# EEG Study: SPN Modulations in a Probabilistic Reinforcement Learning Task

Emma Mandy Van Leckwyck

Master Thesis - Applied Cognitive Neuroscience

*S5431018*
*August 2024*
Department of Psychology
University of Groningen
Examiner/Daily supervisor:
Monicque Lorist and Marc Span

## Abstract

This study explores the modulation of Stimulus-Preceding Negativity (SPN) within a probabilistic reinforcement learning task. The SPN is an event-related potential (ERP) component associated with feedback anticipation, with previous research suggesting its modulation by both uncertainty and emotional valence. We aimed to investigate how SPN is influenced by the level of uncertainty regarding feedback and the emotional relevance of that feedback, particularly in gain versus loss conditions.

Using a probabilistic learning task, we manipulated the uncertainty of potential outcomes through different probability conditions (0.7, 0.5, 0.3) and emotional relevance by introducing gain-focused and loss-focused task conditions. The analysis revealed that SPN amplitudes were more negative in loss-focused conditions, indicating greater emotional impact, and that uncertainty influenced SPN amplitude, with more uncertain outcomes leading to higher SPN amplitudes.

The findings suggest that SPN may serve as an indicator of feedback relevance rather than simply reflecting prediction of outcomes, with emotional valence playing a significant role in its modulation. These results contribute to our understanding of the neural mechanisms underlying reinforcement learning, emphasizing the importance of both uncertainty and emotional context in shaping anticipatory neural responses.

# SPN modulation in a probabilistic reinforcement learning task

Recent advancements in cognitive neuroscience have unveiled various mechanisms underlying the brain's learning processes, particularly learning driven by reinforcement. Reinforcement learning (RL) is the process by which organisms learn, by trial and error, to predict and acquire reward or avoid punishment (Gershman & Daw, 2017). RL explores how neural circuits model the brain's reward system and direct learning by adapting behaviors to maximize rewards and minimize adverse outcomes (Niv, 2009). A practical application of RL principles can be seen in stimulus-association tasks with feedback, which involve several stages: stimulus presentation, response selection, and feedback delivery. During each stage, the brain's neural activities vary, encompassing decision-making, feedback anticipation, and feedback evaluation. This paper investigates the neural processes that occur between decision-making and feedback presentation, specifically focusing on stimulus-preceding negativity (SPN), an Event-Related Potential (ERP) observed during this interval (Chwilla & Brunia, 1991; Brunia & van Boxtel, 2001). ERPs, which are measured electrical brain responses associated with specific sensory, cognitive, or motor events, allow to study the temporal sequence of brain processes and provide key insights into how the brain anticipates and responds to feedback (Luck, 2014). In this context, the research explores the role of SPN in reinforcement learning and examines how it is modulated by different factors. Previous studies have identified the SPN as a slow negative component, right hemisphere dominant, appearing between - 800 ms and 0 ms before feedback (feedback being time 0) (Brunia, 1988; Chwilla & Brunia, 1991; Brunia & van Boxtel, 2001; Brunia et al., 2011; Li et al., 2017), and it is believed to reflect the anticipation of feedback outcomes (Kotani et al., 2017). Initially, the SPN was hypothesized to reflect motor preparation. However, during reaction time (RT) experiments, changes in reflex amplitude (i.e. tendon reflexes in the calf muscles of both legs) were observed during the preparatory period before a response (Brunia et al., 1985), however, when no motor response was required or if the response was delayed, the reflex amplitude stayed constant and the SPN still occurred. This suggests that motor preparation and stimulus anticipation are different processes, characterized by distinct physiological responses (Heart rate deceleration and different negative wave distributions) (Brunia, 1988).

Other researchers have proposed that the SPN serves as an ERP index of reward anticipation during reinforcement learning (Brunia et al., 2011). While the error-related negativity (ERN) and feedback-related negativity (FRN) have been well-studied as ERP components reflecting error detection and the comparison between anticipated and actual feedback outcomes (Holroyd & Coles, 2002), there has been no clear electrophysiological measure of feedback anticipation itself. Brunia et al. (2011) postulated that the SPN might be this anticipatory index, reflecting the expectation of action consequences. They did find a significant correlation between the SPN and FRN, suggesting that the SPN predicts the outcome, while the FRN computes the prediction error when the outcome deviates from expectations. This prediction error is likely processed within the dopaminergic system, where dopamine signals discrepancies in response to erroneous or unexpected outcomes (Holroyd & Coles, 2002). Confirmatively, findings from studies on dopaminergic disorders demonstrated that patients with mild Parkinson's disease showed a reduced SPN amplitude (Mattox et al., 2006), and those with severe Parkinson's disease showed no SPN at all (Brunia et al., 2011).

A contradictory line of evidence for the SPN to reflect a prediction, is that SPN amplitude diminishes as task learning progresses, indicating a negative correlation between SPN amplitude and the level of task learning (Morís et al., 2013). If the SPN is a prediction of the future reward, we would expect a gradual increase of SPN by learning where the more one learns the stronger the prediction will be.

Catena et al. (2012) instead proposed the uncertainty hypothesis, which suggests that the SPN is modulated by the level of uncertainty associated with feedback. As learning progresses and uncertainty about feedback outcome decreases, the SPN amplitude would correspondingly decrease. Their experiment confirmed that the SPN is systematically negatively larger in trials where the outcome was uncertain and unpredictable compared to those that were certain and predictable (Catena et al., 2012; Morís et al., 2013). Another study has found the SPN to be larger for unexpected rewards compared to highly predictable ones (Fuentemilla et al., 2013). Based on these findings, we hypothesize that SPN amplitude will be more negative for choices where the likelihood of receiving a reward is uncertain compared to those with certain outcomes.We further posit that the SPN reflects not only the degree of uncertainty associated with the response but also the subsequent relevance of feedback. After an uncertain response the SPN amplitude will be large if the feedback is highly relevant for learning. Conversely, SPN amplitude will be low if the following feedback has become redundant. Previous findings, supporting this hypothesis, have shown that SPN amplitude was greater for task informative feedback than for task non-informative feedback (Chwilla and Brunia, 1991). Chwilla and Brunia (1991) compared two feedback conditions: one with true/relevant information and the other with false/irrelevant information. Their results showed larger SPN amplitudes in the true feedback condition compared to the false feedback condition. The SPN may also be modulated by the emotional value of rewards. Some evidence shows that the unpredictability of a reward may enhance emotional arousing as it activates reward-related brain areas more robustly than predictable rewards (Fiorillo, Tobler, & Schultz, 2003). We hypothesize that uncertainty and feedback relevance are interconnected with emotional valence. On one hand, anticipation of uncertain outcomes often involves emotional arousal, heightening sensitivity to potential rewards and punishments (Sarinopoulos et al., 2009). On the other hand, the more emotionally arousing a situation is (i.e avoiding electrical shock or monetary loss) the more relevant the feedback becomes. In Chwilla and Brunia's study (1991), the SPN was observed before emotionally relevant feedback, but not following the anticipation of neutral stimulus (e.g., task instruction). Further findings expanded on the SPN being elicited prior to a variety of different emotionally relevant stimuli, such as evocative photos (Poli et al., 2007), aversive noise (Kotani et al., 2001), electrical shocks (Babiloni et al., 2007; Böcker et al., 2001) and monetary rewards (Chwilla and Brunia, 1991a; Kotani et al., 2001, 2003; Masaki et al., 2006, 2010; Ohgami et al., 2004, 2006).

The SPN's right hemisphere dominance suggests a natural link to the emotional valence of feedback. According to the valence hypothesis, the right hemisphere is predominantly involved in processing unpleasant emotions, while the left hemisphere is primarily responsible for processing pleasant emotions. (Ahern and Schwartz, 1979; Reuter-Lorenz and Davidson, 1981; Rodway et al., 2003; Gainotti, 2019). If the right hemisphere predominantly processes unpleasant emotions and the SPN shows right hemisphere dominance, this suggests that relevant negative emotional stimuli may play a significant role in modulating the SPN. Ohgami et al. (2006), indeed showed that during a punishment condition involving financial loss, the SPN exhibited greater amplitude in the right hemisphere (Ohgami et al., 2006). In contrast, during a reward condition where participants gained money, no significant differences in SPN amplitude between the hemispheres were observed. Kotani et al. (2009), hypothesized that the right hemisphere dominance of the SPN is counterbalanced by increased left hemisphere activation due to appetitive/pleasurable stimulation. The SPN could be more sensitive to costing reward than pleasant reward.

Further evidence supporting the role of emotional valence in SPN modulation comes from its generation in the anterior cingulate cortex (ACC) and insular cortex (van Boxtel & Böcker, 2004). Both the ACC and insular cortex are critical brain regions involved in integrating cognitive, emotional, and decision-making processes. They evaluate the emotional significance of stimuli, anticipate outcomes, and regulate responses based on both internal and external cues (Craig, 2009; Etkin, Büchel, & Gross, 2015). Kotani et al. (2015) provided additional insight through a functional magnetic resonance imaging (fMRI) source analysis combined with ERP data, revealing activation approximately 2800 ms before the feedback stimulus in the bilateral anterior insula and around 2600 ms before the feedback stimulus in the ACC. Notably, the left anterior insula showed no significant activation, consistent with the right hemisphere dominance of the SPN.

This right hemisphere preponderance is particularly relevant when considering the broader role of the anterior insula in emotional and bodily processes. Gu et al. (2013) hypothesized that the anterior insula is responsible for the process of interoceptive predictions. Interoceptive predictions are the brain's expectations about the body's internal state and bodily sensations, such as pain and pleasure (Critchley, 2005). The anterior insula is likely to send interoceptive predictions to the autonomic system and then receive prediction errors back from the autonomic system. This mechanism explains why the anterior insula is activated prior to events that elicit interoceptive responses. The anterior insula may transmit these interoceptive predictions in advance, consequently preparing the body for an appropriate response. Since the SPN occurs before a feedback stimulus and is generated in the anterior insula (van Boxtel and

Böcker, 2004), this indicates that the timing of the insula's activation is directly linked to the anticipation of feedback. The brain sends interoceptive predictions before the feedback occurs, suggesting that these predictions are related to affective properties. This relationship explains why the SPN is modulated by emotionally  relevant stimuli and why it is activated before the feedback stimulus.

Another potential explanation for emotional relevance is Bechara and Damasio's (2005) somatic marker hypothesis, suggesting that emotional processes guide (or bias) behavior and decision-making, particularly in uncertain situations. According to this hypothesis, somatic markers are bodily-related signals (like changes in heart rate or skin conductance) that arise in response to certain stimuli. These markers are stored in the brain and reactivated when a certain situation re-occurs, helping to avoid less optimal decisions in the future. This ability to re-activate emotional markers during decision-making is particularly critical, as demonstrated by patients with right hemisphere lesions who, in the absence of such markers, repeatedly made suboptimal choices in gambling and daily life situations (Dunn et al., 2006). Kotani et al. (2009), added evidence for these explanations. Via an fMRI measured task, they hypothesized that increased activation would be observed in the right insula under conditions of greater task difficulty. The results confirmed that the right anterior insula was more activated during more challenging “painful” tasks, supporting the link between the cortical area underlying SPN activity, its right hemisphere dominance, and the somatic marker hypothesis.

This research aims to investigate the factors that modulate SPN amplitude and its role in reinforcement learning. The first factor to be examined is uncertainty, which will be investigated through choices between stimuli, each yielding three different probabilities of rewards and punishments. This approach will create conditions of high and low uncertainty, as well as one uninformative condition. The second factor is emotional relevance, which will be explored through the effects of positive and negative feedback within a simulated monetary gain and loss paradigm. Furthermore, the study will explore the interaction between these two factors. Ultimately, this research seeks to provide deeper insight into the role of SPN in the reinforcement learning process.

## Methods

### Participants

Forty-four healthy adults participated in this study, out of which 10 were excluded due to bad EEG data, and 10 were excluded due to missing relevant data columns in the recorded task data. This resulted in twenty-four adults in the final sample (Mean age = 21.83, SD = 2.14) of which 17 were female, 6 were male, and 1 did not specify (other). Thirteen of the participants were acquaintances of the experimenters. The rest were students in their first year of the Psychology program who received university credits in exchange via the SONA system. All participants received confections for their participation. Most participants reported normal or corrected-to-normal vision. Participants gave written consent before the experiment. This study was exempt from ethical review based on a set of criteria outlined by the Ethics Committee of Psychology at the University of Groningen (code: PSY-2324-S-0298).

### Apparatus

The experiment was performed in a sound-attenuated room with a light intensity of 450 Lux. The task was programmed and implemented using OpenSesame software version 4.0.24 (Mathôt et al., 2012). The stimuli were presented at eye level, on a 27-inch 144 Hertz monitor placed approximately 65 cm from the participants' eyes. Responses were documented with a response box with three buttons. EEG was recorded using a Waveguard original 64-channel EEG cap with sintered Ag/AgCl electrodes placed in a 10-20 layout.

## Experimental Design

A probabilistic reward learning task was designed by adapting aspects of the spatial probabilistic reward learning task by Pütz et al. (2022) and the probabilistic cognitive reinforcement learning task by Frank et al. (2005). The main aim for the participants was to accumulate as many points (in the form of a made-up currency called "Uro") as possible, which they could trade in for confections after the experiment.

### Procedure

After the experiment information was read and understood, participants provided their written consent. Following this, EEG caps were applied. A short verbal introduction was given before more detailed task information was displayed on the screen. They were instructed to focus their gaze on the fixation dot and limit movements. Participants then performed the probabilistic-reward learning task in which they had to select a shape associated with a specific probability of gaining or losing points (0.7, 0.5, or 0.3), giving them a 70%, 50%, or 30% chance of success (i.e. gaining or not losing points (see below)). All participants underwent trials in two different task conditions in a randomized order of presentation. In the *gain-focused* task condition, participants started with zero points which they could *gain* or maintain and had the opportunity to earn points or maintain the ones they had acquired. This setup emphasizes the possibility of gaining rewards, rather than losing them.. In the *loss-focused* task condition, participants started with 210 points which they could maintain or *lose*. This setup emphasizes the possibility of losing rewards. Each task condition comprised five blocks.

The blocks were structured into two phases each: a *learning* and a *testing* phase. In the learning phase, participants were tasked with identifying the associated reward probabilities of three shapes through trial-and-error decision-making. This comprised 24 trials to ensure that all six spatial combinations of the shapes were shown four times. For each choice they received feedback adequate to the right condition. In the testing phase, participants were tasked with choosing any two out of the three shapes presented at once, constituting 24 trials to ensure that each combination was shown six times.

With a total of 48 trials per block, each participant completed 240 trials per condition, resulting in 480 trials per participant. Following a practice round of 24 trials, data collection commenced. Between blocks, participants were allowed to take a brief self-timed break. Halfway through the experiment, there was a two-minute break, after which the task conditions switched. Overall, the task took around 35 minutes to complete. Participants were informed of their total Uro coins and received a corresponding amount of confections.

### Probabilistic Reward-Learning Task

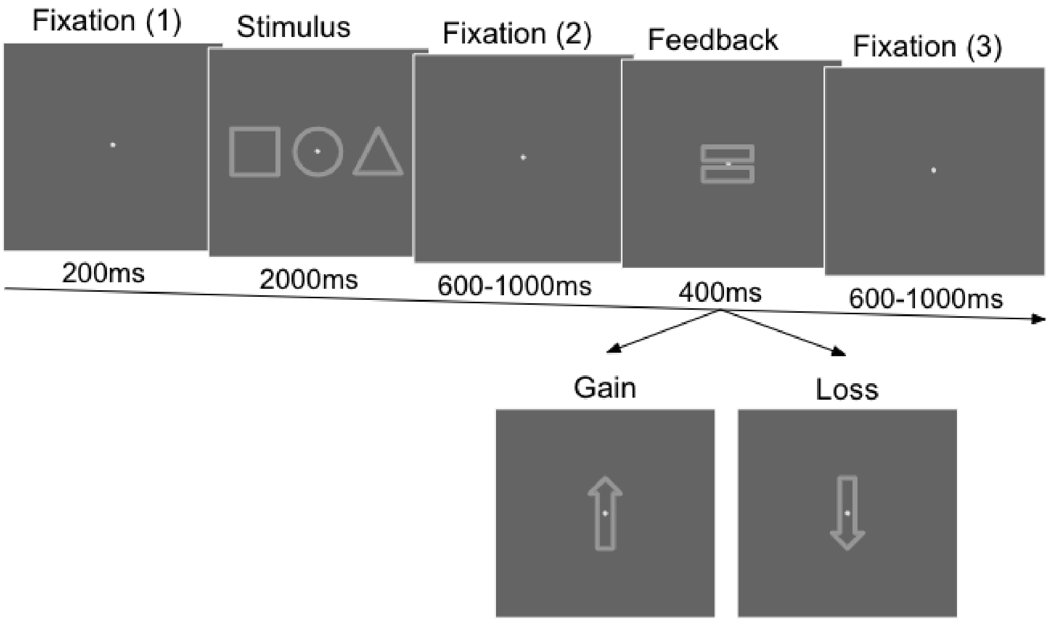
In the learning phase (Figure 1), after a ~200ms fixation screen, participants were presented with three stimuli at once (outlines of a circle, square, and triangle), which were randomized to predetermined locations. Each of the shapes was associated with a different reward probability (0.7, 0.5, or 0.3), randomly assigned to the shapes at the beginning of each block. The shape sizes were as follows: the triangle was 433x328 pixels, the circle was 391x325 pixels, and the square was 357x355 pixels, with a 145-pixel distance between the shapes. Participants had 2000 ms to respond, after which a new fixation screen appeared for 600-1000 ms, followed by feedback for 400 ms and another fixation screen of 600-1000 ms. Feedback depended on the condition participants were in. More specifically, an upward-facing arrow indicated a gain of one

Uro coin and a downward-facing arrow indicated the loss of one Uro Coin. Notably, while the equal sign referred to maintaining one's points in both task conditions, it conveyed an undesirable outcome in the gain-focused condition and a desirable outcome in the loss-focused condition. If participants did not respond to the stimuli on time, a screen showing 'No response!' was shown, and undesirable feedback was given.

At the end of this phase, participants were presented with the total points earned or lost during the phase and the total point score of the block.

**Figure 1**

*Trial Sequence of the Learning Phase With Duration of Each Screen in Milliseconds (ms)*



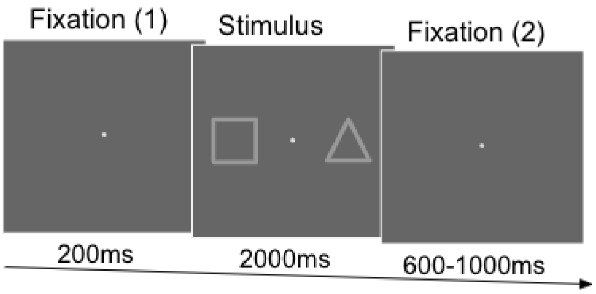
*Note.* For illustrational purposes, stimuli are presented larger than in the actual experiment.

The testing phase (Figure 2) followed a similar pattern. However, after viewing the fixation screen, participants were presented with any two of the three stimuli at once (placed at a 290-pixel distance from each other). The shapes retained their associated reward probabilities, but no feedback was presented. The test phase was used to examine the previously learned reward probabilities. After participants responded, a fixation dot with variable timing between 600 and 1000 ms was shown.

After the testing phase, the participants were presented with the total points gained or lost during the phase and the total point score of the block.

**Figure 2**

*Trial Sequence of the Testing Phase With Duration of Each Screen in Milliseconds (ms)*



*Note.* For illustrational purposes, stimuli are presented bigger than in the actual experiment.

## EEG Recording and Pre-Processing

EEG was recorded with a sampling rate of 512Hz, and the AFz channel was used as the ground. The hardware reference was CPz and data was offline re-referenced to the average of the mastoids: (M1 + M2)/ 2. In three cases one mastoid was removed due to noise or drifts and the data was re-reference to the remaining mastoid. Using a photodiode, we verified that stimulus onset on screen matched the event code timing in the recorded EEG data set.

EEG pre-processing was performed in MATLAB with EEGLAB (Delorme & Makeig, 2004) and the ERPlab (Lopez-Calderon & Luck, 2014) software toolboxes.

Data was filtered with a 0.1-30Hz half-amplitude and 12 dB/oct roll off bandpass filter. Seven datasets with high-frequency noise in the 50Hz range due to the amplifier being charged during recording or possibly line noise were notch filtered with a bandwidth of 49.5-50.5Hz. Independent component analysis (ICA) was performed with EEGLAB (fastICA algorithm; Hyvärinen & Oja, 2000) to remove eyeblinks, horizontal eye movements, and occasional drift component (a maximum of three components per subject) and reconstruct EEG. Epochs with feedback-locked ERPs exceeding a 50 mV amplitude threshold were excluded, and subjects with more than 25% of these rejected trials were removed. Epochs were extracted -800 ms to 0 ms before feedback stimulus onset. The baseline was the average voltage between -800 to -600 ms before feedback stimulus onset.

### *ERP analysis and data extraction*

The region of interest (ROI) for the SPN were the parietal electrodes Pz, P3, P4 based on previous related research (Catena et al., 2012; Kotani et al., 2015 ). A time window of interest (TOI) was defined with an interval between -400 to 0 ms pre-feedback.

## Quantification and Statistical Analysis

### *Trial Exclusion and Model Selection*

A mixed-modeling strategy was employed for all behavioral and neural data analysis using Linear Mixed Effects Models in R software ( Version 2024.04.2+764). Each model incorporated “subject” as a random factor to handle within-subject variability. Descriptions of the models used in the analyses are detailed in the sections below. We considered results statistically significant at $p < 0.05$.

### *Behavioural Analysis*

To investigate the learning trajectory, we examined how accuracy changes over trial number per task condition during the learning phase. Accuracy is defined as the proportion of choosing the “set winner” (0.7) versus “non-set winner” (0.3 & 0.5). We fitted a mixed effects model with the outcome variable “set winner” versus “set runner-up” (0.5) and “set loser” (0.3) as the dependent variable, and “task condition” and  “trial number” as the fixed effect.

Accuracy per condition during the testing phase was analyzed by modeling the probability of choosing the set winner by trial number. The outcome dependent variable was “set winner” versus “non-set winner” (0.3 & 0.5), and fixed effects included “task condition” and “trial number.

Additionally, to examine how speed varied over the course of the trials, we conducted a one-way ANOVA comparing reaction times between the first 12 trials (early phase) and the last 12 trials (late phase) of the learning phase. .

### *EEG analysis*

To analyze the modulation of SPN amplitude by probabilistic choice and task condition in the learning phase, a mixed-effects model (LME) was employed. This model is particularly suitable for handling variability between subjects and examining changes in SPN amplitude across trials and time. Participants were treated as a random effect to account for individual differences and intra-subject variability.

## Results

### Behavior

#### *Proportion of set chosen over trials*

In the learning phase (figure 1), as expected we see a positive learning curve for the set winner across trials. With increasing trial numbers the set winner (0.7) was chosen more often than the sets runner-up (0.5) and loser (0.3) (linear mixed model, $\beta = 0.433$, SE = 0.02045, $t(4104) = 21.208$, $p < 2e\text{-}16$). Non-significant differences were seen between the gain-focused and loss-focused task conditions ($\beta = -0.0083$, SE = 0.01576,, $t = -0.529$, $p = 0.597$).

**Figure 1**

*Grand Average of chosen variables by trials and task condition*

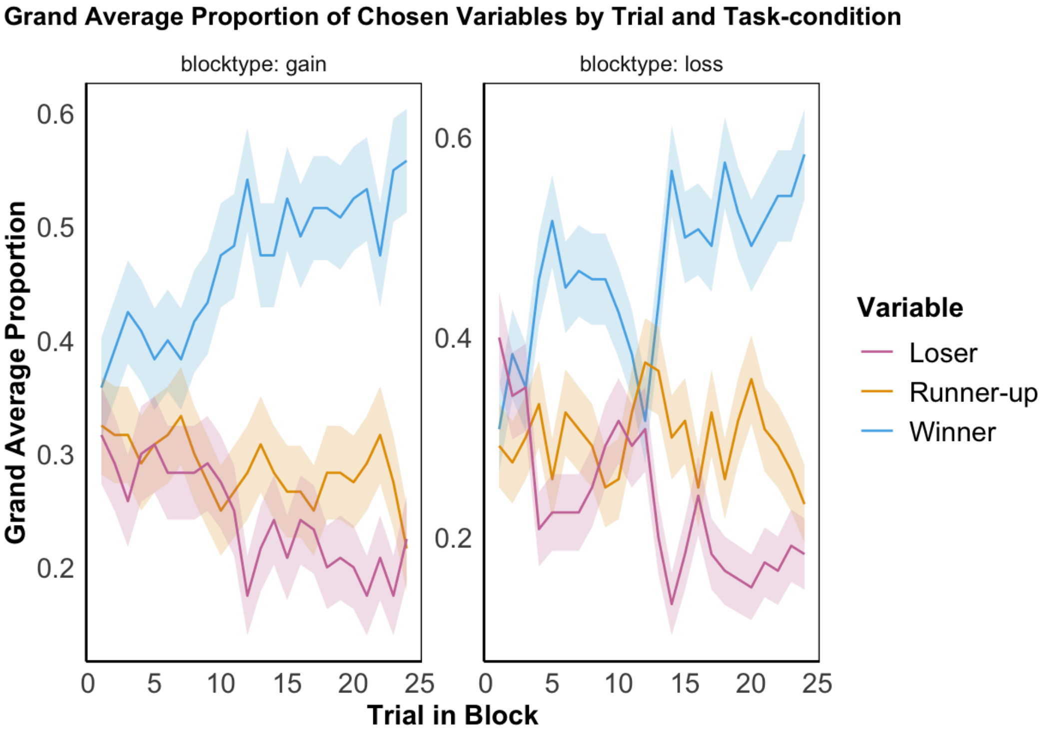


***Figure 1****: Grand Average Proportion of Chosen Variables by Trial and task condition during the learning phase. Set winner (0.7) was chosen more often than the two other sets showing a positive learning curve. There is no significant difference between the two task conditions.*

### *Reaction time in Early vs. Late Trials*

In the learning phase, the effect of trial stage (early vs. late) on reaction time was analyzed using a one-way analysis of variance (ANOVA) (Figure 2). The mean reaction time in the early trials was 638 ms, which decreased to 614 ms in the late trials. The ANOVA indicated a significant main effect of the trials phase on reaction time, ( F(1,5758)=9.565, p=0.00199). The sum of squares associated with trials phase (817,078) indicates that the trials phase accounted for a significant portion of the total variance in reaction times when compared to the residual sum of squares.

**Figure 2**

*Reaction Time early vs. late trials*

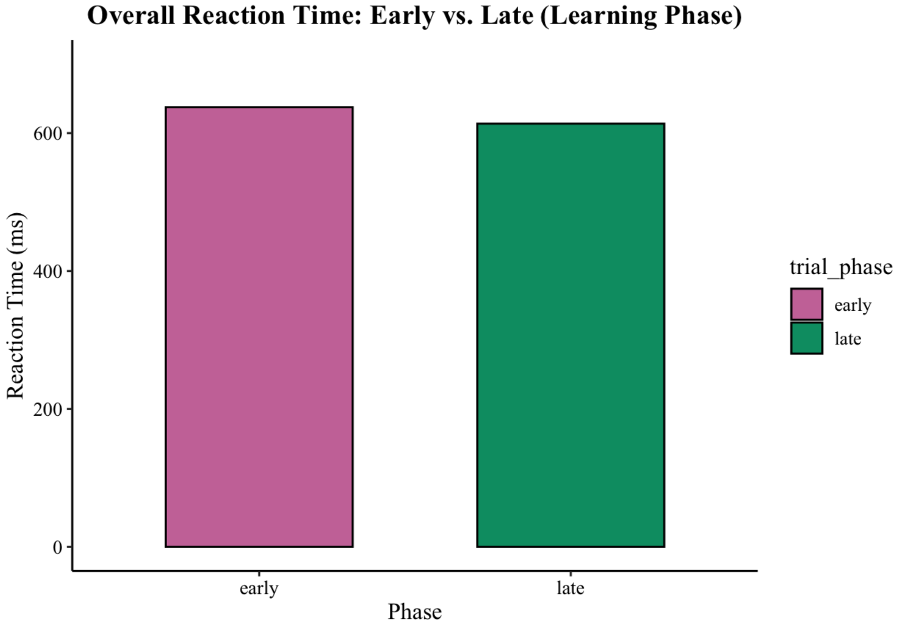


***Figure 2.*** *Mean reaction times (ms) during the early and late stages of the learning phase. The average reaction time during the early trials (1-12) was higher than in the later trials (13-24), indicating a decrease in reaction time as the trials progressed.*

## EEG ANALYSIS

The mixed-effects models evaluated the fixed effects of Task Condition (Gain vs. Loss) and Probability (Winner (0.7), Runner-up (0.5), Loser(0.3)) on the response variable SPN (amplitude), with Subject included as a random effect.

### Main effect of task condition

The analysis revealed a significant main effect of task condition on SPN, indicating that the type of task condition (Gain-focused vs. Loss-focused) had a substantial impact on SPN amplitude. Specifically, the Loss-focused task condition was associated with significantly more negative SPN values compared to the Gain-focused task condition, as evidenced by a Wald chi-square test,( $\chi2(1)=13.15$, $p=0.0003$) and a pairwise comparison (estimate = 2.08, SE = 0.375, $t(115)=5.558$ $p<0.0001$). The effect size for this main effect was moderate, with a partial eta-squared value of 0.21, suggesting that 21% of the variance in SPN can be attributed to the differences between the Gain-focused and Loss-focused conditions.

**Figure 3**

*Effect of task condition on SPN*

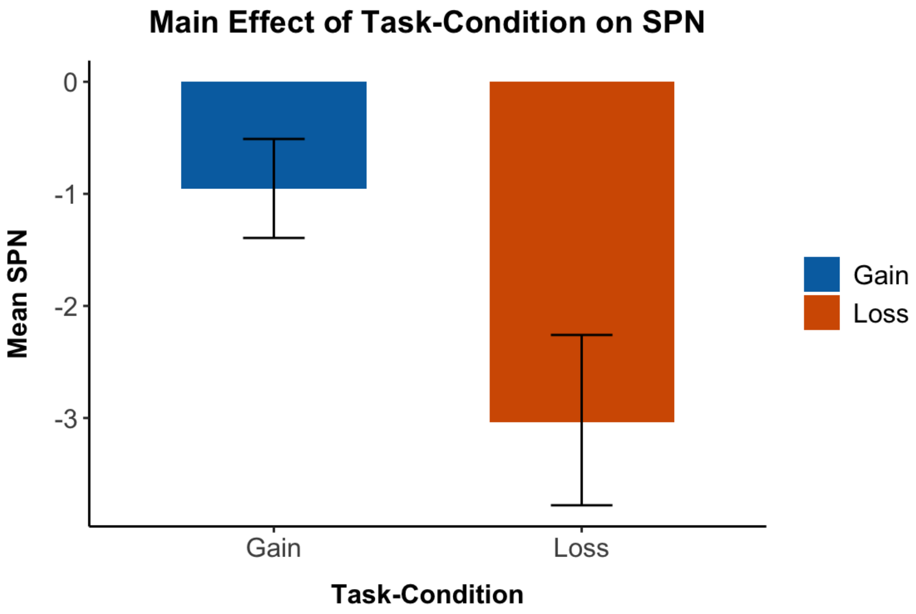


***Figure 3.*** *Mean SPN (Stimulus-preceding-negativity) as a function of task condition (Gain-focused vs. Loss-focused). The loss-focused task condition showed more negative SPN amplitude compared to the loss-focused task condition indicating a greater negative response in the loss-focused task condition. Error bars represent the standard error of the mean (SEM).*

**Figure 4**

### ***Scalp maps in the gain-focused and loss-focused task conditions***

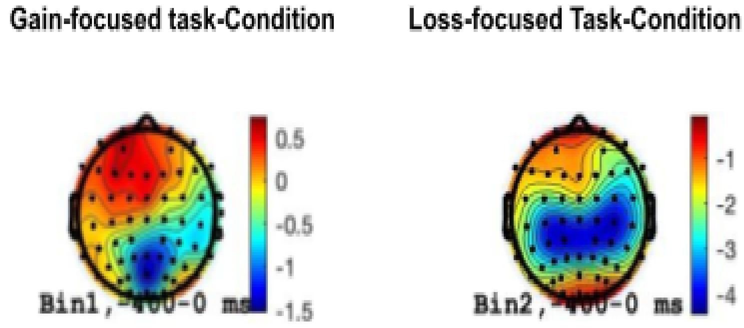


***Figure 4.*** *The scalp maps illustrate the topographical distribution of brain activity during the Gain-focused and Loss-focused task conditions, as measured by electroencephalography (EEG). The maps show the distribution of event-related potentials (ERPs) across the scalp, with red areas indicating higher positive values and blue areas indicating lower (or more negative) values. The topographies are shown for the Region of interest (ROI) between -400ms to 0ms. The loss-focused task condition visually shows a higher proportion of negative activity. Differences in the distribution patterns between the Gain-focused and Loss-focused task conditions suggest variations in neural processing depending on the context of potential gains or losses.*

## Main effect of Probability

The Probability factor, which categorized trials into Winner (0.7), Runner-up (0.5), and Loser (0.3), also showed a significant main effect on SPN, ($\chi2(2)=6.14$, $p=0.0464$). The effect size, though smaller than that of task condition, was still notable with a partial eta-squared value of 0.04, indicating that 4% of the variance in SPN is explained by the probability chosen. Further investigation of the different pairwise comparisons were conducted on SPN amplitude across the probabilities choices—Winner (0.7), Runner-up (0.5), and Loser (0.3) (Figure 5). The comparison between the Winner (0.7) and Runner-up (0.5) revealed a non-significant difference in SPN amplitude (estimate = -0.469, $t(115)=-1.022$ $p=0.5645$), indicating that the SPN was not substantially different between these two levels. Similarly, the comparison between the Winner (0.7) and Loser (0.3) choices showed no significant difference in SPN amplitude (estimate = 0.546, t(115)=1.190 p=0.4618). However, there was a trend observed in the comparison between the Runner-up (0.5) and Loser (0.3) choices, where the SPN amplitude was more negative for the Loser (0.3) and more positive for the Runner-up (0.5) (estimate = 1.015, t(115)=2.212 p=0.0735), though this difference did not reach statistical significance. These results suggest that while there are no significant differences in SPN amplitude between the Winner (0.7) and the other probability levels, there may be a tendency for the SPN to be more negative on the lowest probability, the Loser (0.3) compared to the Runner-up (0.5).

**Figure 5**

*Effect of probability on SPN amplitude*

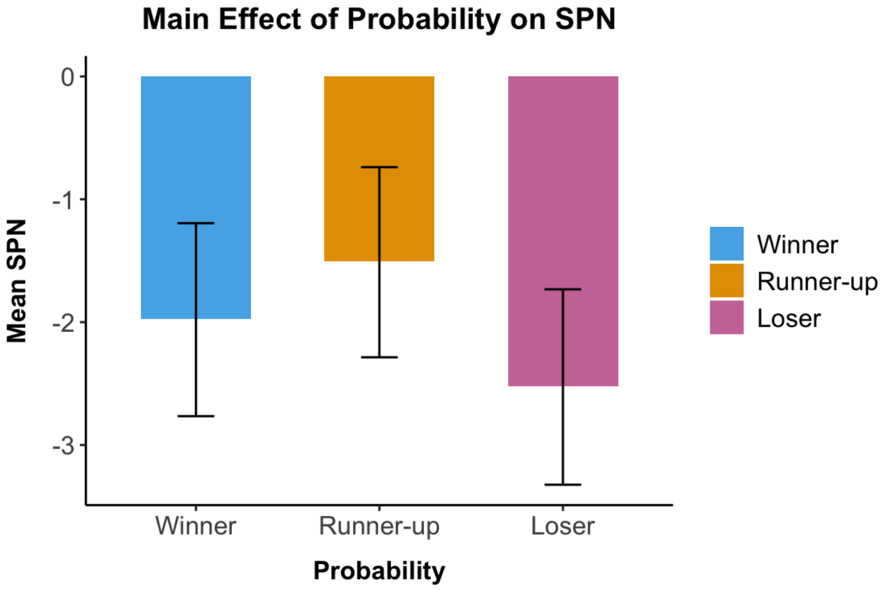


***Figure 5.*** *Effect of probabilities sets (Winner (0.7), Runner-up(0.5), Loser(0.3)) on Mean SPN amplitude. The SPN amplitude was most negative for the Loser (0.3) set. The SPN amplitude was*

*the least negative for the Runner-up set (0.5), indicating the weakest neural response, followed by the Winner set (0.7). Error bars represent the standard error of the mean (SEM).*

**Figure 6**

*SPN during Task- Conditions and Probabilities set*

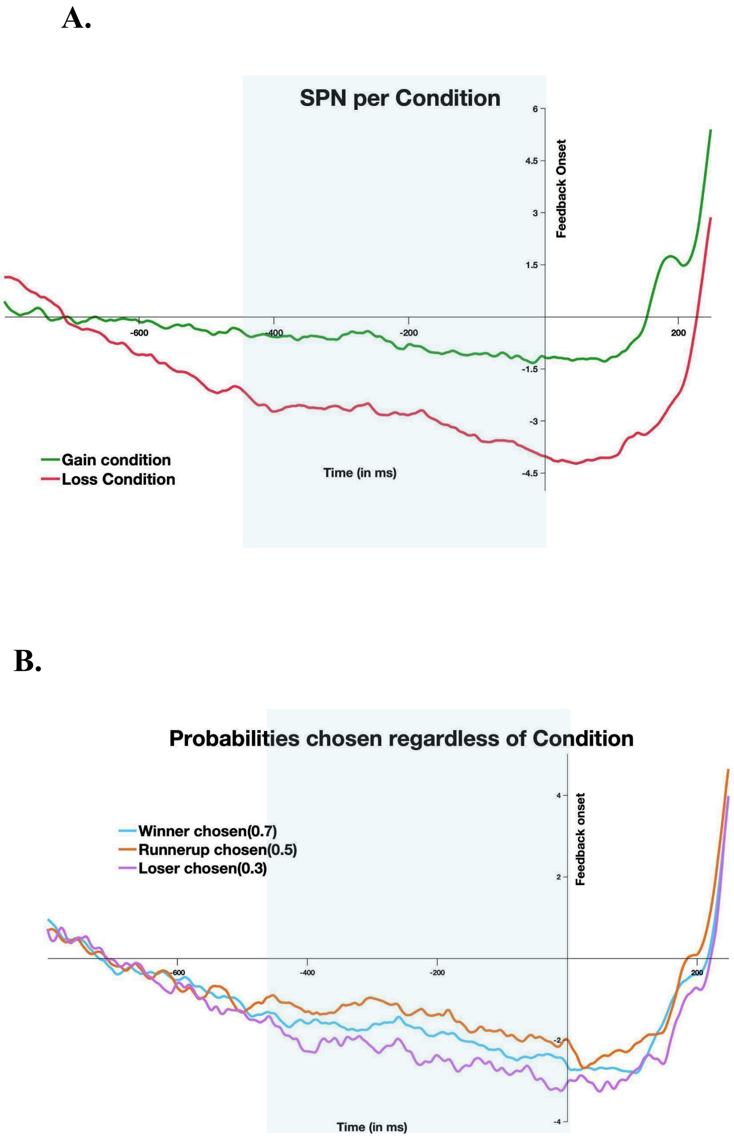


***Figure 5. (A)*** *Effect of task condition (Gain-focused vs. Loss-focused) on SPN amplitude. The green line represents the SPN for the Gain-focused task condition, and the red line represents the SPN for the Loss-focused task condition. Time is plotted in milliseconds (ms) relative to feedback*

*onset at 0 ms. The data show that SPN amplitude is more negative in the Loss-focused task condition compared to the Gain-focused task condition, indicating a stronger neural response during loss-related feedback anticipation.*

***(B)*** *Effect of probability set on SPN amplitude. The blue line represents the SPN when the highest probability (0.7) was chosen, the orange line represents the SPN when the middle probability (0.5) was chosen, and the purple line represents the SPN when the lowest probability (0.3) was chosen. Time is plotted in milliseconds (ms) relative to feedback onset at 0 ms. The data indicate that SPN amplitude becomes more negative as the chosen probability decreases, with the lowest probability (0.3) showing visually the most negative SPN amplitude.*

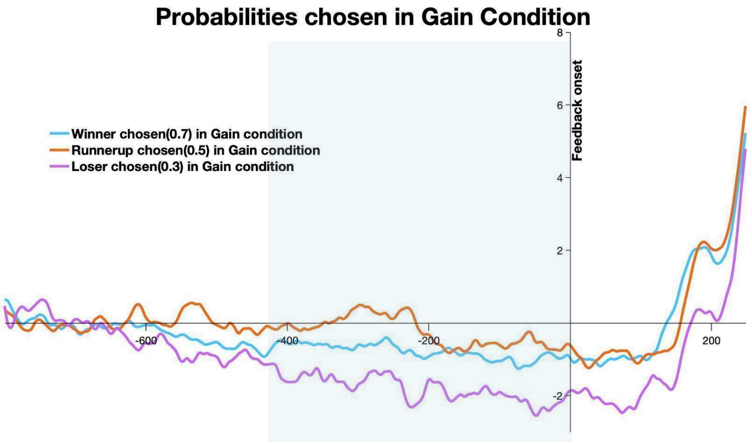
**Interaction Between task condition and Probability**

The interaction between Condition and Probability was not statistically significant (χ2(2)=1.80, p=0.4066), indicating that the effect of task condition on SPN does not vary significantly across different levels of Probability. However, upon visual inspection, the data suggest a different pattern. In the Gain-focused task condition, SPN amplitude appears to become more negative at both the highest and lowest probabilities, with the most negative SPN amplitude observed at the loser set (0.3). In contrast, in the Loss-focused task condition, SPN amplitude shows less variability depending on the chosen probability.

**Figure 7**

***SPN interaction of probability set and task condition***

**A.**



**B.**

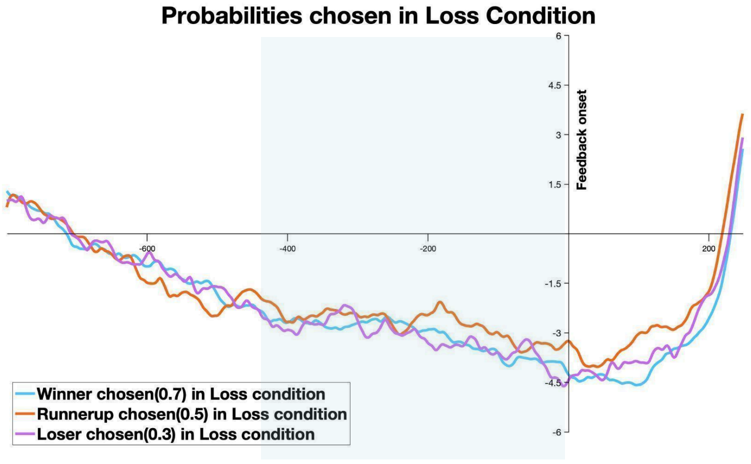


***Figure 7. (A)*** *Effect of chosen probabilities in the Gain-focused task condition on SPN amplitude. The blue line represents the SPN when the highest probability (0.7) was chosen, the orange line represents the SPN when the middle probability (0.5) was chosen, and the purple line represents the SPN when the lowest probability (0.3) was chosen. Time is plotted in milliseconds (ms) relative to feedback onset at 0 ms. The data show that, in the Gain-focused task condition, the SPN amplitude becomes more negative at the highest and lowest chosen probability with the lowest probability (0.3) showing the most negative SPN amplitude.*

***(B)*** *Effect of chosen probabilities in the Loss-focused task condition on SPN amplitude. The blue line represents the SPN when the highest probability (0.7) was chosen, the orange line represents the SPN when the middle probability (0.5) was chosen, and the purple line represents the SPN when the lowest probability (0.3) was chosen. Time is plotted in milliseconds (ms) relative to feedback onset at 0 ms. In the Loss-focused task condition, the data suggest that the SPN amplitude shows variability based on the chosen probability, with the lowest probability (0.3) seemingly showing the most negative SPN amplitude.*

## Exploratory analysis

### *Feedback types*

A repeated measures ANOVA was conducted to examine the effect of feedback types on SPN amplitude, recorded over the parietal region. The analysis revealed a significant main effect of feedback types on SPN, F(1.60, 36.82) = 12.26, $p < .001$, $\eta^2 = .188$, indicating that the SPN differed significantly across the conditions. However, visual analysis indicates that these results are largely driven by the task condition rather than other factors.

**Figure 8**

*Feedback types*

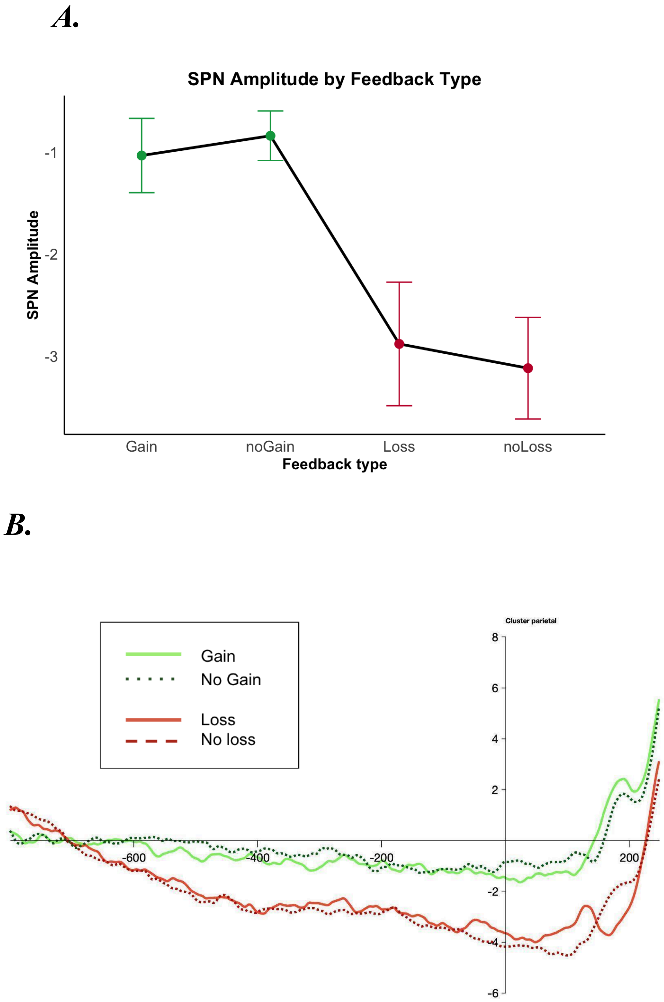


***Figure 8.*** *(A) Effect of SPN on feedback type. The green line represents conditions where feedback involves potential gains (Gain, noGain), while the red line represents conditions involving potential losses (Loss, noLoss). The SPN amplitude is plotted on the y-axis, with more negative values indicating stronger neural responses. Error bars represent the standard error of the mean (SEM). The data show a clear decrease in SPN amplitude from Gain to Loss feedback types, with the most negative SPN observed in the Loss condition.*

*(B)Effect of SPN on feedback types displayed for Gain, no Gain, Loss, and no Loss. The solid green line represents the SPN for Gain, while the dotted green line represents the SPN for no Gain. Similarly, the solid red line represents the SPN for Loss, and the dashed red line represents the SPN for no Loss. Time is plotted in milliseconds (ms) relative to feedback onset at 0 ms. The data illustrate distinct SPN patterns depending on the feedback type, with the Loss and no Loss conditions showing more negative SPN amplitudes compared to Gain and no Gain conditions.*

## Discussion

In a reinforcement learning tasks involving positive and negative feedback, the stimulus-preceding negativity (SPN), a slow cortical negativity that, appears pre-feedback is believed to reflect feedback anticipation (Brunia, 1988; Chwilla and Brunia, 1991; Brunia and van Boxtel, 2001; Brunia et al., 2011; Kotani et al., 2017). However, the modulation and function of SPN are still ambiguous.We aimed to clarify how the SPN is influenced by two key factors: emotional valence and the uncertainty of potential feedback. Considering these findings, we also discuss the potential role of SPN in reinforcement learning.

We manipulated two main factors in our task: choosing a shape associated with a probability set representing chance of success (targeting uncertainty) and gain/loss focused task conditions (targeting emotional valence).

Our data indicate a small effect of probabilistic choices on SPN amplitude. Visual inspection suggests that the lowest amplitude was observed for the loser (0.3) choice, followed by the winner (0.7) and then the runner-up (0.5) choices. In line with the uncertainty hypothesis, which posits that feedback uncertainty modulates the SPN (Catena et al., 2012), it seems that the choice with the highest uncertainty—the 30% probability of success—resulted in the most negative amplitude. Interestingly, the runner-up (0.5) choice appears to have the most positive amplitude. One might argue that the constant randomness of feedback, as it is always a 50% chance of success, made the feedback received always the same level of uncertainty and hence non-informative. This aligns with the result of Chwilla and Brunia (1991) where SPN amplitude was greater for task informative feedback than for task non-informative feedback.

These results are particularly intriguing when considering the learning curve of the best and worst choice. Our data shows a steep increase in the learning curve for the winner (0.7) choice, indicating that participants successfully learned the best option. In contrast, loser (0.3) choices did not show a distinct learning curve. This finding supports Morís et al. (2013), who suggested that SPN amplitude is contingent on learning, with SPN magnitude decreasing as learning progresses. In addition, we see that reaction time decreases in the 24 trials set of the learning phase. If Morís et al. (2013) and Catena et al., (2012) are right, this reflects learning and reduced uncertainty as one tends to be faster when more certain (McDougle & Collins, 2020).

However, uncertainty might be too restrictive in modulating the SPN. Potentially it is the broader context of feedback relevance thus how important the subsequent feedback is for the learning and not only the uncertainty of the response that plays a role (Morís et al., 2013;Walentowska et al., 2018) .

We also hypothesize that SPN is modulated by emotionally relevant feedback (task condition) (Chwilla and Brunia, 1991a; Böcker et al., 2001; Kotani et al., 2001, 2003; Ohgami et al., 2004; Poli et al., 2007; Masaki et al., 2006, 2010; Ohgami et al., 2006; Babiloni et al., 2007). We indeed found differences in SPN amplitude between the gain-focused and loss-focused task conditions. The SPN showing in the loss-focused condition was markedly more negative and increased distribution of negative activity on the topographic map. Our findings suggest that losses, in general, are more salient than gains, possibly due to loss aversion (Kahneman et al., 1990). Loss aversion refers to the tendency for people to prefer avoiding losses over acquiring equivalent gains, as they anticipate that losses will have a greater emotional impact (Kermer et al., 2006). This expectation, or affective forecast, aligns with the SPN's role in anticipating the emotional consequences of a potential loss. It’s possible that participants in the loss-focused task condition anticipated losses more intensely than they did gains in the gain-focused condition. One might suggest these results connect to the somatic marker hypothesis stating that emotional processes guide decision-making: the increased negativity of SPN amplitude in the loss-focused task condition may indicate the reactivation of emotional somatic markers (i.e changes in heart rate or skin conductance), which arise to guide decisions in minimizing negative outcomes.

Remarquable is that in the loss-focused task condition, SPN amplitudes appear more consistent across probabilities, suggesting that losses trigger a uniform effect, independent of probability. This may reflect the brain's consistent reinforcement of negative outcomes, aligning with the principle that reinforcing a specific outcome (in this case, loss) leads to faster and more stable learning (Staddon, 1995). In contrast, the gain-focused condition may lead to more variability in SPN amplitudes because the potential for positive reinforcement (gains) could be more directly linked to the probabilities in turn leading to a broader range of neural responses as participants adjust their expectations based on the likelihood of gaining.

In light of these findings, what role might the SPN play in reinforcement learning? Our exploratory analysis provided contradictory evidence to the hypothesis that the SPN serves as an index of reward anticipation reflecting the expectation of action consequences in relation to the FRN (Brunia et al., 2011). Specifically, it seems there was no difference in SPN amplitude between feedback types beyond the effect of task condition. Additionally, we observed an effect of learning and a moderate effect of uncertainty. This suggests that the SPN may not represent a prediction but rather serve as a signal for the necessity of feedback before it is received, indicating how far they are in the learning process. This line of reasoning could explain why the SPN amplitudes in the loss-focused condition are significantly more negative compared to the gain-focused condition. As losses are generally more emotionally impactful, all feedback in the loss-focused condition could be  deemed highly relevant for the learning process.

Future research should more clearly differentiate the concepts of relevance, uncertainty, and emotion in the context of SPN modulation. If SPN serves as an indicator of relevance, it would be beneficial to examine its role throughout the entire reinforcement learning process, particularly in comparison with Feedback-Related Negativity (FRN), which reflects prediction error. Notably, if feedback initially anticipated as relevant proves otherwise, the FRN response may differ from that observed when expectations are met

In conclusion, this research investigated factors modulating the SPN and discussed its role in the reinforcement learning process. Our findings indicate that the SPN is strongly influenced by emotional valence and, to a lesser extent, by uncertainty. Finally, we suggest that the SPN's role in reinforcement learning is more about signaling the importance of feedback than predicting it.

# References

Ahern, G. L., & Schwartz, G. E. (1979). Differential lateralization for positive versus negative emotion. *Neuropsychologia*, *17*(6), 693–698. https://doi.org/10.1016/0028-3932(79)90045-9

Babiloni, F., Astolfi, L., Cincotti, F., Mattia, D., Tocci, A., Tarantino, A., Marciani, M., Salinari, S., Gao, S., Colosimo, A., & De Vico Fallani, F. (2007). Cortical Activity and Connectivity of Human Brain during the Prisoner's Dilemma: an EEG Hyperscanning Study. *Conference Proceedings*. https://doi.org/10.1109/iembs.2007.4353452

Bach, D. R. (2012). Knowing how much you don't know: a neural organization of uncertainty estimates. *Nature Reviews Neuroscience*, *13(8)*, 572–586.

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, *52*(2), 336–372. https://doi.org/10.1016/j.geb.2004.06.010

Bell, A. J., & Sejnowski, T. J. (1995). An Information-Maximization approach to blind separation and blind Deconvolution. *Neural Computation*, *7*(6), 1129–1159. https://doi.org/10.1162/neco.1995.7.6.1129

Brunia, C. (1988). Movement and stimulus preceding negativity. *Biological Psychology*, *26*(1–3), 165–178. https://doi.org/10.1016/0301-0511(88)90018-x

Brunia, C. H., Hackley, S. A., Van Boxtel, G. J., Kotani, Y., & Ohgami, Y. (2011). Waiting to perceive: Reward or punishment? *Clinical Neurophysiology*, *122*(5), 858–868. https://doi.org/10.1016/j.clinph.2010.12.039

Catena, A., Perales, J. C., Megías, A., Cándido, A., Jara, E., & Maldonado, A. (2012). The brain network of expectancy and uncertainty processing. *PloS One*, *7*(7), e40252. https://doi.org/10.1371/journal.pone.0040252

Chwilla, D. J., & Brunia, C. H. (1991). Event‐Related potentials to different feedback stimuli. *Psychophysiology*, *28*(2), 123–132. https://doi.org/10.1111/j.1469-8986.1991.tb00400.x

Craig, A. D. B. (2009). How do you feel — now? The anterior insula and human awareness. *Nature Reviews. Neuroscience*, *10*(1), 59–70. https://doi.org/10.1038/nrn2555

Critchley, H. D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. *Journal of Comparative Neurology*, *493*(1), 154–166. https://doi.org/10.1002/cne.20749

Damen, E. J. P., & Brunia, C. H. M. (1994). Is a stimulus conveying task‐relevant information a sufficient condition to elicit a stimulus‐preceding negativity? *Psychophysiology*, *31*(2), 129–139. https://doi.org/10.1111/j.1469-8986.1994.tb01033.x

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Etkin, A., Büchel, C., & Gross, J. J. (2015). The neural bases of emotion regulation. *Nature Reviews. Neuroscience*, *16*(11), 693–700. https://doi.org/10.1038/nrn4044

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, *299*(5614), 1898–1902. https://doi.org/10.1126/science.1077349

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-Related negativity predicts reinforcement learning and conflict biases. *Neuron*, *47*(4), 495–501. https://doi.org/10.1016/j.neuron.2005.06.020

Fryer, S., Roach, B. J., Holroyd, C. B., Paulus, M. P., Sargent, K., Boos, A., Ford, J. M., & Mathalon, D. H. (2021). Electrophysiological investigation of reward anticipation and outcome evaluation during slot machine play. *NeuroImage*, *232*, 117874. https://doi.org/10.1016/j.neuroimage.2021.117874

Fuentemilla, L., Cucurell, D., Marco-Pallarés, J., Guitart-Masip, M., Morís, J., & Rodríguez-Fornells, A. (2013). Electrophysiological correlates of anticipating improbable but desired events. *NeuroImage*, *78*, 135–144. https://doi.org/10.1016/j.neuroimage.2013.03.062

Gainotti, G. (2019). The role of the right hemisphere in emotional and Behavioral Disorders of Patients with frontotemporal lobar degeneration: an updated review. *Frontiers in Aging Neuroscience*, *11*. https://doi.org/10.3389/fnagi.2019.00055

Gershman, S. J., & Daw, N. D. (2017). Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annual Review of Psychology*, *68*(1), 101–128. https://doi.org/10.1146/annurev-psych-122414-033625

Gray, M. A., & Critchley, H. D. (2007). Interoceptive basis to craving. *Neuron*, *54*(2), 183–186. https://doi.org/10.1016/j.neuron.2007.03.024

Gray, M. A., Harrison, N. A., Wiens, S., & Critchley, H. D. (2007). Modulation of Emotional Appraisal by False Physiological Feedback during fMRI. *PloS One*, *2*(6), e546. https://doi.org/10.1371/journal.pone.0000546

Gu, X., Hof, P. R., Friston, K. J., & Fan, J. (2013). Anterior insular cortex and emotional awareness. *Journal of Comparative Neurology*, *521*(15), 3371–3388. https://doi.org/10.1002/cne.23368

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037/0033-295x.109.4.679

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Networks*, *13*(4–5), 411–430. https://doi.org/10.1016/s0893-6080(00)00026-5

Kermer, D. A., Driver-Linn, E., Wilson, T. D., & Gilbert, D. T. (2006). Loss aversion is an affective forecasting error. *Psychological Science*, *17*(8), 649–653. https://doi.org/10.1111/j.1467-9280.2006.01760.x

Kotani, Y., Hiraku, S., Suda, K., & Aihara, Y. (2001). Effect of positive and negative emotion on stimulus‐preceding negativity prior to feedback stimuli. *Psychophysiology*, *38*(6), 873–878. https://doi.org/10.1111/1469-8986.3860873

Kotani, Y., Ohgami, Y., Ishiwata, T., Arai, J., Kiryu, S., & Inoue, Y. (2015). Source analysis of stimulus-preceding negativity constrained by functional magnetic resonance imaging. *Biological Psychology*, *111*, 53–64. https://doi.org/10.1016/j.biopsycho.2015.08.005

Kotani, Y., Ohgami, Y., Kuramoto, Y., Tsukamoto, T., Inoue, Y., & Aihara, Y. (2009). The role of the right anterior insular cortex in the right hemisphere preponderance of stimulus-preceding negativity (SPN): An fMRI study. *Neuroscience Letters*, *450*(2), 75–79. https://doi.org/10.1016/j.neulet.2008.11.032

Kotani, Y., Ohgami, Y., Yoshida, N., Kiryu, S., & Inoue, Y. (2017). Anticipation process of the human brain measured by stimulus-preceding negativity (SPN). *Journal of Physical Fitness and Sports Medicine*, *6*(1), 7–14. https://doi.org/10.7600/jpfsm.6.7

Lai, E. R. (2011). *Motivation: A Literature Review*. https://d1wqtxts1xzle7.cloudfront.net/40756174/Motivation_Review_final-libre.pdf?1450092242=&response-content-disposition=inline%3B+filename%3DMotivation_A_literature_review.pdf&Expires=1709914992&Signature=Bf3u4RGGC7r91cCEseKWMnRKw2tgz~MIoA4KBKbLTmqJ89JRi4DBtyQb9on-LhfihrdQ8FS0nyQhSJYBTErcBVb9bUKjQNob~4p9fOH--~eNpzQCmEFpx2ptLew-qvqDamg5MGMsUhLWlBwRjGRJBotpMzOcO~1kl6qEdnQI~-4yV5zzFsRihs70et2r5Baoz1yYLjFx3GQYbGl-V81y605eastx~zY6YMYlboBRIbjKFGcSEjG93Xea8iitHeUgpOHwMZZL45gZe2SAIejhUFhSlw-vhUtFMFybmOTyN68HjJPEIL-s5F6pMQmZNs1F0tTppOB9setHA9jVHqkDyg__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA

Li, Q., Xu, J., & Zheng, Y. (2017). Stimulus-preceding negativity: An electrophysiological index of reward anticipation. *Xinli Kexue Jinzhan*, *25*(7), 1114. https://doi.org/10.3724/sp.j.1042.2017.01114

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00213

Masaki, H., Takeuchi, S., Gehring, W. J., Takasawa, N., & Yamazaki, K. (2006). Affective-motivational influences on feedback-related ERPs in a gambling task. *Brain Research*, *1105*(1), 110–121. https://doi.org/10.1016/j.brainres.2006.01.022

Masaki, H., Yamazaki, K., & Hackley, S. A. (2010). Stimulus-preceding negativity is modulated by action-outcome contingency. *Neuroreport*, *21*(4), 277–281. https://doi.org/10.1097/wnr.0b013e3283360bc3

Mathôt, S., Schreij, D., & Theeuwes, J. (2011). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

McDougle, S. D., & Collins, A. G. E. (2020). Modeling the influence of working memory, reinforcement, and action uncertainty on reaction time and choice during instrumental learning. *Psychonomic Bulletin & Review*, *28*(1), 20–39. https://doi.org/10.3758/s13423-020-01774-z

Morís, J., Luque, D., & Rodríguez‐Fornells, A. (2013). Learning‐induced modulations of the stimulus‐preceding negativity. *Psychophysiology*, *50*(9), 931–939. https://doi.org/10.1111/psyp.12073

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*, 139–154. https://doi.org/10.1016/j.jmp.2008.12.005

Ohgami, Y., Kotani, Y., Tsukamoto, T., Omura, K., Inoue, Y., Aihara, Y., & Nakayama, M. (2006). Effects of monetary reward and punishment on stimulus‐preceding negativity. *Psychophysiology*, *43*(3), 227–236. https://doi.org/10.1111/j.1469-8986.2006.00396.x

Pütz, C., Van Den Berg, B., & Lorist, M. M. (2022). Dynamic modulation of neural feedback processing and attention during spatial probabilistic learning. *iScience*, *25*(5), 104302. https://doi.org/10.1016/j.isci.2022.104302

Reuter-Lorenz, P., & Davidson, R. J. (1981). Differential contributions of the two cerebral hemispheres to the perception of happy and sad faces. *Neuropsychologia*, *19*(4), 609–613. https://doi.org/10.1016/0028-3932(81)90030-0

Rodway, P., Wright, L., & Hardie, S. (2003). The valence-specific laterality effect in free viewing conditions: The influence of sex, handedness, and response bias. *Brain and Cognition*, *53*(3), 452–463. https://doi.org/10.1016/s0278-2626(03)00217-3

Sarinopoulos, I., Grupe, D. W., Mackiewicz, K. L., Herrington, J. D., Lor, M., Steege, E. E., & Nitschke, J. B. (2009). Uncertainty during Anticipation Modulates Neural Responses to Aversion in Human Insula and Amygdala. *Cerebral Cortex*, *20*(4), 929–940. https://doi.org/10.1093/cercor/bhp155

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L., & Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *the Journal of Neuroscience/the Journal of Neuroscience*, *27*(9), 2349–2356. https://doi.org/10.1523/jneurosci.5587-06.2007

Tanovic, E., & Joormann, J. (2019). Anticipating the unknown: The stimulus-preceding negativity is enhanced by uncertain threat. *International Journal of Psychophysiology*, *139*, 68–73. https://doi.org/10.1016/j.ijpsycho.2019.03.009

Walentowska, W., Paul, K., Severo, M. C., Moors, A., & Pourtois, G. (2018). Relevance and uncertainty jointly influence reward anticipation at the level of the SPN ERP component. *International Journal of Psychophysiology*, *132*, 287–297. https://doi.org/10.1016/j.ijpsycho.2017.11.005