher confidence calibration.

Statistical analyses were performed by means of linear mixed models (LMMs) using R [102] and the lme4 package [103]. We chose LMMs, similar to linear multiple regression models, as they allow for parametric analyses of single-trial measures. Further, LMMs are robust to unequally distributed numbers of observations across participants, and simultaneously estimate fixed effects and random variance between participants in both intercepts and slopes. For all dependent variables, full models, including all predictors, were reduced step-wise until model comparisons indicated significantly decreased fit.

We report model comparisons and fit indices: Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), which decrease with improving model fit. Random effect structures were determined using singular value decomposition. Variables explaining zero variance were removed from the random effects structure [104,105].

Prior to the analyses, error magnitude, RPE and SPE were scaled from ms to seconds and confidence and block were also scaled to a range of ± 1 for similar scaling of all predictors. Furthermore, block, error magnitude, confidence, and SPE were centered on their medians for accurate intercept computation. RPE was not centered, as zero represents a meaningful value on the scale (predicted and actual error magnitude are the same), and positive and negative values are qualitatively different (negative and positive values represent outcomes that are, respectively, worse or better than expected). Model formulas are reported in the respective tables. Fixed effects are visualized using the effects package [106].

**Data Availability**

The datasets generated and analyzed during the current study are available under https://github.com/froemero/Outcome-Predictions-and-Confidence-Regulate-Learning.

**Code Availability**

Scripts for all analyses are available under https://github.com/froemero/Outcome-Predictions-and-Confidence-Regulate-Learning.

**Acknowledgments**

We thank Lena Fliedner and Lara Montau for support in data acquisition and helpful discussions during the setup of the task, Rainer Kniesche for advice on programming the stimulus government, Markus Ullsperger, Adrian Haith, Martin Maier, and Rasha Abdel Rahman for valuable discussions, and Mehrdad Jazayeri for valuable feedback on a previous draft. RF is further grateful for the continuous scientific and personal support by her office mates at Humboldt-University, Benthe Kornrumpf and Florian Niefind, who made her life and work a lot more fun and happened to also have inspired the title of this paper.

**Author Contributions**

RF, NY, WS and BS conceived the study. RF collected the data. RF analyzed the data. MRN, RF, and RB conducted the computational modeling. The manuscript was written by RF, with contributions of MRN and RB. All authors discussed results and edited the manuscript.

**Competing Interests**

The authors declare no competing interests.

**References**

1 McGuire, J. T., Nassar, M. R., Gold, J. I. & Kable, J. W. Functionally dissociable influences on learning rate in a dynamic environment. *Neuron* **84**, 870-881, doi:10.1016/j.neuron.2014.10.013 (2014).

2 Yu, A. J. & Dayan, P. Uncertainty, neuromodulation, and attention. *Neuron* **46**, 681-692, doi:10.1016/j.neuron.2005.04.026 (2005).

3 Behrens, T. E., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. Learning the value of information in an uncertain world. *Nature neuroscience* **10**, 1214-1221, doi:10.1038/nn1954 (2007).

4 Diederen, K. M. & Schultz, W. Scaling prediction errors to reward variability benefits error-driven learning in humans. *Journal of neurophysiology* **114**, 1628-1640, doi:10.1152/jn.00483.2015 (2015).

5 Diederen, K. M., Spencer, T., Vestergaard, M. D., Fletcher, P. C. & Schultz, W. Adaptive Prediction Error Coding in the Human Midbrain and Striatum Facilitates Behavioral Adaptation and Learning Efficiency. *Neuron* **90**, 1127-1138, doi:10.1016/j.neuron.2016.04.019 (2016).

6 Pearce, J. M. & Hall, G. A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review* **87**, 532-552, doi:10.1037/0033-295x.87.6.532 (1980).

7 Faisal, A. A., Selen, L. P. & Wolpert, D. M. Noise in the nervous system. *Nature reviews. Neuroscience* **9**, 292-303, doi:10.1038/nrn2258 (2008).

8 Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction*. Vol. 1 (MIT press, 1998).

9 Wolpert, D. M., Diedrichsen, J. & Flanagan, J. R. Principles of sensorimotor learning. *Nature reviews. Neuroscience* **12**, 739-751, doi:10.1038/nrn3112 (2011).

10 McDougle, S. D. *et al.* Credit assignment in movement-dependent reinforcement learning. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 6797-6802, doi:10.1073/pnas.1523669113 (2016).

11 McDougle, S. D. *et al.* Neural Signatures of Prediction Errors in a Decision-Making Task Are Modulated by Action Execution Failures. *Current biology : CB* **29**, 1606-1613 e1605, doi:10.1016/j.cub.2019.04.011 (2019).

12 Parvin, D. E., McDougle, S. D., Taylor, J. A. & Ivry, R. B. Credit Assignment in a Motor Decision Making Task Is Influenced by Agency and Not Sensory Prediction Errors. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **38**, 4521-4530, doi:10.1523/JNEUROSCI.3601-17.2018 (2018).

13 Maier, M. E., Yeung, N. & Steinhauser, M. Error-related brain activity and adjustments of selective attention following errors. *Neuroimage* **56**, 2339-2347, doi:10.1016/j.neuroimage.2011.03.083 (2011).

14 Yeung, N., Botvinick, M. M. & Cohen, J. D. The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol Rev* **111**, 931-959, doi:10.1037/0033-295X.111.4.939 (2004).

15 Riesel, A., Weinberg, A., Endrass, T., Meyer, A. & Hajcak, G. The ERN is the ERN is the ERN? Convergent validity of error-related brain activity across different tasks. *Biol Psychol* **93**, 377-385, doi:10.1016/j.biopsycho.2013.04.007 (2013).

16 Maier, M. E., di Pellegrino, G. & Steinhauser, M. Enhanced error-related negativity on flanker errors: error expectancy or error significance? *Psychophysiology* **49**, 899-908, doi:10.1111/j.1469-8986.2012.01373.x (2012).

17 Murphy, P. R., Robertson, I. H., Harty, S. & O'Connell, R. G. Neural evidence accumulation persists after choice to inform metacognitive judgments. *Elife* **4**, doi:10.7554/eLife.11946 (2015).

18 Luft, C. D., Takase, E. & Bhattacharya, J. Processing graded feedback: electrophysiological correlates of learning from small and large errors. *Journal of cognitive neuroscience* **26**, 1180-1193, doi:10.1162/jocn_a_00543 (2014).

19 Ulrich, N. & Hewig, J. A miss is as good as a mile? Processing of near and full outcomes in a gambling paradigm. *Psychophysiology* **51**, 819-823, doi:10.1111/psyp.12232 (2014).

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

21 Arbel, Y. & Donchin, E. How large the sin? A study of the event related potentials elicited by errors of varying magnitude. *Psychophysiology* **48**, 1611-1620, doi:10.1111/j.1469-8986.2011.01264.x (2011).

22 Kononowicz, T. W., Roger, C. & van Wassenhove, V. Temporal Metacognition as the Decoding of Self-Generated Brain Dynamics. *Cerebral cortex* **29**, 4366-4380, doi:10.1093/cercor/bhy318 (2019).

23 Akdoğan, B. & Balcı, F. Are you early or late?: Temporal error monitoring. *Journal of Experimental Psychology: General* **146**, 347-361, doi:10.1037/xge0000265 (2017).

24 Kononowicz, T. W. & van Wassenhove, V. Evaluation of Self-generated Behavior: Untangling Metacognitive Readout and Error Detection. *Journal of cognitive neuroscience* **31**, 1641-1657, doi:10.1162/jocn_a_01442 (2019).

25 Wolpert, D. M. & Flanagan, J. R. Motor prediction. *Current biology : CB* **11**, R729-732 (2001).

26 Schiffer, A. M., Siletti, K., Waszak, F. & Yeung, N. Adaptive behaviour and feedback processing integrate experience and instruction in reinforcement learning. *Neuroimage* **146**, 626-641, doi:10.1016/j.neuroimage.2016.08.057 (2017).

27 Bland, A. R. & Schaefer, A. Different varieties of uncertainty in human decision-making. *Front Neurosci* **6**, 85, doi:10.3389/fnins.2012.00085 (2012).

28 Nassar, M. R., Wilson, R. C., Heasly, B. & Gold, J. I. An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **30**, 12366-12378, doi:10.1523/JNEUROSCI.0822-10.2010 (2010).

29 O'Reilly, J. X. Making predictions in a changing world-inference, uncertainty, and learning. *Front Neurosci* **7**, 105, doi:10.3389/fnins.2013.00105 (2013).

30 Vaghi, M. M. *et al.* Compulsivity Reveals a Novel Dissociation between Action and Confidence. *Neuron* **96**, 348-354 e344, doi:10.1016/j.neuron.2017.09.006 (2017).

31 Meyniel, F., Schlunegger, D. & Dehaene, S. The Sense of Confidence during Probabilistic Learning: A Normative Account. *PLoS Comput Biol* **11**, e1004305, doi:10.1371/journal.pcbi.1004305 (2015).

32 Pouget, A., Drugowitsch, J. & Kepecs, A. Confidence and certainty: distinct probabilistic quantities for different goals. *Nature neuroscience* **19**, 366-374, doi:10.1038/nn.4240 (2016).

33 Boldt, A. & Yeung, N. Shared neural markers of decision confidence and error detection. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **35**, 3478-3484, doi:10.1523/JNEUROSCI.0797-14.2015 (2015).

34 van den Berg, R. *et al.* A common mechanism underlies changes of mind about decisions and confidence. *Elife* **5**, e12192, doi:10.7554/eLife.12192 (2016).

35 Meyniel, F. & Dehaene, S. Brain networks for confidence weighting and hierarchical inference during probabilistic learning. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E3859-E3868, doi:10.1073/pnas.1615773114 (2017).

36 Desender, K., Boldt, A. & Yeung, N. Subjective Confidence Predicts Information Seeking in Decision Making. *Psychological science* **0**, 956797617744771, doi:10.1177/0956797617744771 (2018).

37 Boldt, A., Blundell, C. & De Martino, B. Confidence modulates exploration and exploitation in value-based learning. *Neurosci Conscious* **2019**, niz004, doi:10.1093/nc/niz004 (2019).

38 Sarafyazd, M. & Jazayeri, M. Hierarchical reasoning by neural circuits in the frontal cortex. *Science* **364**, eaav8911, doi:10.1126/science.aav8911 (2019).

39 Carlebach, N. & Yeung, N. Flexible use of confidence to guide advice requests. *PsyArXiv*, doi:10.31234/osf.io/ctyqp (2020).

40 Yeung, N. & Summerfield, C. Metacognition in human decision-making: confidence and error monitoring. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **367**, 1310-1321, doi:10.1098/rstb.2011.0416 (2012).

41 Ullsperger, M., Fischer, A. G., Nigbur, R. & Endrass, T. Neural mechanisms and temporal dynamics of performance monitoring. *Trends Cogn Sci* **18**, 259-267, doi:10.1016/j.tics.2014.02.009 (2014).

42 Balcı, F. *et al.* Optimal Temporal Risk Assessment. *Frontiers in Integrative Neuroscience* **5**, doi:10.3389/fnint.2011.00056 (2011).

43 Hadjiosif, A. M., Krakauer, J. W. & Haith, A. M. Did we get sensorimotor adaptation wrong? Implicit adaptation as direct policy updating rather than forward-model-based learning. *bioRxiv*, 2020.2001.2022.914473, doi:10.1101/2020.01.22.914473 (2020).

44 Holroyd, C. B. & Coles, M. G. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* **109**, 679-709, doi:10.1037/0033-295X.109.4.679 (2002).

45 Polich, J. Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* **118**, 2128-2148, doi:10.1016/j.clinph.2007.04.019 (2007).

46 Fischer, A. G. & Ullsperger, M. Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron* **79**, 1243-1255, doi:10.1016/j.neuron.2013.07.006 (2013).

47 Miltner, W. H., Braun, C. H. & Coles, M. G. 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, doi:10.1162/jocn.1997.9.6.788 (1997).

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

49 Holroyd, C. B., Hajcak, G. & Larsen, J. T. The good, the bad and the neutral: electrophysiological responses to feedback stimuli. *Brain Res* **1105**, 93-101, doi:10.1016/j.brainres.2005.12.015 (2006).

50 Walsh, M. M. & Anderson, J. R. Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neurosci Biobehav Rev* **36**, 1870-1884, doi:10.1016/j.neubiorev.2012.05.008 (2012).

51 Holroyd, C. B., Nieuwenhuis, S., Yeung, N. & Cohen, J. D. Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport* **14**, 2481-2484, doi:10.1097/01.wnr.0000099601.41403.a5 (2003).

52 Sambrook, T. D. & Goslin, J. A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological bulletin* **141**, 213-235, doi:10.1037/bul0000006 (2015).

53 Lohse, K. R., Miller, M. W., Daou, M., Valerius, W. & Jones, M. Dissociating the Contributions of Reward-Prediction Errors to Trial-Level Adaptation and Long-Term Learning. *Biol Psychol*, 107775, doi:10.1016/j.biopsycho.2019.107775 (2019).

54 Pfabigan, D. M. *et al.* Context-sensitivity of the feedback-related negativity for zero-value feedback outcomes. *Biol Psychol* **104**, 184-192, doi:10.1016/j.biopsycho.2014.12.007 (2015).

55 Sambrook, T. D. & Goslin, J. Mediofrontal event-related potentials in response to positive, negative and unsigned prediction errors. *Neuropsychologia* **61**, 1-10, doi:10.1016/j.neuropsychologia.2014.06.004 (2014).

56 Talmi, D., Atkinson, R. & El-Deredy, W. The feedback-related negativity signals salience prediction errors, not reward prediction errors. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **33**, 8264-8269, doi:10.1523/JNEUROSCI.5695-12.2013 (2013).

57 Nieuwenhuis, S., De Geus, E. J. & Aston-Jones, G. The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology* **48**, 162-175, doi:10.1111/j.1469-8986.2010.01057.x (2011).

58 Walentowska, W., Moors, A., Paul, K. & Pourtois, G. Goal relevance influences performance monitoring at the level of the FRN and P3 components. *Psychophysiology* **53**, 1020-1033, doi:10.1111/psyp.12651 (2016).

59 Frömer, R., Stürmer, B. & Sommer, W. (Don't) Mind the effort: Effects of contextual interference on ERP indicators of motor preparation. *Psychophysiology* **53**, 1577-1586, doi:10.1111/psyp.12703 (2016).

60 Butterfield, B. & Mangels, J. A. Neural correlates of error detection and correction in a semantic retrieval task. *Brain Res Cogn Brain Res* **17**, 793-817, doi:http://dx.doi.org/10.1016/S0926-6410(03)00203-9 (2003).

61 O'Reilly, J. X. *et al.* Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E3660-3669, doi:10.1073/pnas.1305373110 (2013).

62 Severo, M. C., Paul, K., Walentowska, W., Moors, A. & Pourtois, G. NEUROPHYSIOLOGICAL EVIDENCE FOR EVALUATIVE FEEDBACK PROCESSING DEPENDING ON GOAL RELEVANCE. *NeuroImage*, 116857, doi:https://doi.org/10.1016/j.neuroimage.2020.116857 (2020).

63 Yeung, N. & Sanfey, A. G. Independent coding of reward magnitude and valence in the human brain. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **24**, 6258-6264, doi:10.1523/JNEUROSCI.4537-03.2004 (2004).

64 Sailer, U., Fischmeister, F. P. & Bauer, H. Effects of learning on feedback-related brain potentials in a decision-making task. *Brain Res* **1342**, 85-93, doi:10.1016/j.brainres.2010.04.051 (2010).

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

66 Yeung, N., Holroyd, C. B. & Cohen, J. D. ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral cortex* **15**, 535-544, doi:10.1093/cercor/bhh153 (2005).

67 Fischer, A. G. & Ullsperger, M. When is the time for a change? Decomposing dynamic learning rates. *Neuron* **84**, 662-664, doi:10.1016/j.neuron.2014.10.050 (2014).

68 Pfabigan, D. M., Zeiler, M., Lamm, C. & Sailer, U. Blocked versus randomized presentation modes differentially modulate feedback-related negativity and P3b amplitudes. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* **125**, 715-726, doi:10.1016/j.clinph.2013.09.029 (2014).

69 Ernst, B. & Steinhauser, M. Effects of feedback reliability on feedback-related brain activity: A feedback valuation account. *Cogn Affect Behav Neurosci* **18**, 596-608, doi:10.3758/s13415-018-0591-7 (2018).

70 Ullsperger, M., Danielmeier, C. & Jocham, G. Neurophysiology of performance monitoring and adaptive behavior. *Physiol Rev* **94**, 35-79, doi:10.1152/physrev.00041.2012 (2014).

71 Bellebaum, C. & Colosio, M. From feedback- to response-based performance monitoring in active and observational learning. *Journal of cognitive neuroscience* **26**, 2111-2127, doi:10.1162/jocn_a_00612 (2014).

72 Bultena, S., Danielmeier, C., Bekkering, H. & Lemhöfer, K. Electrophysiological Correlates of Error Monitoring and Feedback Processing in Second Language Learning. *Frontiers in human neuroscience* **11**, 29-29, doi:10.3389/fnhum.2017.00029 (2017).

73 Kruger, J. & Dunning, D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol* **77**, 1121-1134 (1999).

74 Fleming, S. M. & Lau, H. C. How to measure metacognition. *Front Hum Neurosci* **8**, 443, doi:10.3389/fnhum.2014.00443 (2014).

75 Correa, C. M. C. *et al.* How the Level of Reward Awareness Changes the Computational and Electrophysiological Signatures of Reinforcement Learning. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **38**, 10338-10348, doi:10.1523/JNEUROSCI.0457-18.2018 (2018).

76 Butterfield, B. & Metcalfe, J. The correction of errors committed with high confidence. *Metacognition Learning* **1**, 69-84, doi:10.1007/s11409-006-6894-z (2006).

77 Butterfield, B. & Metcalfe, J. Errors committed with high confidence are hypercorrected. *Journal of experimental psychology. Learning, memory, and cognition* **27**, 1491-1494, doi:10.1037/0278-7393.27.6.1491 (2001).

78 Lebreton, M., Abitbol, R., Daunizeau, J. & Pessiglione, M. Automatic integration of confidence in the brain valuation signal. *Nature neuroscience* **18**, 1159-1167, doi:10.1038/nn.4064 (2015).

79 Nassar, M. R. *et al.* Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature neuroscience* **15**, 1040-1046, doi:10.1038/nn.3130 (2012).

80 Arbuzova, P. *et al.* Measuring Metacognition of Direct and Indirect Parameters of Voluntary Movement. *bioRxiv*, 2020.2005.2014.092189, doi:10.1101/2020.05.14.092189 (2020).

81 Charles, L. & Yeung, N. Dynamic sources of evidence supporting confidence judgments and error detection. *Journal of experimental psychology. Human perception and performance* **45**, 39-52, doi:10.1037/xhp0000583 (2019).

82 Desender, K., Boldt, A., Verguts, T. & Donner, T. H. Post-decisional sense of confidence shapes speed-accuracy tradeoff for subsequent choices. *bioRxiv*, 466730, doi:10.1101/466730 (2018).

83 Desender, K., Murphy, P., Boldt, A., Verguts, T. & Yeung, N. A post-decisional neural marker of confidence predicts information-seeking in decision-making. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 2620-2618, doi:10.1523/JNEUROSCI.2620-18.2019 (2019).

84 Nassar, M. R., Bruckner, R. & Frank, M. J. Statistical context dictates the relationship between feedback-related EEG signals and learning. *Elife* **8**, e46975, doi:10.7554/eLife.46975 (2019).

85 Colizoli, O., de Gee, J. W., Urai, A. E. & Donner, T. H. Task-evoked pupil responses reflect internal belief states. *Scientific Reports* **8**, 13702, doi:10.1038/s41598-018-31985-3 (2018).

86 Chiviacowsky, S. & Wulf, G. Feedback after good trials enhances learning. *Res Q Exerc Sport* **78**, 40-47, doi:10.1080/02701367.2007.10599402 (2007).

87 Chiviacowsky, S. & Wulf, G. Self-controlled feedback: Does it enhance learning because performers get feedback when they need it? *Research quarterly for exercise and sport* **73**, 408-415 (2002).

88 Chiviacowsky, S. & Wulf, G. Self-controlled feedback is effective if it is based on the learner's performance. *Res Q Exerc Sport* **76**, 42-48, doi:10.1080/02701367.2005.10599260 (2005).

89 Cockburn, J. & Holroyd, C. B. Feedback information and the reward positivity. *International Journal of Psychophysiology*, doi:https://doi.org/10.1016/j.ijpsycho.2017.11.017 (2017).

90 Langdon, A. J., Sharpe, M. J., Schoenbaum, G. & Niv, Y. Model-based predictions for dopamine. *Curr Opin Neurobiol* **49**, 1-7, doi:10.1016/j.conb.2017.10.006 (2018).

91 Dabney, W. *et al.* A distributional code for value in dopamine-based reinforcement learning. *Nature*, doi:10.1038/s41586-019-1924-6 (2020).

92 Wolpert, D. M. & Flanagan, J. R. Computations underlying sensorimotor learning. *Curr Opin Neurobiol* **37**, 7-11, doi:10.1016/j.conb.2015.12.003 (2016).

93 Holroyd, C. B. & Yeung, N. Motivation of extended behaviors by anterior cingulate cortex. *Trends Cogn Sci* **16**, 122-128, doi:10.1016/j.tics.2011.12.008 (2012).

94 Lieder, F., Shenhav, A., Musslick, S. & Griffiths, T. L. Rational metareasoning and the plasticity of cognitive control. *PLoS Comput Biol* **14**, e1006043, doi:10.1371/journal.pcbi.1006043 (2018).

95 Oldfield, R. C. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* **9**, 97-113 (1971).

96 Costa, P. T. & McCrae, R. R. *Revised NEO personality inventory (NEO-PI-R) and NEO Five-Factor inventory (NEO-FFI): Professional Manual*. (Psychological Assessment Resources, 1992).

97 Strobel, A., Beauducel, A., Debener, S. & Brocke, B. Eine deutschsprachige Version des BIS/BAS-Fragebogens von Carver und White. *Zeitschrift für Differentielle und Diagnostische Psychologie* **22**, 216-227, doi:10.1024//0170-1789.22.3.216 (2001).

98 Raven, J. The Raven's Progressive Matrices: Change and Stability over Culture and Time. *Cognitive Psychology* **41**, 1-48, doi:10.1006/cogp.1999.0735 (2000).

99 Frömer, R., Maier, M. & Abdel Rahman, R. Group-Level EEG-Processing Pipeline for Flexible Single Trial-Based Analyses Including Linear Mixed Models. *Frontiers in Neuroscience* **12**, doi:10.3389/fnins.2018.00048 (2018).

100 Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* **134**, 9-21, doi:10.1016/j.jneumeth.2003.10.009 (2004).

101 Ille, N., Berg, P. & Scherg, M. Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *Journal of Clinical Neurophysiology* **19**, 113-124, doi:10.1097/00004691-200203000-00002 (2002).

102 R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing (2014).

103 lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-8 (2014a).

104 Bates, D., Kliegl, R., Vasishth, S. & Baayen, H. Parsimonious mixed models. *arXiv preprint arXiv:1506.04967* (2015).

105 Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H. & Bates, D. Balancing Type I error and power in linear mixed models. *Journal of Memory and Language* **94**, 305-315, doi:10.1016/j.jml.2017.01.001 (2017).

106 Fox, J. & Weisberg, S. *An R companion to applied regression*. 3rd edn, (Sage, Thousand Oaks CA, 2019).