Psychophysiology

WILEY

PSYCHOPHYSIOLOGY

SPECIAL ISSUE ARTICLE OPEN ACCESS

Transcutaneous Vagus Nerve Stimulation Enhances Probabilistic Learning

Resul Çakır[1,2] | İlkim Büyükgüdük[2] | Petek Bilim[3] | Ataberk Erdinç[1] | Maria Geraldine Veldhuizen[2,4,5]

[1]Department of Psychology, Economics, Administrative and Social Sciences, Toros University, Türkiye | [2]Department of Psychology, Science and Letters, Mersin University, Türkiye | [3]Department of Psychology, Economics, Administrative and Social Sciences, KTO Karatay University, Türkiye | [4]Department of Anatomy, Medicine, Mersin University, Türkiye | [5]National Magnetic Resonance Research Center (UMRAM), Bilkent University, Türkiye

Correspondence: Maria Geraldine Veldhuizen (margaveldhuizen@gmail.com)

Received: 1 April 2024 | **Revised:** 24 January 2025 | **Accepted:** 29 January 2025

Funding: The authors received no specific funding for this work.

Keywords: extinction | monetary rewards | probabilistic learning | tVNS

ABSTRACT

tVNS enhances various memory and learning mechanisms, but there is inconclusive evidence on whether probabilistic learning can be enhanced by tVNS. Here, we tested a simplified version of the probabilistic learning task with monetary rewards in a between-participants design with left and right-sided cymba conchae and tragus stimulation (compared to sham stimulation) in a sample of healthy individuals ($n = 80$, 64 women, on average 26.38 years old). tVNS enhances overall accuracy significantly ($p = 4.09 \times 10^{-04}$) and reduces response times ($p = 1.1006 \times 10^{-49}$) in the probabilistic learning phase. Reinforcement learning modelling of the data revealed that the tVNS group uses a riskier strategy, dedicates more time to stimulus encoding and motor processes and exhibits greater reward sensitivity relative to the sham group. The learning advantage for tVNS relative to sham persists ($p = 0.005$ for accuracy and $p = 9.2501 \times 10^{-27}$ for response times) during an immediate extinction phase with continued stimulation in which feedback and reward were omitted. Our observations are in line with the proposal that tVNS enhances reinforcement learning in healthy individuals. This suggests that tVNS may be useful in contexts where fast learning and learning persistence in the absence of a reward is an advantage, for example, in the case of learning new habits.

1 | Introduction

Learning how to respond to new stimuli in our environment, despite a lack of prior knowledge of the ensuing outcomes, is an essential task in daily life. Stimulus–outcome contingencies are gradually learned by trial and error, resulting in an intuitive preference for stimuli associated with a higher reward probability (Frank and Kong 2008). Widespread results from animal and human studies suggest that increased dopamine release in the brain is associated with reinforcement (Schultz 2001). The vagus nerve, among other autonomic parasympathetic functions, carries information about our internal state from the gut to our brain. The vagus nerve also has sympathetic catecholaminergic fibres that release noradrenaline and dopamine (Seki et al. 2014; Yang et al. 1999; Verlinden et al. 2016). In recent studies, it has been shown that vagus nerve stimulation induces dopamine release from the ventral tegmental area (VTA; Han et al. 2018; Manta et al. 2013; Fernandes et al. 2020). This suggests that vagal afferents play a big role in controlling behaviour related to rewards (Tellez et al. 2013), and they have a vital function in regulating reward-related behaviours (Teckentrup and Kroemer 2023; Berthoud 2008a, 2008b, de Lartigue 2016). Studies conducted on rats have shown that vagus stimulation improves taste and location preference learning by modulating dopamine release in the brain (Han et al. 2018). In addition, Fernandes et al. (2020) discovered that damage to the hepatic branch of the vagus nerve hinders the functioning of dopamine in the

VTA, which in turn affects the behaviour of seeking food. Furthermore, the connections between brain areas involved in dopamine regulation and memory play a vital role in improving memory formation when exposed to pleasurable stimuli (Dalley and Everitt 2009). Collectively, these discoveries indicate that vagus neurons play a crucial role in reward pathways.

Transcutaneous auricular vagus nerve stimulation (tVNS) is a noninvasive, safe technique of electrical stimulation of the auricular branch of the vagus nerve (Ellrich 2011; Redgrave et al. 2018), which connects the gut and the brain (Berthoud, 2008; de Lartigue 2016; Yuan and Silberstein 2016). Compared to invasive vagus stimulation (VNS), tVNS provides similar outcomes and activation in similar brain structures (Kaniusas et al. 2019; Frangos and Komisaruk 2017; Kraus et al. 2013).

There is a growing body of evidence of tVNS causing changes in psychological processes. Although the mechanism of action of tVNS is not yet fully understood, it has been shown to have an enhancing effect on many cognitive functions through its wide connections with different regions of the brain (Jandackova et al. 2023; McIntire et al. 2021; Giraudier et al. 2020; Colzato et al. 2018). Recent studies showed that tVNS enhances learning and memory in various domains including associative memory (Sellaro et al. 2018; Jacobs et al. 2015), working memory (Sun et al. 2017), language learning (Aguiar and Kaan 2019; Kaan et al. 2021), lexical learning (Phillips et al. 2021) and implicit learning (Jongkees et al. 2018). To our awareness, so far two studies have tested the effect of tVNS on probabilistic reinforcement learning (Weber et al. 2021; Kühnel et al. 2020). Weber et al. (2021) tested the effects of left-sided cymba conchae stimulation (versus sham stimulation in the left ear lobe) in eight participants with epilepsy (1 year seizure-free) in a probabilistic learning task with varying levels of probabilities between visual symbols and monetary rewards in a within-participants design. Weber et al. (2021) observed greater cumulative accuracy in the learning phase and greater preservation of accuracy rates during the extinction phase for tVNS relative to sham. Kühnel et al. (2020) used a probabilistic go–no go task with visual patterns and monetary reward in 39 healthy participants after an overnight fast and performance in various other tasks. They tested the effect of left-sided cymba conchae stimulation relative to sham stimulation in the left earlobe in a within-participants design. They observed lower learning in the tVNS group, driven by lower learning from punishment. Kühnel et al. (2020) also noted that when tVNS was applied in the first session, the learning reduction was greater. Clearly, there are inconsistencies in the few observations that exist, and more insight into the effect of probabilistic reinforcement learning is necessary. Lack of sufficient participants, possible influences of cognitive dysfunction, relatively high within-participant variation and lack of adequate control conditions may have contributed to the inconsistencies in observations in these previous studies.

Here we tested the effect of tVNS versus sham in a modified probabilistic learning task (mPLT). Given the order effect that Kühnel et al. (2020) observed, we opted for a between-participants design to reduce within-participant order effects as a source of variation. To control for between-participants sources of variation in cognitive functioning, we measured performance on secondary tasks that reflect cognitive flexibility and inhibition ability. To reduce within-participants sources of variation related to the varying levels of probability, which create a higher difficulty, we opted for a constant level of reinforcement across three different patterns that are associated with monetary reward. Lastly, we attempted to maximise the tVNS stimulation efficacy by stimulating both left- and right sides in both cymba conchae and tragus. Both left- and right-sided stimulation are equally safe (Redgrave et al. 2018) as the signal from both auricular branches of the vagus nerve is integrated before activating vagal efferents to the heart (Chen et al. 2015; Kim et al. 2022). The auricular branch of the vagus nerve innervates both the tragus and the cymba conchae (Peuker and Filler 2002) and while many studies have opted to stimulate one of these two locations, it may be most effective to stimulate both. We hypothesised that tVNS relative to sham would enhance probabilistic reinforcement learning and lead to the persistence of the learned stimulus–outcome associations postreinforcement.

2 | Methods

2.1 | Study Participants

There were no published effect sizes of tVNS effects on a similarly modified probabilistic learning task available at the time of the design of this study. To give the reader context for the effect size that our sample size is powered to observe, we performed a post hoc power analysis. A post hoc power analysis [with G*Power 3.1.9.7, Faul et al. 2007] for a fixed effects omnibus one-way F-test, with inputs of two groups, with a desired power of 0.95, an alpha of 0.05 and an effect size $f = 0.4$, yielded a calculated desired sample size of 84. This analysis shows that this study is powered to observe a large effect size.

We recruited 80 participants from Toros University undergraduate students and academicians. Inclusion criteria were as follows: being healthy and between the ages of 18–40 (mean = 26.45, SD = 5.43, min = 19, max = 40; 56 women and 24 men). We excluded participants with heart diseases (e.g., coronary artery disease, cardiac arrhythmia and bradycardia), epilepsy, blood pressure disease, significant skin diseases, drug addiction, alcohol addiction, neurological and psychiatric disorders and pregnancy from the study. The data for the study were collected between June and November of 2023. Depending on their performance in the probabilistic learning task, participants earned between 0 and 120 Turkish Liras in the study. Ethical approval was received for the study from Toros University Social Science Institute (number 2023/90) and all participants signed a written informed consent form. The study was conducted in agreement with the Declaration of Helsinki.

2.2 | Procedure

Participants took part in a single-blind randomised controlled experiment investigating the efficacy of tVNS on a modification of the probabilistic learning task (mPLT) (Weber et al. 2021; Kunisato et al. 2012), which included only a learning phase (with feedback), an extinction phase (without feedback) and a constant level of reward probability. After arriving at the laboratory, participants were seated in front of a table with a 21.5-in. computer screen and a standard Turkish keyboard. The computer screen was positioned such that participants sat approximately 30 cm away from it. All tests were conducted in the same room at a constant temperature (22°C). Before the administration of the mPLT and tVNS, we measured cognitive traits with the Wisconsin card sorting test (WCST) and Flanker test to exclude differences between the groups in terms of cognitive functions. Participants were randomly divided into two groups: the vagus stimulation group (tVNS) and the sham stimulation group (sham) for the mPLT. The WCST and Flanker test were administered using the psychology experiment building language (PEBL2) computer application (Mueller and Piper 2014). The mPLT was conducted using custom-written Python (Python Software Foundation 2019) code. After the mPLT, we also measured whether the presented symbols led to any conscious associations and whether participants had the impression that some pairs were chosen harder than others. The Python code used to run the mPLT experiment and Supporting Information on the data obtained are available at https://osf.io/w5vtu/.

2.3 | tVNS and Sham Stimulation

The auricular branch of the vagus nerve innervates the cymba concha and tragus of the outer ear (Yakunina et al. 2017; Kreuzer et al. 2012), and researchers frequently target these locations in basic cognitive neuroscience studies with tVNS (Yuan and Silberstein 2016). In this experimental design, tVNS-naive participants received transcutaneous vagus nerve stimulation on the left and right ear cymba concha and tragus and sham stimulation on the left and right ear lobes. In the study, a tVNS device (Vagustim Sağlık Teknolojileri A.Ş., İstanbul, Türkiye) connected via Bluetooth and providing electrical stimulation with two titanium electrodes was used for transcutaneous stimulation. For the tVNS group, two electrodes in opposite directions were placed inside the cavum concha (see Figure 2), so both the cymba conchae and tragus received stimulation. We used standard ear clips with electrodes for sham stimulation. The stimulation area for the sham group was chosen from the upper scapula and earlobe region, where there is no vagus nerve distribution. The electrodes provide transcutaneous stimulation with the following parameters: a biphasic square wave pulse at 25 Hz and a pulse width of 250 μs, with a duty cycle of 30 s on, 30 s off and constant voltage (conform reporting guidelines Farmer et al. (2021)). The stimulation intensity was adjusted for each participant for each ear separately. During the adjustment procedure, the stimulation amplitude was gradually increased. The participant was asked to report when a pricking or burning sensation was felt. The stimulation intensity was then immediately decreased until the participant reported feeling an innocuous and comfortable sensation. Stimulation intensity remained at the selected level during the session. The average stimulation amplitudes for sham were 5.88 (right) and 5.94 mA (left) and for tVNS, 6.035 (right) and 6.210 mA (left) (see Supplementary analyses for an analysis of the effect of tVNS and side of stimulation on stimulation amplitude).

2.4 | Cognitive Control Measures

The cognitive control measures of cognitive flexibility and inhibition ability were measured before the tVNS electrodes were attached to the ear.

The WCST (Heaton 1993) was used to evaluate participants' cognitive flexibility. The WCST used in this study was an adaptation based on the WCST from PEBL2 (Mueller and Piper 2014). The WCST is essentially a card-matching task. In each trial, a response card is presented above four stimulus cards that vary along three dimensions: colour (red, blue, yellow and green), shape (circles, triangles, stars and crosses) and quantity (one, two, three and four). During each trial, the participant was required to 'match' the response card with one of the four stimulus cards without any specific guidance from the administrator. The sorting rule, that is, the dimension by which the card should be correctly matched, is something the participant needs to figure out through a process of trial and error. For example, a response card showing two blue triangles can potentially be matched based on colour (blue), shape (triangle) or quantity (two). Following each response, the participant received feedback in the form of 'correct' or 'incorrect', which aids in determining the correct sorting rule. Each participant completed the 64-card version of the WCST, which was administered and scored in the standard version (Heaton 1993). In the standard version of the WCST, correct response, total errors, perseverative responses, perseverative errors, nonperseverative errors, unique errors, trials to complete the first cat, failure to maintain set, learning to learn and conceptual level response parameters can be evaluated. The perseverative error parameter, which is defined as persistence in previous erroneous trials and which is reported to predict cognitive flexibility, was evaluated in the current study (Carruthers et al. 2019; Çuhadaroğlu 2016). Thus, a higher amount of perseverative errors was interpreted as lower cognitive flexibility.

We used the Flanker test to evaluate participants' inhibition ability. The Flanker test used in this study was an adaptation of the Flanker test from PEBL (Mueller and Piper 2014). We presented a target array consisting of five arrows in the centre of the screen. The participant had to respond quickly and accurately in the direction of the central arrow in congruent and incongruent conditions. In the congruent condition, the central arrow faced the same direction with the surrounding arrows (e.g., ←←←←← or →→→→→) whereas in the incongruent condition, the central arrow faced the opposite direction with the surrounding arrows (e.g., ←←→←← or →→←→→). The participant completed a total of 160 trials. In each trial, the congruent, neutral and incongruent flanker arrays were presented in random order with an equal probability of occurrence. Prior to the experimental trials, the participant completed a practice session of eight trials to become familiar with the procedure. The participant was instructed to respond to the targets quickly and accurately. During the practice session, the participant received onscreen feedback regarding their performance at the end of each trial. In the Flanker test, total errors, mean accuracy, mean response time, congruent mean and incongruent mean can be obtained. In this study, the total number of errors and mean accuracy were taken into account to assess the participants' disinhibition levels. A higher total number of errors and a lower mean accuracy were interpreted as a lower inhibition ability.

2.5 | Pilot Study to Assess Appropriateness of the Symbols Used in the mPLT

In this pilot study, three abstract symbol pairs were generated using the Blossom font for use in the mPLT task (See Figure 1). We conducted this pilot study to ensure that the chosen symbols did not evoke any conscious associations among the participants. The pilot study was conducted with 300 participants; none of these pilot study participants took part in the main experiment. In the pilot study, six abstract symbols of the same size and with equal spacing were printed on an A4 sheet and presented to the participants, and they were asked to indicate whether each abstract symbol evoked any association by choosing one of the yes or no options under each symbol.

2.6 | Experimental Procedure of the Modified Probabilistic Learning Task (mPLT)

Participants were divided into two groups, assigned randomly: active vagus stimulation group (cymba conchae and tragus) and sham stimulation group (earlobe). Before the experiment, the experimenter used the rand() and balance() commands in Excel to determine which group each participant would be included in, such that 40 tVNS and 40 shams were randomly assigned.

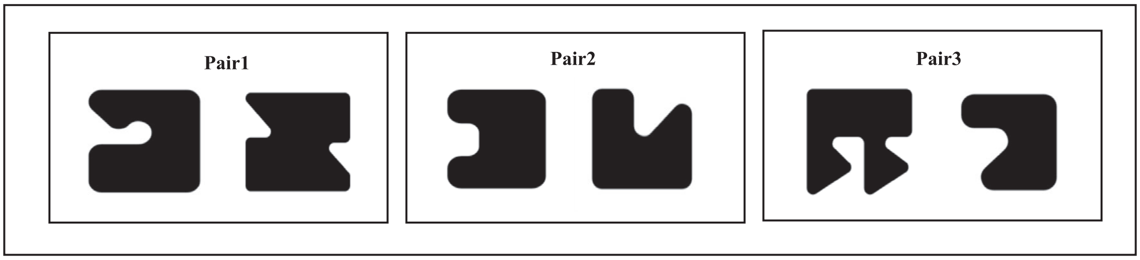

FIGURE 1 | Example of symbol pairs from the Blossom font that were used in the probability learning task.

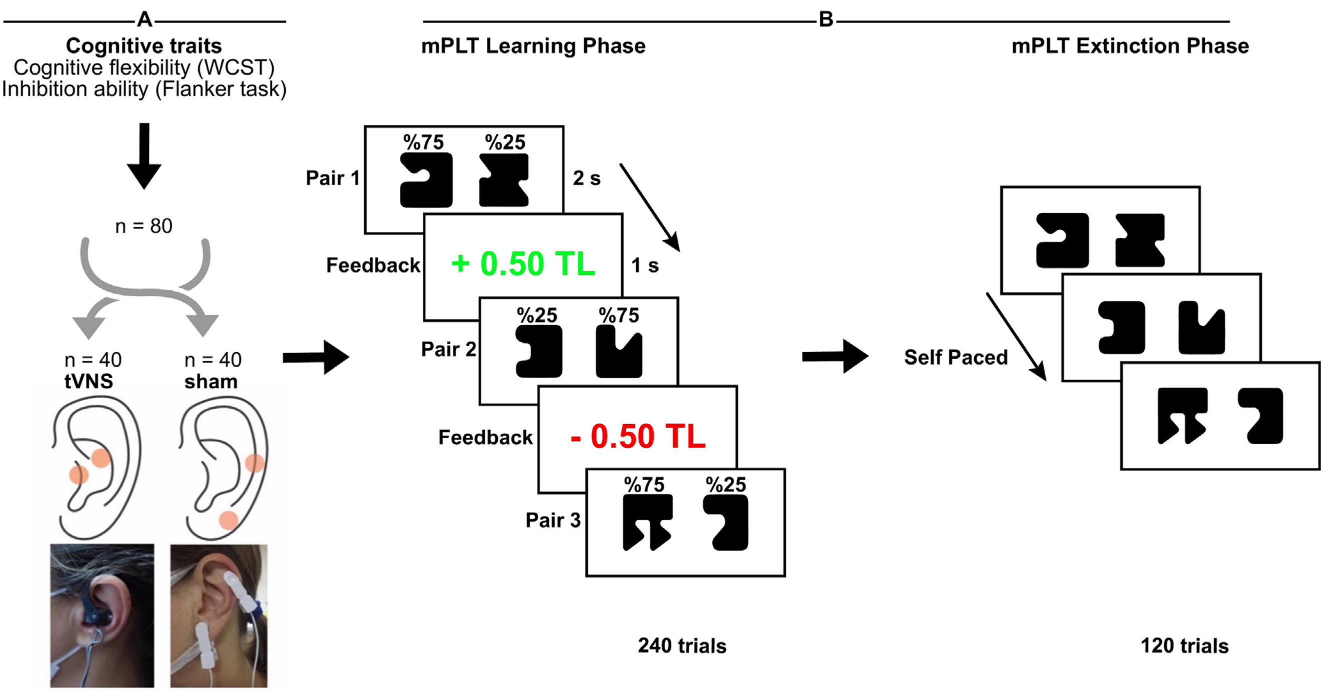

FIGURE 2 | Schematic overview of the experimental design. A: Participants first completed cognitive tasks and were then randomly assigned to the tVNS or sham group and outfitted with the electrodes as shown in the drawing and photographs. The stimulation lasted for the duration of the experiment in 30-s on–off cycles. B. The experimental task consisted of two sequential parts: The modified probabilistic learning and Extinction phases. The stimuli consisted of six letters from the Blossom alphabet (https://www.dafont.com). In the modified probabilistic learning phase, participants were repeatedly presented with fixed pairs of stimuli (240 trials, 80 trials per set). Participants were instructed to answer as quickly as possible and that their answer after 2 s would be considered incorrect. To maximise the total payoff, participants had to learn which stimuli were more highly associated with reward. Stimulus pairs during the learning phase were fixed, but the side of the presentation was randomised for each trial. After the modified probabilistic learning phase, the Extinction phase followed (120 trials, 40 trials per set), during which no feedback was given. Each trial presented a randomised combination of six previously presented stimuli in the extinction phase. The whole experiment took 15–20 min in total.

Stimulation (tVNS or sham) was activated before the practice phase and continued until the extinction phase was completed.

The task consists of three consecutive phases: the practice phase, the learning phase and the extinction phase, each described in detail below. The design was a modification of the probabilistic learning task based on Weber et al. (2021); (Kunisato et al. 2012). In total, the duration of the paradigm was 20 min. The experimental design is summarised in Figure 2. The experimenter ensured that each participant understood the task before starting with the paradigm.

2.7 | mPLT Practice Phase

The participant received written and oral instructions and several practice trials. The practice phase consisted of 12 trials. Different symbols were used in this phase of the study and the learning phase.

2.8 | mPLT Learning Phase

During the learning phase, the participant was repeatedly presented with three fixed pairs of stimuli (240 trials, 80 trials per set). The stimuli consisted of six abstract symbols from the artificial Blossom alphabet (https://www.dafont.com). In each trial, the participant saw two abstract symbols, one on the centre-right and the other on the centre-left of the screen, and they were instructed to learn through trial and error which symbol was associated with a high reward. The symbols were displayed for 2 s. One of the symbols displayed on the screen is associated with a reward with a probability of 75%, while the other is associated with a reward with a probability of 25%. This contingency never changed during the experiment. Our justification for choosing this specific reinforcement probability is as follows: Our study design was primarily based on the study by Weber et al. (2021). Since they showed an absence of performance greater than chance in the sham condition with three different contingency levels (80%–20%, 70%–30% and 60%–40%), and they also showed that learning in the tVNS condition occurred only after 123 trials, we wanted to create an easier-to-learn version of the probability learning task to better examine the effect of tVNS on reinforcement learning in a healthy population. We also wanted to ensure that we could observe any learning (performance greater than chance) under sham stimulation (unlike Weber et al.). As the learning rate generally tracks the level of the contingency (faster learning for more extreme reward contingency levels), we chose a relatively high reward probability, that is, a level not close to 50% versus 50%; in our case, 25% versus 75%. Reinforcement learning and evidence accumulation models also often assume noisy accrual of evidence for responses to a stimulus. If we make another assumption of limited processing resources, one could predict that learning would be slower in the context of variable contingency levels that include conditions closer to 50% reward contingencies. Therefore we elected to keep the reward probability constant across the three pairs but included three pairs to exclude idiosyncratic influences of one specific symbol pair. The participant was instructed to indicate which of the two stimuli was associated with a higher reward as fast as possible by pressing one of two arrow keys. The participant had 2 s to respond, responses slower than 2 s were automatically considered incorrect. When a participant chose a symbol associated with a reward, they received feedback on the next screen (displayed for 2 s) stating ‘You won 0.5 TL’, and if they chose the wrong symbol, they received feedback saying, ‘You lost 0.5 TL’. The locations of symbols on the screen changed throughout the repetitions, but the correct answer remained the same. The assignment of the reward to a symbol was counterbalanced among participants. The intertrial interval was 1 s. Additionally, the participant was able to see their current balance throughout the experiment. Before starting the experiment, the participant was informed that the reward they would receive for their correct answers in the experiment would accumulate in their balance and that this accumulated amount would be paid to them at the conclusion of the experiment. In each trial of the probabilistic learning task, the accuracy and response time of the participant's answers were measured. Higher accuracy represents higher learning.

2.9 | mPLT Extinction Phase

After completion of the learning phase, an extinction phase followed, in which no feedback was given. The extinction phase aims to examine the effect of continued vagus nerve stimulation on performance without reinforcement. The extinction phase is performed using the same images used in the learning phase. However, at this stage, participants are not rewarded for their correct answers or penalised for their incorrect or slow responses. In addition, the participant did not receive any feedback on the accuracy of their responses. The extinction task took 120 trials. During this phase, the three same symbol pairs from the learning phase were presented 40 times each, in a random order. In the extinction phase, each trial's accuracy and response time were measured. Higher accuracy represents higher persistence of previously learned associations.

2.10 | Postexperiment Questionnaire

Following the experimental phase, two questionnaires printed on A4 paper were shown to the participants. The goal of administering this questionnaire was 1) to test whether the symbols used in the mPLT evoked any associations and 2) to indicate whether some pairs were more difficult to choose than others. In the first questionnaire, we showed the participants the symbols used in the mPLT and asked them to indicate whether each of the individual symbols evoked any association by choosing ‘yes’ or ‘no’. In the second questionnaire, the participants were first asked whether it was more difficult to choose some pairs than other pairs by choosing one of the options ‘yes’ or ‘no’. If the answer was ‘yes’ we asked the participants to indicate which pair they considered most difficult.

2.11 | Data Analyses

All analyses were conducted using JASP 0.18.1 (JASP Team 2023) and R 4.2.2 (R Core Team 2022). The response time and accuracy dependent variable analyses are primary, as well as the control variable analyses, the demographic variable analyses, the pilot study data analyses and the postexperimental questionnaire analyses. The RLDDM analyses are exploratory analyses. The response time and accuracy dependent variable analyses, the control variable analyses, the demographic variable analyses, the pilot study data analyses and the postexperimental questionnaire analyses were conducted using JASP 0.18.3, while the RLDDM and its associated analyses were conducted using R.

2.12 | Data Cleaning and Organisation

Ratcliff and Tuerlinckx (2002) emphasise that responses faster than 0.2 s indicate fast guesses. Such responses are invalid for data analysis. Additionally, such short RTs cause serious problems for diffusion model fitting and extreme biases in parameter estimates, so trials with RTs faster than 0.2 s were excluded from all subsequent analyses using the mPLT data (see Ratcliff and Childers 2015). In total, 1.09% of the total number of trials for the tVNS group and 2.40% for the sham group were excluded from the data set (the participant with the most data excluded had 18 trials excluded).

We divided the trials in the learning and extinction phases into 4 blocks of 60 trials. The decision to include block as a factor was exploratory and the choice for the number of trials in a block was based on convenience. The justification for exploring the effect across blocks is that upon initially graphing the results, we wondered if performance was similar between the conditions at the beginning of the task and whether performance stabilised over time. We had no specific predictions regarding the pattern over time, but including block as a factor would allow us to perform 'manipulation checks'. For example, if there is learning, then there should be chance-level performance at first, as well as an increase in accuracy over time, as well as no decrease in accuracy over time.

2.13 | Learning Phase Accuracy and Response Time

Trial-level measures (response time and accuracy) were analysed using a frequentist multilevel modelling approach. We specified a binomial (logit linked) generalised linear mixed model (GLMM) with a likelihood ratio significance test method for the dependent variable accuracy. We specified the GLMM with fixed effects variables group (tVNS vs. sham) and block (1–4) and a random effects grouping factor 'subject ID'. We specified a Gaussian (identity-linked) linear mixed model (LMM) with the Satterthwaite significance test method for the dependent variable response time. We specified the LMM with fixed effects variables group (tVNS vs. sham) and block (1–4) and a random effects grouping factor 'subject ID'.

Each model included an interaction term for group by block. We report the simple effects and the interaction effects. Additionally, in order to examine the effect of groups over the course of time, we inspected custom post hoc comparisons of the groups within each block. We also specified a linear contrast of blocks to examine if there was a general effect of learning regardless of group (see Table S1). For the post hoc comparisons conducted in the accuracy GLMM and response time LMMs, *p*-value adjustment was performed using the Holm method to account for four multiple comparisons. Additionally, for response times, the Satterthwaite method was used to estimate degrees of freedom.

We note that in the first block of 60 trials, to examine whether accuracy rates started at 0.5 and whether the groups performed similarly in the first trials of the learning phase, we performed supplementary analyses with mini-blocks of six trials each for the first block only, which we describe and report in detail in the Supporting Information (for detailed information see Tables S18 and S19). We also include Bayesian statistics for response time and accuracy in Table S2, but these are not included in the results section due to high agreement with the frequentist approach.

2.14 | Reinforcement Learning Drift-Diffusion Modelling (RLDDM)

Since the GLMM revealed that most learning took place within the first block and performance within the group was largely stable in the other blocks, we included only Block 1 in the RLDDM. hBayesDM (Ahn et al. 2017) generally gives stable results, especially for models with few repetitions, providing consistent estimates even for $n < 100$ (Lerche et al. 2017; Wiecki et al. 2013). For inspection of the stability of the RLDDM model results, we refer the interested reader to the Supplementary Analysis for an examination of tVNS effects on RLDDM for other blocks, which were generally consistent with the effects observed in block 1 (Tables S5, S6, S8 and Figure S2).

The drift-diffusion model (DDM) and reinforcement learning (RL) models are combined to understand greater details of decision-making processes in, for example, instrumental learning tasks. We fitted an RLDDM (Model 6, Pedersen et al. 2017) with R (R Core Team 2022) and the hBayesDM package (Ahn et al. 2017). This RLDDM uses trial-based prediction errors to assign values to stimulus–action pairs in a delta learning model framework (Rescorla 1972). RLDDM is initialised with priors that reflect established findings in the literature. Then, the Markov Chain Monte Carlo method fits the model to the data by estimating the joint posterior distribution for all parameters. These benefits make a Bayesian hierarchical framework especially valuable when estimating individual parameters for complex models based on limited data (Ahn et al. 2017). In our RLDDM, we drew 4000 samples from the posterior distribution and discarded the first 2000 (burn-in samples; Kruschke 2014).

2.15 | Reinforcement Learning Drift-Diffusion Model Fit

2.15.1 | Relative Model Fit

To evaluate model quality, we used several key metrics: the Rhat values (Gelman–Rubin statistic; Gelman and Rubin 1992), the expected log pointwise predictive density (ELPD; Vehtari et al. 2017) and the leave-one-out information criterion (LOOIC). Detailed results are presented in Table S3.

The Rhat statistic (Gelman et al. 1996) was employed to assess convergence across chains. This statistic measures the variation between chains relative to the variation within chains, with values close to 1 (max 1.1) indicating satisfactory convergence. We computed the ELPD and LOOIC metrics for each model to assess model fit and generalisation performance. These values were derived from an 8000×40 log-likelihood matrix. The ELPD and LOOIC are critical measures for evaluating how well a model fits observed data and generalises to new data (Vehtari et al. 2017). An increase in ELPD indicates improved model fit and generalisation capacity, while a decrease in LOOIC—due to its negative scaling—reflects better model performance (Gelman et al. 2014). Models with higher ELPD and lower LOOIC are preferred, as they provide stronger evidence of predictive accuracy. Conversely, a decrease in ELPD or an increase in LOOIC suggests weaker performance, potentially due to overfitting or underfitting, indicating the need for model refinement (Vehtari et al. 2017). Thus, improvements in ELPD and reductions in LOOIC are interpreted as indicators of superior model performance.

We estimated hyper-group and individual parameters dependent on the vagus nerve stimulation manipulation (active VNS or sham). Group-level parameters of the between-subjects vagus nerve stimulation effects were used to assess how the stimulation influenced performance. We also examined whether the two models fit the RLDDM model better than a pure DDM model, which is assumed to have no learning processes, that is, static decision parameters. To compare the two drift-diffusion modelling approaches, we calculated LOOIC and ELPD values (see Table S3).

2.15.2 | Absolute Model Fit

In addition to a relative model comparison, we further investigated the best model using absolute model fit measures to determine whether it could capture key features of the data. For this purpose, we compared the observed data with data simulated based on the estimated parameters. The simulated data were generated using the hBayesDM package (Ahn et al. 2017). To generate the simulated data, we utilised the output of the hBayesDM package, which provided simulated response time (RT) data in a three-dimensional array format of $8000 \times 40 \times 60$. This structure represents 8000 simulations of RTs for 40 participants across 60 trials in an experimental design. We conducted this procedure separately for the tVNS and sham groups. Following the generation of simulated data, we computed the mean RTs across the 8000 simulations for each participant within each group. This allowed us to obtain a set of RT values for each participant over 60 trials (trials with RT below 2000 ms that were excluded in data cleaning were not simulated here as well, so for some participants, this number is less than 60). These aggregated RT values were then used for further comparison with the observed RT distributions to assess model performance. To evaluate the model's performance, we conducted posterior predictive checks (Gelman et al. 2013; Gelman and Hill 2007), which involved comparing the observed distributions of response times (RTs) with those generated from posterior predictive distributions. This procedure allowed us to assess how well the model captured the underlying cognitive processes and reproduced the observed RT distributions.

2.15.3 | Parameter Recovery

To determine whether the tVNS and sham group data fit the RLDDM model, we performed a parameter recovery study by estimating the posterior distributions of the parameters on the simulated data with the hBayesDM package (Ahn et al. 2017). With the two data sets obtained from the simulation (separately for the tVNS and sham groups), the RLDDM model was run with two MCMC chains, with 2000 burn-in samples and 4000 iterations for each chain. We examined the means of the posterior distributions for the estimated and simulated parameter values together.

2.15.4 | Estimated Parameters From RLDDM

From the RLDDM analysis, we obtained the drift rate (v), boundary separation (a), nondecision time (tau), negative learning rate (η^-) and positive learning rate (η^+) estimations. These parameters can be interpreted as follows: Drift rate is related to the rate of learning and reward sensitivity and represents the level of exploitation of information. Larger values reflect better performance (Pedersen and Frank 2020; Pedersen et al. 2017). The boundary separation parameter 'a' describes the amount of evidence needed until a decision threshold is reached and thus controls the speed–accuracy trade-off—a smaller boundary separation emphasises speed over accuracy (more error prone), whereas the reverse is true for a larger boundary separation (more conservative) (Chakroun et al. 2023; Pedersen et al. 2017). The nondecision time parameter ('tau') captures the time spent on sensory perception, motor preparation and motor output, which are components of the RT unrelated to the evidence accumulation process (Pedersen and Frank 2020; Pedersen et al. 2017). The negative learning rate and positive learning rate estimations reflect reward prediction error, the discrepancy between observed and predicted reward, which plays a central role in the delta learning model framework (Rescorla 1972). We represented RLDDM parameters in Figure 3. We estimated hyper-group and individual parameters depending on the stimulation group (tVNS vs. sham) and block (1–4). We used group-level estimation of the between-subjects stimulation effect to assess how tVNS affected cognitive performance. To assess the effect of tVNS, we used GLMMs. The GLMMs were specified with fixed factors 'group' (tVNS vs. sham) and 'block' (1–4) and random effects factor 'participant' on each of the dependent variables of RLDDM posterior estimation values (a, v, tau, η^+, η^-).

To assess the effect of tVNS on RLDDM, we calculated the posterior means for each group (tVNS and sham) 95% highest-density intervals (HDI) and the Bayes factor as evidence for the directional effect. A posterior probability distribution that shows the probability of a parameter value based on the experimental data is produced by Bayesian parameter estimation. High-density intervals (HDI, Kruschke 2018; Körber et al. 2016) are the intervals of this posterior distribution. Values within the HDI indicate the parameter's 95% most credible value and are more probable than values outside of it (Kruschke 2018). Bayes factors

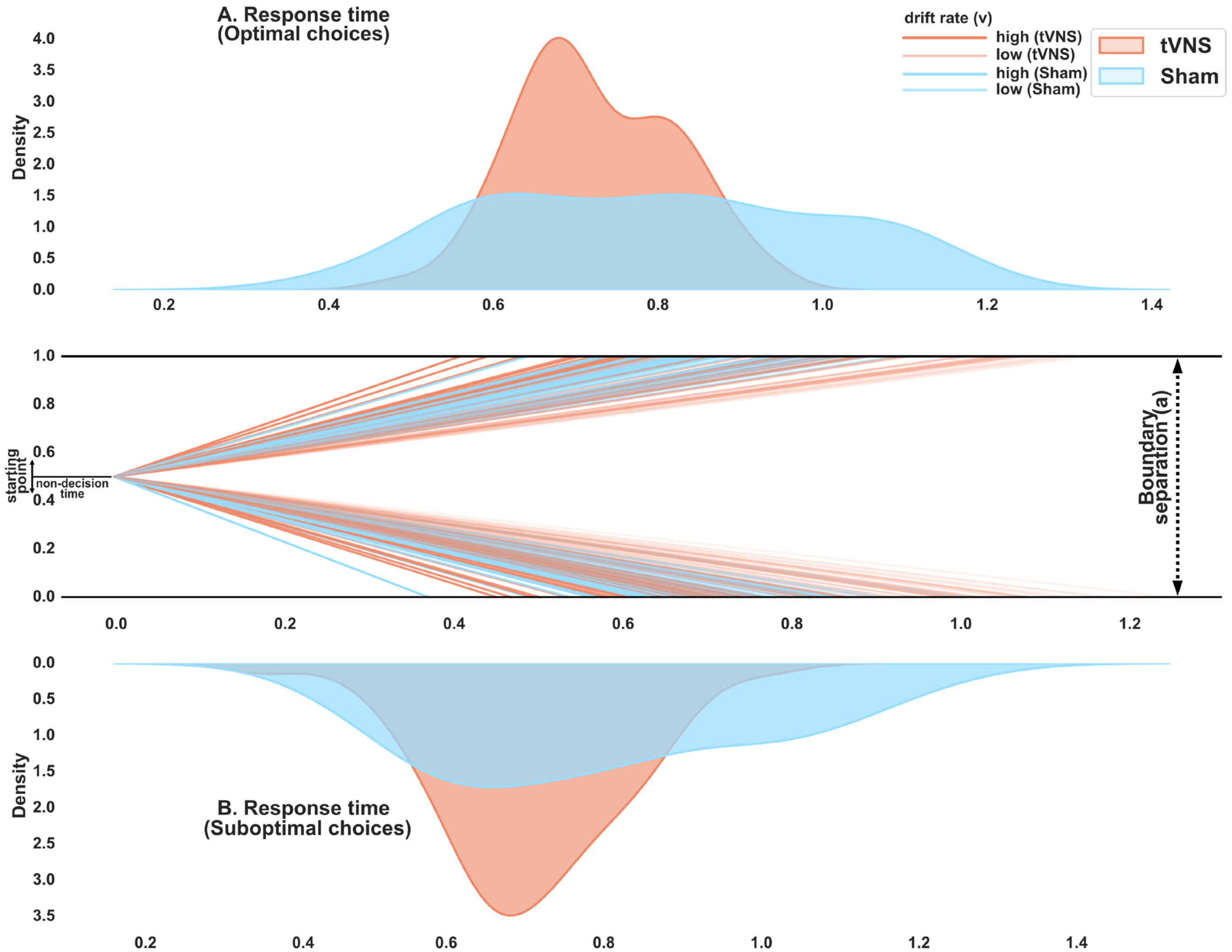

FIGURE 3 | **Drift-diffusion model**. The drift-diffusion model's key characteristics include the process of evidence accumulation, which starts from an initial point (z). Evidence is collected along sample paths incorporating Gaussian noise, continuing until a decision threshold—either upper or lower—is met; at this point, a response is triggered. Drift rates are illustrated by lines, with higher drift rates shown as more saturated colours and lower drift rates depicted with less saturation. The upper and lower panels depict response time (RT) distributions for optimal and suboptimal choices for the tVNS and sham stimulation groups. These distributions, derived from observed data, highlight differences between the two groups in their decision-making processes. Python was employed to visualise, allowing a clear representation of the DDM parameters and response time distributions. The results demonstrate distinct characteristics in the tVNS group compared to the sham group across both optimal and suboptimal responses. The figure aligns with established frameworks in decision-making research (Wiecki et al. 2013) and serves as a comprehensive visualisation of the DDM.

(BF) for directional effects are calculated as i/(1—i), where i is the integral of the posterior distribution from 0 to + (Marsman and Wagenmakers 2017; Jeffreys 1998).

2.15.5 | Extinction Phase Accuracy and Response Time

The statistical analysis for accuracy and response time in the extinction phase was identical to the analysis for the learning phase with the following adjustment: We included only two levels of the fixed effects variable block (Blocks 1–2). We also included Bayesian statistics for response time and accuracy in Table S10, but these are not included in the results section due to high agreement with the frequentist approach. Additionally, we ran an analysis to assess whether accuracy exceeds chance levels in this phase (see Table S20).

2.15.6 | Cognitive Control Measures and Demographic Data

Bayesian independent samples t-tests were used to examine whether the data obtained from WCST and Flanker and the age of the participants differed between the tVNS and sham groups. A Bayesian contingency table analysis was used to examine whether the gender of the participants differed between the tVNS and sham groups.

2.15.7 | Pilot Study for Preexisting Associations With the Symbols Used in the mPLT

The data obtained from this pilot study were analysed using a Bayesian one-sample binomial test for each symbol. We expected participants' choice of no (no evoked associations experienced) to be greater than the test value of 0.5 in a Bayesian one-sample binomial analysis.

2.15.8 | Postexperiment Questionnaire

A Bayesian one-sample binomial test was used to assess whether the symbols evoked any associations in mPLT participants, and a Bayesian contingency table analysis was used to determine whether there were differences between the tVNS and sham groups. Then, a Bayesian contingency table analysis test was used to examine whether participants' judgments about the difficulty of the symbols and their decisions about which symbol pair was more difficult to choose than others differed between the tVNS and sham groups.

3 | Results

3.1 | mPLT Learning Phase Accuracy and Response Time Show Greater Evidence of Learning for the tVNS Group Relative to the Sham Group

The average accuracy (Panel A) and response time (Panel B) across blocks in the learning phase for each group are displayed in Figure 4. The average accuracy starts around 50% for each group (see also Supplementary Analyses for effects within the first block) and increases to ~75% for the tVNS group and increases to ~65% for the sham stimulation group (see Table S1 for descriptive statistics). The average response time in the learning phase for the tVNS group starts at about ~700, while the sham group starts at about ~800 msec. With subsequent blocks, the tVNS group's response times decrease to about 600, while the sham group increases to about 900 msec. To test whether tVNS improved learning compared to sham, we conducted a generalised linear mixed model (GLMM) with a likelihood ratio test method. We conducted a linear mixed model (LMM) with the Satterthwaite method for response time. We specified each model with fixed factors 'group' (tVNS vs. sham) and 'block' (1–4) and the random effects factor 'participant' on the dependent variables of accuracy and response times.

For response accuracy, we observed a main effect of group (tVNS/sham) [$\chi2$ (1) = 12.491, $p = 4.09 \times 10^{-4}$] and block [$\chi2$ (3) = 47.139, $p = 3.247 \times 10^{-10}$] such that the tVNS group had a higher probability of being accurate regardless of block. There was no significant interaction between the group and block [$\chi2$ (3) = 4.405, $p = 0.221$]. We examined estimated marginal means contrasts (Holm adjusted for multiple comparisons) to test which experimental blocks differed between groups, and we observed that in Blocks 3 and 4, the difference between groups was significant (Table S1). An additional custom contrast for a linear effect of the block was not significant (Table S1).

For response time, we observed a main effect of group (tVNS/sham) [$F = (1, 77.57) = 1257.223$, $p = 1.1006 \times 10^{-49}$], such that the tVNS group had a lower response time regardless of the block. We also observed a significant effect of block [$F = (3, 77.50) = 57.383$, $p = 1.2196 \times 10^{-19}$], such that response times decreased after block 1. In addition, we observed a significant interaction between group and block [$F = (3, 77.50) = 33.890$, $p = 4.2472 \times 10^{-14}$]. The estimated marginal means contrasts showed that the response times were lower for each block in the tVNS group (with Satterthwaite correction, Table S1). An additional custom contrast for a linear effect of block showed that there was a significant decrease in response times with progressive blocks (Table S1).

To summarise, we observe that across blocks the tVNS group learns quickly, in the sense that accuracy improves and participants give faster responses. For the sham stimulation group, the accuracy also improves, but less than for the tVNS group, and participants give slower responses. We refer the interested reader to the Supplementary Analyses for an examination of learning effects within the first block, which shows that there is no significant learning until ~24 trials.

3.2 | Reinforcement Learning Drift-Diffusion Model Fit

3.2.1 | Relative Model Fit

First, to address the relative fit of the RLDDM, we assessed Rhat values, LOOIC and ELPD values. All parameters in both models demonstrated excellent convergence, with maximum Rhat values of 1.003. For the tVNS group, hyper-group Rhat values were as follows: $\mu_a = 1.000$, $\mu_v = 0.999$, $\mu_{\text{tau}} = 0.999$, $\mu_{n+} = 0.999$ and $\mu_{n-} = 0.999$. Similarly, for the sham group, the Rhat values were $\mu_a = 1.001$, $\mu_v = 0.999$, $\mu_{\text{tau}} = 1.000$, $\mu_{\eta+} = 0.999$ and $\mu_{\eta-} = 1.000$. These results indicate that all MCMC chains successfully converged within each model (Gelman and Rubin 1992). When comparing the two models, the tVNS model showed superior performance, with an ELPD_{LOO} of −1887.2, compared to

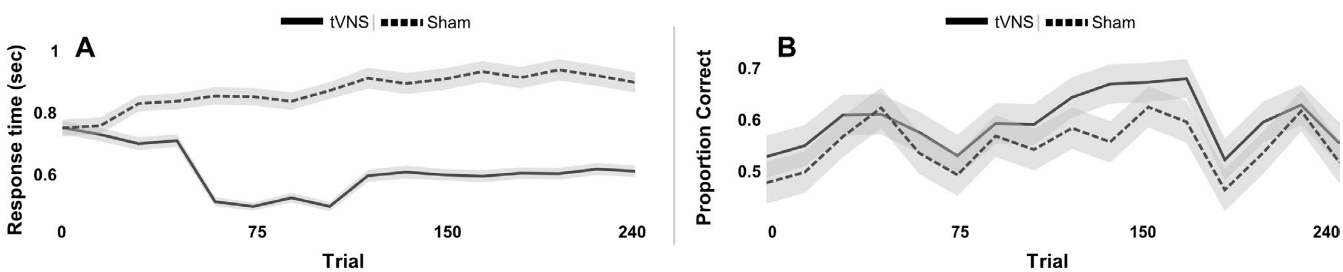

FIGURE 4 | Performance in the mPLT learning phase. (A) Mean response time (B) Proportion correct. Plots show estimated marginal means ±95% confidence interval (CI). The groups are indicated with different line types: dashed line—sham/ear lobe, solid line—tVNS.

−2467.5 for sham, suggesting that the tVNS model provides a better explanation of the data. Similarly, the LOOIC for the tVNS model was substantially lower (3774.3) than for the sham model (4935.1), further supporting the superior fit and balance between explanatory power and complexity in the tVNS model. The Pareto k diagnostic values revealed notable differences in model reliability. For the tVNS model, 92.5% of the data points fell within the 'good' range ($-\infty$ to 0.7), with only 7.5% categorised as 'bad' (0.7–1) and none in the 'very bad' range ($k > 1$). In contrast, the sham model had only 22.5% of data points in the 'good' range, with 65% in the 'bad' range and 12.5% in the 'very bad' category. These findings suggest that the LOOIC estimates for the sham model may be unreliable due to problematic data points, whereas the estimates for the tVNS model are robust. Overall, the tVNS model provides a stronger and more reliable explanation of the data compared to the sham model. To further assess model performance, we investigated whether both models were better fitted by the RLDDM compared to a pure DDM, which assumes static decision parameters without learning processes. The ELPD comparison indicated that both the tVNS and sham models were better fitted by the RLDDM than by the pure DDM (see Tables S3, S7, and Figure S1 for detailed results).

3.2.2 | Absolute Model Fit

Second, to establish absolute model fit, we evaluated posterior predictive checks (Gelman et al. 2013). This procedure entails comparing observed data with data generated from posterior predictive distributions. In this study, the models were developed to account for underlying cognitive processes by integrating response times (RTs). We examined the RT distributions for both observed and simulated data to assess model fit. A well-fitting model should accurately predict the RT distributions associated with observed choices. In Figure 5, we present density plots of posterior predictive RTs overlaid on histograms of the observed RT distributions. Responses favouring suboptimal options are represented as negative RT values. A visual comparison between the observed and predicted distributions reveals a close alignment for both groups. These results indicate that the models successfully capture the RT distributions underlying choice behaviour.

3.2.3 | Parameter Recovery

We conducted a parameter recovery study to validate the fit of the individual estimated parameters of the RLDDM model to the tVNS and sham groups. This involved estimating the posterior distributions of the model parameters using simulated data generated with the hBayesDM package (Ahn et al. 2017). Simulated datasets were generated separately for the tVNS and sham groups, and the RLDDM was applied to each using Markov chain Monte Carlo (MCMC) with 2000 burn-in samples and 4000 iterations per chain. Posterior convergence was assessed using Rhat values, which ranged from 0.999 to 1.003 for the tVNS group and 0.999 to 1.002 for the sham group, indicating successful convergence for all parameters. The means of the posterior distributions closely approximated the simulated values across all parameters, supporting the validity of the RLDDM model for both groups (see Figure 6). Among the simulated models, the simulated tVNS group demonstrated superior performance with lower LOOIC and higher ELPD, as well as smaller standard errors. In contrast, the simulated sham group showed higher LOOIC and lower ELPD, with greater uncertainty. These results indicate that the simulated parameters from the tVNS model better fit the data compared to sham (for detailed information, see Table S3).

3.2.4 | Estimated Learning Parameters From RLDDM for the mPLT Learning Phase

To evaluate whether the results from the RL model were influenced by the tVNS condition, we estimated group- and individual-level parameters for the tVNS and sham groups. Group-level parameters for the effects of ear stimulation were used to assess how tVNS impacts performance (see Figure 7 and Table S4).

Following Jeffreys's evidence categories for Bayes factors (Jeffreys 1998), the analysis indicated strong or very strong evidence that tVNS increased drift rate scaling and nondecision time while decreasing boundary separation. The findings also showed no evidence of a difference between tVNS and sham in terms of positive and negative learning rates; although tVNS numerically resulted in slightly smaller values for these

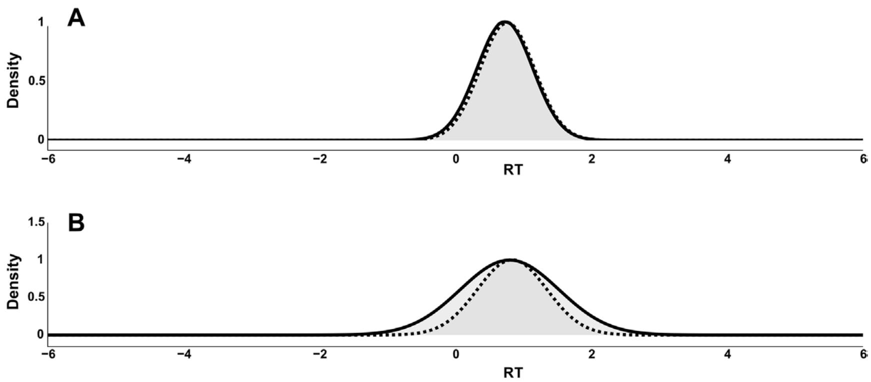

FIGURE 5 | **Posterior predictive RT distributions**. (A) tVNS and (B) sham groups. Density lines represent observed results and the simulation method in solid and dashed black lines, respectively. Choices in favour of the suboptimal option are coded as negative.

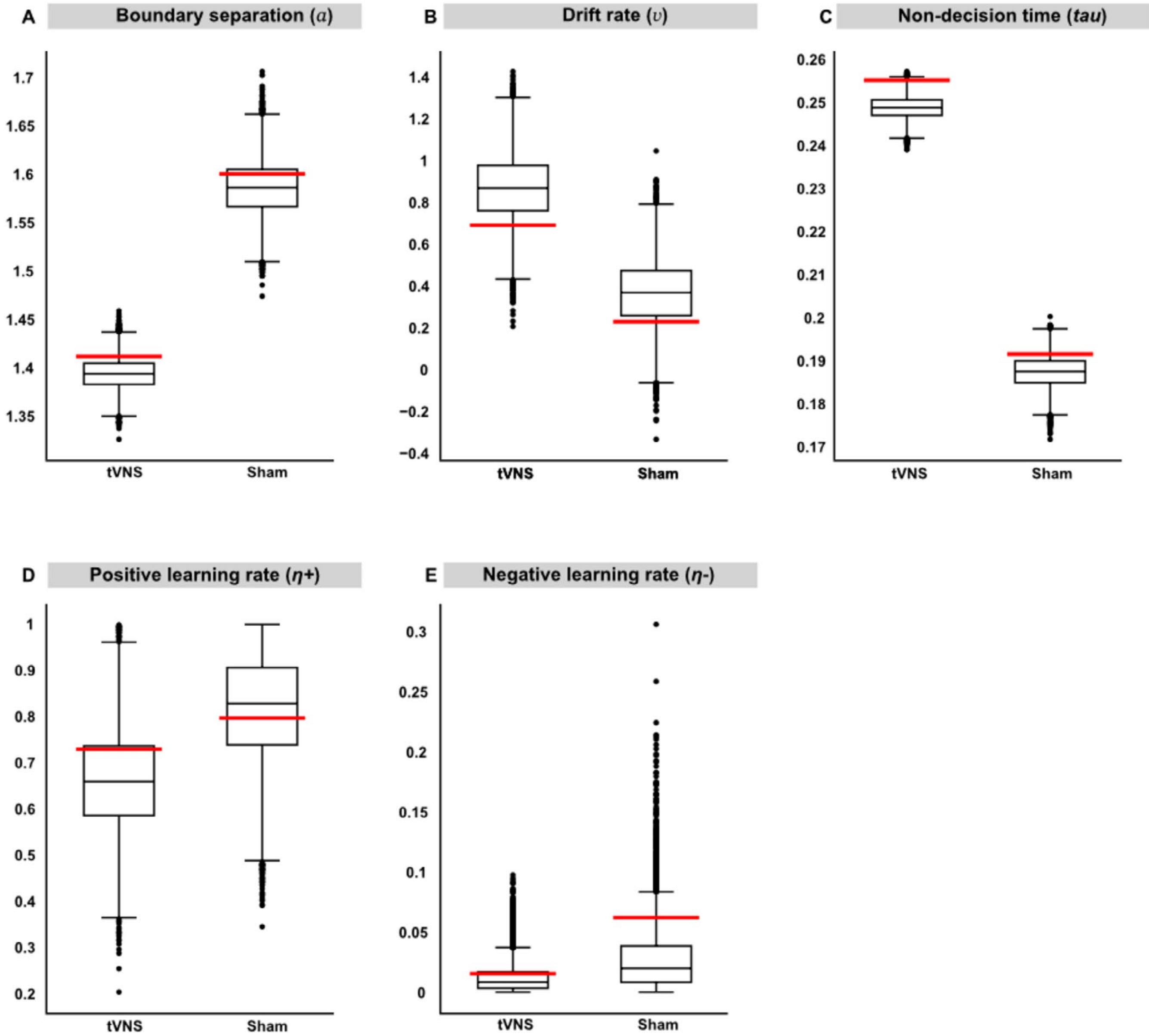

FIGURE 6 | **Parameter recovery result**. The horizontal black and red lines represent the mean parameter values obtained from the observed data and simulated data, respectively. Each plot was divided in half to represent two groups: left tVNS and right sham.

parameters. Finally, we compared the parameter estimates from the two models using the 95% highest density interval (HDI). If the HDI did not overlap with zero, this was taken to indicate a difference between the groups.

The HDI results showed that tVNS increased drift rate scaling (mean difference = 0.488, HDI95% [0.049, 0.919], BF = 70.42) and nondecision time (mean difference = 0.061, HDI95% [0.0518, 0.0703], BF = ∞) while decreasing boundary separation Mean difference = −0.192, (HDI95% [−0.2569, −0.1252], BF = ∞). On the other hand, there was no evidence to indicate a difference between tVNS and sham with respect to positive (mean difference = −0.153, HDI95% [−0.4582, 0.1437], BF = 0.20) and negative (mean difference = −0.015, HDI95% [−0.0846, 0.0323], BF = 0.42) learning rates, although tVNS produced slightly smaller values for these measures (for detailed information, see Table S4). This finding is not so uncommon, as similar findings have been reported in previous studies (e.g., Pedersen et al. 2017; Chakroun et al. 2023; Wiehler and Peters 2024), where positive and negative learning parameters showed less robust effects than the other parameters. These findings indicate that there might be subtle effects that could reflect individual differences or some specific subtleties in task-related processes.

3.3 | mPLT Extinction Phase Accuracy and Response Time Show Greater Perseverance of Learned Behaviour in the Absence of Reward for the tVNS Group Relative to the Sham Group

The average accuracy (Panel A) and response time (Panel B) across blocks in the extinction phase for each group are displayed in Figure 8. Similar to the learning phase, the average accuracy starts at around 50% and increases to ~55% for the tVNS group,

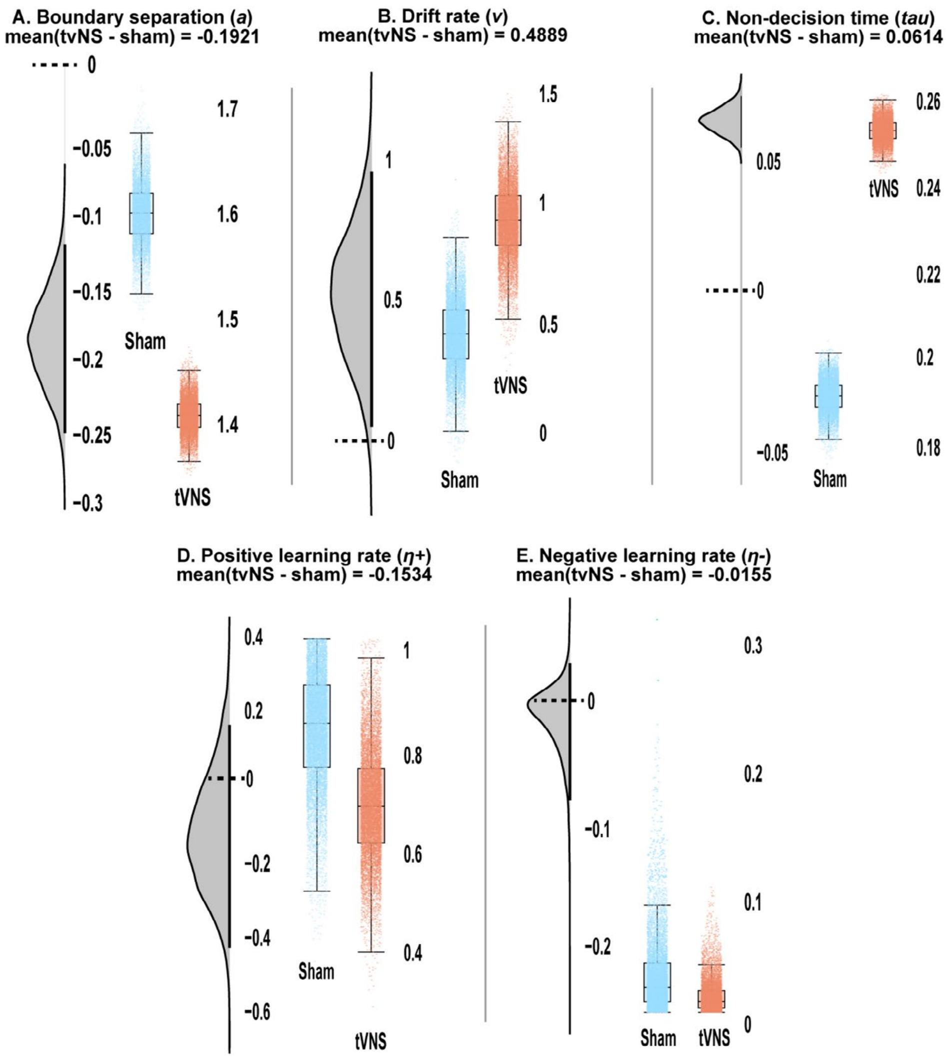

FIGURE 7 | **Posterior distributions of differences for tVNS and sham groups**. (A) boundary separation, (B) drift rate, (C) nondecision time, (D) positive learning rate and (E) negative learning rate. Thick horizontal bars below the distributions represent 95% high-density intervals. Additionally, boxplots illustrate the individual data points and group averages for each parameter, offering a detailed depiction of the distribution characteristics and central tendencies for the tVNS and sham groups.

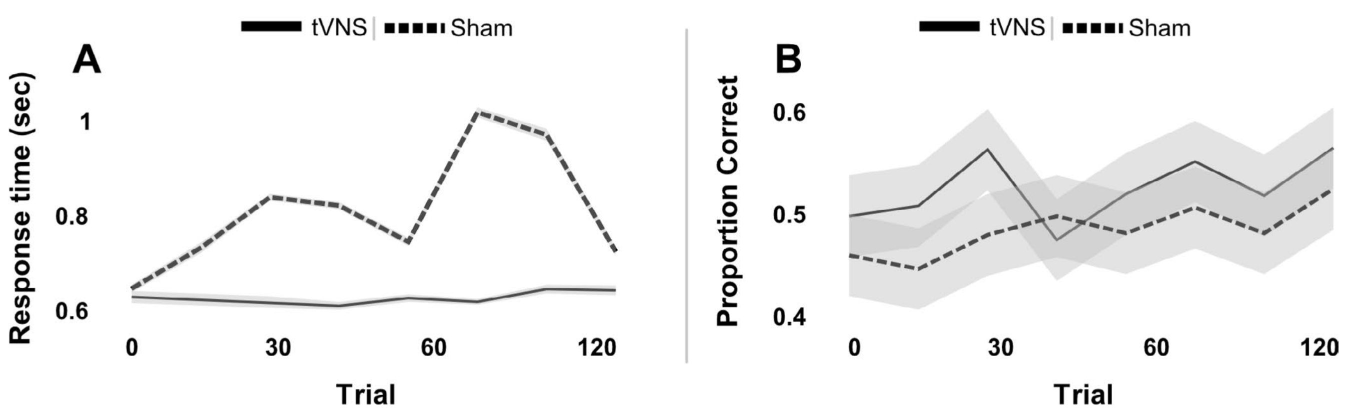

FIGURE 8 | Performance in the mPLT extinction phase. (A) Mean response time (B) Proportion correct. Graphs show estimated marginal means ±95% confidence interval (CI). The groups are indicated with different line types: dashed line—sham/ear lobe and solid line—tVNS.

and for the sham group, it starts at ~47% and increases to ~50% (see Table S9 for descriptive statistics). The average response time in the extinction phase for the tVNS group starts at ~620 msec, while the sham group starts at ~750 msec. The tVNS group response time between blocks does not change considerably, whereas the sham group response time increases to ~850 msec.

To test whether tVNS showed greater perseverance of learned behaviour in the absence of reward compared to sham, we conducted a GLMM regression with fixed factors 'group' (tVNS vs. sham) and 'block' (1–2) and a random effects factor 'participant' on the dependent variables of participants' accuracy and we conducted an LMM with fixed factors 'group' (tVNS vs. sham) and 'block' (1–2) and random effects factor 'participant' on the dependent variables of participants' response times.

For response accuracy, we observed a main effect of group (tVNS/sham) [$X^2(1) = 7.927$, $p = 0.005$], such that the tVNS group had a higher probability of being accurate regardless of block. We also observed a main effect of block [$X^2(1) = 27.842$, $p = 1.316 \times 10^{-7}$], which reflects learning of the stimulus reward contingency. There was no significant interaction between group and block [$X^2(1) = 1.995 \times 10^{-5}$, $p = 0.996$]. We examined estimated marginal means contrasts (Holm adjusted for two comparisons) to test which blocks differed between groups and we observed that both Blocks 1 and 2 were significant (Table S9). We refer the interested reader to the Supplementary Analyses for one-sample t-tests against chance, which show that the tVNS group's accuracy exceeded chance levels in Block 2 of the extinction phase, while for the sham group accuracy was lower than chance in Block 1 and not different from chance in Block 2 of the extinction phase.

For response time, we observed a main effect of group (tVNS/sham) [F (1,78.00) = 264.156, $p = 9.2501 \times 10^{-27}$], such that the tVNS group had a lower response time regardless of block. We also observed a main effect of block [F (1,78.00) =30.995, $p = 3.5280 \times 10^{-7}$] and an interaction of group and block [F (1,78.00) = 18.474, $p = 4.9267 \times 10^{-5}$]. The estimated marginal means contrasts (Holm adjusted for multiple comparisons, and Satterwaite correction) showed that the response times were lower for each block in the tVNS group (Table S9). To summarise, in a phase without feedback and reward in the mPLT, we observed that under continued tVNS, the learned information is maintained and performance remains higher than under sham.

3.4 | Control Analyses: No Cognitive and Demographic Differences Between Groups

To exclude differences between groups in baseline cognitive function, we analysed the performance of participants on two tasks measuring the cognitive functions of cognitive flexibility (WCST perseverative errors) and inhibition ability (Flanker test mean accuracy). Bayesian independent samples t-tests revealed that WCST perseverative errors and Flanker Test mean accuracy did not differ significantly between groups (Table S11). For WCST perseverative errors, there is weak-to-moderate evidence ($\log BF_{10} = -1.141$) in favour of the null hypothesis (that they are similar) relative to the alternative hypothesis (that they are not similar). For Flanker mean accuracy, there is weak to moderate evidence ($\log BF_{10} = -0.9974$) in favour of the null hypothesis (that they are similar) relative to the alternative hypothesis (that they are not similar). These results confirm that the groups did not differ in their cognitive flexibility and response inhibition ability and consequently, these variables were not included as covariates in any additional analyses.

To exclude differences between groups in demographic variables that may influence cognitive function, we analysed the age and gender distributions of participants. A Bayesian independent samples t-test (Table S12) revealed that for age there is weak to moderate evidence in favour of the null hypothesis (that they are similar between groups) relative to the alternative hypothesis (that they are not similar) ($\log BF_{10} = -1.071$). A Bayesian contingency table analysis (Table S13) also shows that for gender there is weak evidence ($\log BF_{10} = -0.923$) in favour of the null hypothesis (that they are similar) relative to the alternative hypothesis (that they are not similar). These results confirm that the groups did not differ in their age and gender and consequently, these variables were not included as covariates in any additional analyses.

3.5 | Control Analyses: No Preexisting Associations With the Symbols in a Pilot Study

To ensure that symbols did not evoke any conscious associations with participants, we analysed the association judgments in an independent group of pilot participants (Table S14). Bayesian one-sample binomial test results revealed that 'no' responses provided evidence in favour of the alternative hypothesis (that occur more frequently than expected by chance) relative to the null hypothesis (that occurs equally frequently than chance) ($\log BF_{10} = 1.905$–46.498) for all symbols except for 'visual 1' for which we observed only weak evidence.

3.6 | Control Analyses: No Differences Between Groups for Associations With the Symbols and Difficulty Biases

To assess awareness of the mPLT task, we asked various questions about the symbols after the experiment. When we asked participants whether they had any association with each of the symbols separately, a Bayesian one-sample binomial test (Table S15) revealed that 'no' choices of the participants for each symbol provided weak to very strong evidence in favour of the alternative hypothesis (that they are not similar) relative to the null hypothesis (that they are similar) ($\log BF_{10} = 0.271$–2.964) These results suggest that there were no significant explicit associations with the symbols.

Next, participants were asked if all pairs seemed equally easy to respond to. The Bayesian contingency table analysis revealed that a few more people in the tVNS group felt that all pairs were equally easy to respond to, while in the sham group, more people indicated that some pairs were more difficult to respond to. However, a Bayesian contingency table analysis (Table S16) provides strong evidence in favour of the null hypothesis (that the groups are similar) relative to the alternative hypothesis (that they are not similar) ($\log BF_{10} = -0.4153$).

In addition, when we analysed the 'difficult pairs' in the participants who stated to the previous question that they experienced difficulty choosing, a Bayesian contingency tables analysis (Table S17) provides strong evidence in favour of the null hypothesis (that the groups find the same pairs most difficult) relative to the alternative hypothesis (that the groups find different pairs most difficult), which indicates there was no significant bias for or against a specific pair, nor a difference between groups ($\log BF_{10} = -2.7096$).

4 | Discussion

The objective of this study was to examine the influence of transcutaneous vagus nerve stimulation (tVNS) on the process of probabilistic reinforcement learning. Our findings indicate that tVNS significantly enhances overall accuracy and decreases response times in a simple probabilistic learning task compared to sham stimulation. In the extinction phase of the experiment, when rewards and feedback to the participants were omitted but stimulation was continued, we observed a continued enhanced performance for the tVNS group relative to the sham group. This is consistent with tVNS enhancing cognitive processing (Giraudier et al. 2020) and modulating reward-related (Han et al. 2018) neural circuits in the brain, potentially through the enhancement of dopaminergic and noradrenergic signalling pathways.

4.1 | Improved Accuracy With a Low Reward and an Easy Task

We operationalised reinforcement learning with a simple variation of the probabilistic learning task (Frank et al. 2004). Our study design was primarily based on the study by Weber et al. (2021). Since they showed an absence of performance greater than chance in the sham condition with three different contingency levels, and they also showed that learning in the tVNS condition occurred only after 123 trials, we wanted to create an easier-to-learn version of the probability learning task. We also wanted to ensure that we could observe any learning (performance greater than chance) under sham stimulation (unlike Weber et al.). Therefore, we chose a relatively high reward probability and constant reward probability across the three pairs. As intended, we observed an increase in performance, that is, evidence of learning in both groups. But the reader should note that the performance is relatively low, that is, accuracy levels out at about 60% in the tVNS group and about 55% in the sham group (compared to accuracy levels of around 75% and 60% in tVNS and sham respectively by Weber et al. and around 80% for tVNS and around 85% for sham by Kühnel et al. (2020)). A lower accuracy rate than Weber et al. which is still significantly different from chance is not surprising, as our study has a higher sample size. Is the relatively low accuracy rate we reported exceptional? To our awareness, no reference or baseline values exist on what constitutes 'normal' performance and what constitutes a deviation from normal performance in probabilistic learning. However, in the practice phase of the transitive inference tasks, which is relatively similar to what we use here (but transitive inference tasks usually include less deterministic pairs making the task more challenging), participants should usually reach 67%–80% accuracy to proceed in the study and most participants master this task without problems within 150–200 trials, but it should also be noted that participants usually have more time to respond in the training phase. In the current study, both groups started around chance level. The tVNS group improved significantly as well as the sham group, but performance in both groups is generally lower than observed for the same contingency level in the classical probability learning task, where the performance per criterion reaches 67%–80% across contingency levels. It is possible that the relatively small reward and punishment values in the current experiment may have contributed to a relatively low overall performance. While the low-value outcome may have contributed to lower overall performance, a low reward may also be one of the contributing factors for tVNS to boost learning, as previous studies have shown that tVNS increases willingness to work for less wanted rewards (Neuser et al. 2020), liking for less liked puddings (Öztürk et al. 2020) and liking for less liked food images in depression (Koepp et al. 2021). This suggests that tVNS is more effective in the context of low reward values. Future studies may intentionally manipulate outcome values to test this hypothesis. We used monetary rewards here, but results from Neuser et al. (2020) and Ferstl et al. (2022) suggest that primary food rewards may boost willingness to work more, so future studies may also explore using the probabilistic learning tasks with food rewards.

4.2 | A Unified Account Across Reinforcement Learning Studies: Normalisation?

Despite relatively low accuracy rates, response times in our study are relatively fast and show more pronounced effects of tVNS, while the two prior studies did not observe any effects on response times. Besides a presumably easier learning task, we note the following differences between our study and the previous studies. Both previous studies used a within-participants cross-over design, and in both studies, the reinforcement learning task came after prolonged stimulation and performance of multiple other tasks (~3 h in Weber et al. and ~1 h in Kühnel et al. 2020). In our case, stimulation was started only at the start of the mPLT task, and this novel sensation may have served as a slight distractor leading to low (but greater than chance) performance overall. It is also possible that, due to the low absolute reward value in our study (1 tL at the time of the experiment was about $0.04), participants may have had relatively low motivation, leading to relatively low performance. Another comparable study, Kühnel et al. (2020), also used a small reward (5ct per correct response), but there were various other methodological differences between our task and theirs, which are that they had a more difficult task and that they used a go–no go paradigm. An additional explanation may be that in 2023 the average inflation rate in Türkiye was at around 54% compared to the previous year, and 73% in 2022. This means that participants may have generally had a mindset that the reward value is decreasing in real time. An acute sense of currency depreciation may affect motivation differently relative to a sense of a stable currency, which is certainly in line with observations that motivation is lower when reward devalues due to state changes such as fullness vs. hunger. If our speculation of a combination of a low reward value and a relatively easier task explains the relatively low accuracy rate, then in a similar but more difficult task with a higher reward, we may observe improved performance. Weber et al. used a much more comparable task to ours compared to Kühnel et al. (2020). They used 20 c rewards and three different levels of reward contingencies and observed ~80% accuracy after learning in the tVNS group. To better interpret our results, a future study may directly assess whether increased difficulty and increased reward, separately and combined, lead to changes in accuracy rates.

Why do we observe relatively low absolute performance and then relatively large effects of tVNS? It is possible that the relatively large effect of tVNS on learning in our study is related to the proximity in time to the start of stimulation, that is, with prolonged stimulation, there may be habituation to stimulation. And how is it possible that Kühnel et al. (2020) observed reductions in reinforcement learning, while we and others (Weber et al. 2021) observe greater improvements in learning? Previously, Koepp et al. (2021) observed that tVNS reduced variance in liking ratings across participants, such that tVNS increased relatively low liking ratings and reduced relatively high liking ratings. Koepp et al. speculated that tVNS may have a 'normalizing' effect and that net effects of tVNS may depend on different states, traits and/or different populations of neurons actuated in the NTS. This proposal may explain the contradictory results across studies of tVNS effects on reinforcement learning: under relatively low performance, tVNS may improve learning, while under relatively high performance, tVNS may reduce learning. Future studies may test this hypothesis directly.

We also note that performance in our study reached stable levels relatively quickly, after about 24 trials for both groups. This means that future studies using a similar design as ours can consider a much shorter version with improved efficiency, with a proposed maximum of 120 trials.

4.3 | A Computational Model of Reinforcement Learning Suggests Learning Reflects Mostly Intuitive Responding and Greater Reward Sensitivity

To gain more insight into the mechanisms of the improved performance of the tVNS groups, we fitted an RLDDM model to our data. Such models can give an indication of whether improved learning is the result of different speed–accuracy strategies, increased reward sensitivity, improved stimulus encoding and/or response preparation and increased learning from greater-than-expected rewards or from less-than-expected rewards. From the estimated RLDDM parameters, we observed that the tVNS group benefited from a riskier or impulsive (relatively fast but also possibly more error prone) style, dedicating more time to stimulus encoding and motor processes and having greater reward sensitivity relative to the sham group. We did not observe differences between tVNS and sham stimulation on the two types of reward learning rates: the positive and negative learning rates. This indicates that in our study, tVNS improves overall learning by sensory and attentional processes (nondecision time), by reward processes (drift rate) and by shifting internal decision thresholds (boundary separation), but not by a bias towards a positive or negative error signal. Drift rate and boundary separation are often correlated but can also be independently manipulated. For example, Voss et al. (2004) observed a higher boundary separation when participants were more motivated to achieve higher accuracy, but no change in drift rate. Conversely, when the task was made more difficult, drift rate increased, but not boundary separation (Voss et al. 2004). This means that risky strategies (lower boundary separation) can be compatible with a high drift rate in easy-to-perform tasks in which participants are encouraged to make faster decisions. In our case, a more risky strategy (lower boundary separation) can be compatible with greater reward sensitivity if the more risky strategy primarily leads to faster responding and not to more error-prone responding, which is consistent with the faster response times that we observed for the tVNS group. We propose that this strategy would be better described as 'intuitive' rather than 'risky' because there is faster decision making in the absence of increased error rates (relative to the sham group).

How do our RLDMM model parameters compare to those in other studies of tVNS during reinforcement learning? Weber et al. (2021), similar to our findings, observed a higher drift rate and longer nondecision time under tVNS compared to sham. Their model did not include estimates of positive and negative learning rates. Unlike us, they also observed a higher boundary separation for tVNS, indicating a shift towards more conservative responding. Kühnel et al. (2020) observed that tVNS affected the following parameters from an RL model: tVNS reduced the learning rate, increased the noisiness of choices and decreased learning from punishment. To interpret differences in the observation of these computational models is not straightforward as the model structures differ in various regards, and future studies should work towards comparing and/or finding consensus on the types of models that are appropriate for reinforcement learning studies. However, we may speculate that employing different speed–accuracy trade-off strategies may modulate the effects of tVNS on learning. In Kühnel et al. (2020) participants had 1000 msec to respond in Go trials. In Weber et al. (2021) the time to respond is not explicitly reported, but it is described that the task was 'self-paced'. In the current study reported here, participants had 2000 msec for their forced choice. It is possible that participants in the study by Kühnel et al. (2020) could not easily have shifted towards faster responding since they were perhaps already responding at their fastest (supported by a lack of effect on response times), and a shift of resources towards a faster strategy induced by tVNS may have led to more errors. Conversely, if participants have a longer time to respond, as in our study, faster response strategies may be successful without sacrificing accuracy. In the study by Weber et al. (2021), response times may be relatively slow, for example, up to 3.49 on average for one participant. Such slow response times may indicate that the participants biased their strategies very conservatively, which is supported by the increased boundary separation in tVNS. Such different speed–accuracy strategies may be related to different learning mechanisms in the brain, such as slow striatal learning and fast amygdala learning (Averbeck and Costa 2017). Future studies may directly assess the effect of speed–accuracy trade-off manipulations of tVNS-modulated reinforcement learning on neural responses and behaviour.

The effect of tVNS on improved learning or the RLDDM parameters does not necessarily reflect changes in one neurotransmitter pathway versus another; that is to say, the mechanism by which tVNS contributes to learning may stem from dopamine, noradrenaline or serotonin pathways. In addition, since we observed no differential effects on computational estimates of learning from positive vs. negative outcomes, we cannot speculate about different dopamine receptor contributions. Future studies could test the effect of tVNS with the classical probabilistic learning task, including the transitive inference phase, to gain more insights into the effects of tVNS on learning from positive (D1 receptor signalling) versus negative outcomes (D2 receptor signalling) (Cohen and Frank 2009), but regardless of the task used, it should be noted that tVNS effects are relatively unspecific.

4.4 | Continued Stimulation After Learning: Continued Enhanced Performance

After the learning phase of the reinforcement learning task, participants completed an additional 120 trials with the same forced-choice task, the 'extinction' phase. In this task, the same symbol pairs were presented, and continued tVNS stimulation was delivered as in the learning phase, but with the crucial difference that correct responses were no longer rewarded and feedback on performance was no longer given. The goal of the inclusion of this phase in the study was to assess the effect of continued vagus nerve stimulation on performance without reinforcement (modelled after Weber et al. 2021). Both the tVNS group and the sham group start this phase with decreased performance relative to the end of the learning phase. Both groups increase their performance slightly over time during the extinction phase. The tVNS group consistently shows greater accuracy and faster response times relative to the sham group in this phase, and during Block 2, the tVNS group's performance is greater than chance. During this phase, response times become slower over time too. The greater performance of the tVNS group is probably reflective of persistence of greater learning from the learning phase. Weber et al. also observed greater accuracy for the tVNS group relative to the sham group in the extinction phase. Why does accuracy not decline over time during the extinction phase, and why is even greater than chance accuracy achieved in the tVNS group? We speculate that the participants in this new mPLT phase without feedback may have initially taken a strategy of making different choices from the mPLT learning phase, but then later—in the absence of feedback—reverted back to their old response pattern to minimise cognitive fatigue. The relatively high accuracy during this phase indicates relatively stable levels of learning, which may have depended on the continued delivery of the tVNS. Extinction of conditioned fear in humans can be increased by tVNS immediately (Burger et al. 2016) or after a delay (Genheimer et al. 2017). Continued tVNS stimulation during this extinction phase may reinforce the acquired associations (a continuation of the effects in the learning phase), facilitate extinctions (like conditioned fear extinction studies showed) or a combination of both. As contributions from these two factors cannot be disentangled in the current design, we recommend that future studies use an extinction phase without tVNS stimulation. Future studies may also test extinction on subsequent test days to gain more insight into the effects of tVNS on probability learning persistence.

4.5 | tVNS Stimulation Locations

Aside from differences in task design, our study also differs from previous probabilistic reinforcement studies in our choice of tVNS stimulation locations. We used bilateral tVNS in both the tragus and cymba conchae. Neuser et al. (2020) previously observed greater motivation to obtain rewards, which was driven more by left-sided stimulation for food rewards (but also see Lucchi et al. 2024), but Ferstl et al. (2022) showed similar effects with only right-sided stimulation. While there may be laterality differences for VN signals related to diverging motivational circuits (Grimm and See 2000; Han et al. 2018; Neuser et al. 2020;), tVNS effects in both ears seem to affect probability learning, and using both ears for stimulation may be the most effective approach. Similarly, various studies have investigated tragus vs. cymba conchae tVNS effectiveness. Some studies prefer using cymba conchae over tragus because contributions from other auricular nerves can be excluded (Burger and Verkuil 2018; Peuker and Filler (2002) but also see (Badran et al. 2019, 2018). However, an agnostic approach that stimulates both the tragus and cymba conchae (and inclusion of an appropriate active sham condition), as we used here, may be most effective.

4.6 | Limitations

Our study has several limitations. First, we did not control for internal state in this study. Various other studies have shown that internal state or food consumption may influence the effects of tVNS (Altınkaya et al. 2023; Kozorosky et al. 2022). It is conceivable that variation in fasting time could contribute to noise in the study, although it may not be systematic variation. Since we worked with secondary (monetary) rewards, not food rewards, internal state may not be a confounding variable here. However, we recommend that future studies looking at probability learning should measure internal state or actively control it by asking participants to fast for a specific amount of time.

Second, we did not measure awareness of the relative contingency in this experiment. However, our postexperiment questionnaire indicates that participants in the sham group felt less confident about choosing certain pairs (although all pairs had a similar contingency), which may indicate a relatively higher awareness in the tVNS group. Awareness of the contingencies (Frank et al. 2005) may indicate that the intuitive preference that results from probability learning may be subject to more explicit logical reasoning; however, our overall low accuracy rates argue against conscious awareness. Future studies should include more in-depth measures of reward expectancies, parallel to Burger et al. (2016).

Third, we did not assess blinding awareness in the participants, which is an important improvement for future studies to include, for example, by employing the Bang blinding index (Bang et al. 2004) or a custom tVNS awareness questionnaire (Altınkaya et al. 2023).

Fourth, our experiment was unnecessarily inefficient, since the learning effects already occurred within the first block of the learning phase and the advantage for the tVNS groups was present in the subsequent blocks. We recommend that future studies (assuming all other parameters stay the same) use at most 120 trials to capture the most important changes in performance, yet still have enough trials available to fit computational models.

Lastly, we reported effects of tVNS stimulation in each block separately with custom post hoc tests because we had a specific interest in when (i.e., in which block) differences in learning as a function of tVNS occur, an important consideration since the study (in hind-sight) did not need as many as 240 trials and the post hoc tests per block show this. Normally, only significant interaction between block and group would justify the inspection of such custom tests, but this interaction was not present for all dependent variables. We would like to warn the reader to be cautious about the interpretation of these significant post hoc tests in the absence of a significant interaction of block and group, for example, in the case of response accuracy in the learning phase of the mPLT.

5 | Conclusions

To summarise, we show that tVNS enhances choice accuracy and decreases response times in a probabilistic learning task in healthy individuals. We also observed that this learning persists in the absence of rewards in a subsequent phase of the task. Reinforcement learning computational modelling suggests that the learning enhancement may be the result of greater reward sensitivity, more impulsive decisions and more time dedicated to sensory and attentional processes. As tVNS is relatively low-cost and easily applicable, this intervention may be explored for future use in enhancing new habit formation or faster learning of simple tasks.

Author Contributions

Resul Çakır: conceptualization, data curation, investigation, methodology, project administration, software, visualization, writing – original draft, writing – review and editing. **İlkim Büyükgüdük:** investigation, writing – review and editing. **Petek Bilim:** investigation, writing – review and editing. **Ataberk Erdinç:** investigation, software, visualization, writing – review and editing. **Maria Geraldine Veldhuizen:** conceptualization, supervision, writing – review and editing.

Ethics Statement

The study was conducted in accordance with the permission obtained from the ethics committee of Toros University (2023/90). All participants completed an informed voluntary consent form.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data of this study will be made publicly available in the Open Science Framework at https://osf.io/w5vtu.

References

Aguiar, I. D., and E. Kaan. 2019. "Statistical Language Learning and Vagus Nerve Stimulation: A Pilot Study." *UF Journal of Undergraduate Research* 21, no. 1: 107919. https://doi.org/10.32473/ufjur.v21i1.107919.

Ahn, W. Y., N. Haines, and L. Zhang. 2017. "Revealing Neurocomputational Mechanisms of Reinforcement Learning and Decision-Making With the hBayesDM Package." *Computational Psychiatry* 1: 24–57. https://doi.org/10.1162/CPSY_a_00002.

Altınkaya, Z., L. Öztürk, İ. Büyükgüdük, et al. 2023. "Non-invasive Vagus Nerve Stimulation in a Hungry State Decreases Heart Rate Variability." *Physiology & Behavior* 258: 114016. https://doi.org/10.1016/j.physbeh.2022.114016.

Averbeck, B. B., and V. D. Costa. 2017. "Motivational Neural Circuits Underlying Reinforcement Learning." *Nature Neuroscience* 20, no. 4: 505–512. https://doi.org/10.1038/nn.4506.

Badran, B. W., B. Y. Alfred, D. Adair, et al. 2019. "Laboratory Administration of Transcutaneous Auricular Vagus Nerve Stimulation (taVNS): Technique, Targeting, and Considerations." *Journal of Visualized Experiments* 2019, no. 143: e58984. https://doi.org/10.3791/58984.

Badran, B. W., O. J. Mithoefer, C. E. Summer, et al. 2018. "Short Trains of Transcutaneous Auricular Vagus Nerve Stimulation (taVNS) Have Parameter-Specific Effects on Heart Rate." *Brain Stimulation* 11, no. 4: 699–708. https://doi.org/10.1016/j.brs.2018.04.004.

Bang, H., L. Ni, and C. E. Davis. 2004. "Assessment of Blinding in Clinical Trials." *Controlled Clinical Trials* 25, no. 2: 143–156. https://doi.org/10.1016/j.cct.2003.10.016.

Berthoud, H. R. 2008a. "Vagal and Hormonal Gut–Brain Communication: From Satiation to Satisfaction." *Neurogastroenterology and Motility* 20: 64–72. https://doi.org/10.1111/j.1365-2982.2008.01104.x.

Berthoud, H.-R. 2008b. "The Vagus Nerve, Food Intake and Obesity." *Regulatory Peptides* 149, no. 1: 15–25. https://doi.org/10.1016/j.regpep.2007.08.024.

Burger, A. M., and B. Verkuil. 2018. "Transcutaneous Nerve Stimulation via the Tragus: Are We Really Stimulating the Vagus Nerve?" *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation* 11, no. 4: 945–946. https://doi.org/10.1016/j.brs.2018.03.018.

Burger, A. M., B. Verkuil, I. van Diest, W. der Does, J. F. Thayer, and J. F. Brosschot. 2016. "The Effects of Transcutaneous Vagus Nerve Stimulation on Conditioned Fear Extinction in Humans." *Neurobiology of Learning and Memory* 132: 49–56. https://doi.org/10.1016/j.nlm.2016.05.007.

Carruthers, S. P., C. T. Gurvich, D. Meyer, et al. 2019. "Exploring Heterogeneity on the Wisconsin Card Sorting Test in Schizophrenia Spectrum Disorders: A Cluster Analytical Investigation." *Journal of the International Neuropsychological Society* 25, no. 7: 750–760. https://doi.org/10.1017/S1355617719000420.

Chakroun, K., A. Wiehler, B. Wagner, et al. 2023. "Dopamine Regulates Decision Thresholds in Human Reinforcement Learning in Males." *Nature Communications* 14, no. 1: 5369. https://doi.org/10.1038/s41467-023-41130-y.

Chen, M., L. Yu, Q. Liu, et al. 2015. "Noninvasive Vagus Nerve Stimulation: A Novel Promising Modulator for Cardiac Autonomic Nerve System Dysfunction." *International Journal of Cardiology* 187: 338–339. https://doi.org/10.1016/j.ijcard.2015.03.317.

Cohen, M. X., and M. J. Frank. 2009. "Neurocomputational Models of Basal Ganglia Function in Learning, Memory and Choice." *Behavioural Brain Research* 199, no. 1: 141–156. https://doi.org/10.1016/j.bbr.2008.09.029.

Colzato, L. S., S. M. Ritter, and L. Steenbergen. 2018. "Transcutaneous Vagus Nerve Stimulation (tVNS) Enhances Divergent Thinking." *Neuropsychologia* 111: 72–76. https://doi.org/10.1016/j.neuropsychologia.2018.01.003.

Çuhadaroğlu, A. 2016. "Bilişsel esnekliğin yordayıcıları." *Cumhuriyet Uluslararası Eğitim Dergisi* 2, no. 1: 86–101.

Dalley, J. W., and B. J. Everitt. 2009. "Dopamine Receptors in the Learning, Memory and Drug Reward Circuitry." *Seminars in Cell & Developmental Biology* 20, no. 4: 403–410. https://doi.org/10.1016/j.semcdb.2009.01.002.

Ellrich, J. 2011. "Transcutaneous Vagus Nerve Stimulation." *European Neurological Review* 6, no. 4: 262–264. https://doi.org/10.17925/ENR.2011.06.04.254.

Farmer, A. D., A. Strzelczyk, A. Finisguerra, et al. 2021. "International Consensus Based Review and Recommendations for Minimum Reporting Standards in Research on Transcutaneous Vagus Nerve Stimulation (Version 2020)." *Frontiers in Human Neuroscience* 14: 568051. https://doi.org/10.3389/fnhum.2020.568051.

Faul, F., E. Erdfelder, A.-G. Lang, and A. Buchner. 2007. "G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences." *Behavior Research Methods* 39: 175–191. https://doi.org/10.3758/BF03193146.

Fernandes, A. B., J. Da Alves Silva, J. Almeida, et al. 2020. "Postingestive Modulation of Food Seeking Depends on Vagus-Mediated Dopamine Neuron Activity." *Neuron* 106, no. 5: 778–788. https://doi.org/10.1016/j.neuron.2020.03.009.

Ferstl, M., V. Teckentrup, W. M. Lin, et al. 2022. "Non-invasive Vagus Nerve Stimulation Boosts Mood Recovery After Effort Exertion." *Psychological Medicine* 52, no. 14: 3029–3039. https://doi.org/10.1017/S0033291720005073.

Frangos, E., and B. R. Komisaruk. 2017. "Access to Vagal Projections via Cutaneous Electrical Stimulation of the Neck: fMRI Evidence in Healthy Humans." *Brain Stimulation* 10, no. 1: 19–27. https://doi.org/10.1016/j.brs.2016.10.008.

Frank, M. J., and L. Kong. 2008. "Learning to Avoid in Older Age." *Psychology and Aging* 23, no. 2: 392–398. https://doi.org/10.1037/0882-7974.23.2.392.

Frank, M. J., J. W. Rudy, W. B. Levy, and R. C. O'Reilly. 2005. "When Logic Fails: Implicit Transitive Inference in Humans." *Memory & Cognition* 33, no. 4: 742–750. https://doi.org/10.3758/BF03195340.

Frank, M. J., L. C. Seeberger, and R. C. O'reilly. 2004. "By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism." *Science* 306, no. 5703: 1940–1943. https://doi.org/10.1126/science.1102941.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2013. *Bayesian data analysis*. 3rd ed. CRC Press.

Gelman, A., and J. Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Gelman, A., J. Hwang, and A. Vehtari. 2014. "Understanding Predictive Information Criteria for Bayesian Models." *Statistics and Computing* 24: 997–1016. https://doi.org/10.1007/s11222-013-9416-2.

Gelman, A., X. L. Meng, and H. Stern. 1996. "Posterior Predictive Assessment of Model Fitness via Realized Discrepancies." *Statistica Sinica* 6: 733–760.

Gelman, A., and D. B. Rubin. 1992. "Inference From Iterative Simulation Using Multiple Sequences." *Statistical Science* 7, no. 4: 457–472. https://doi.org/10.1214/ss/1177011136.

Genheimer, H., M. Andreatta, E. Asan, and P. Pauli. 2017. "Reinstatement of Contextual Conditioned Anxiety in Virtual Reality and the Effects of Transcutaneous Vagus Nerve Stimulation in Humans." *Scientific Reports* 7, no. 1: 17886. https://doi.org/10.1038/s41598-017-18183-3.

Giraudier, M., C. Ventura-Bort, and M. Weymar. 2020. "Transcutaneous Vagus Nerve Stimulation (tVNS) Improves High-Confidence Recognition Memory but Not Emotional Word Processing." *Frontiers in Psychology* 11: 541855. https://doi.org/10.1017/S0033291720005073.

Grimm, J. W., and R. E. See. 2000. "Dissociation of Primary and Secondary Reward-Relevant Limbic Nuclei in an Animal Model of Relapse." *Neuropsychopharmacology* 22, no. 5: 473–479. https://doi.org/10.1016/S0893-133X(99)00157-8.

Han, W., L. A. Tellez, M. H. Perkins, et al. 2018. "A Neural Circuit for Gut-ınduced Reward." *Cell* 175, no. 3: 665–678. https://doi.org/10.1016/j.cell.2018.08.049.

Heaton, R. K. 1993. *Wisconsin Card Sorting Test: Computer Version-2*. Psychological Assessment Resources.

Jacobs, H. I., J. M. Riphagen, C. M. Razat, S. Wiese, and A. T. Sack. 2015. "Transcutaneous Vagus Nerve Stimulation Boosts Associative Memory in Older Individuals." *Neurobiology of Aging* 36, no. 5: 1860–1867. https://doi.org/10.1016/j.neurobiolaging.2015.02.023.

Jandackova, V., V. Vasendova, M. Jackowska, and J. Koenig. 2023. "The Effect of Long-Term Non-invasive Vagus Nerve Stimulation on Cognitive Performance: Results From a Randomized Placebo Controlled Trial." *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation* 16, no. 2: 6. https://doi.org/10.1016/j.brs.2023.03.028.

JASP Team. 2023. "JASP (Version 0.17.1) [Computer Software]." https://jasp-stats.org/.

Jeffreys, H. 1998. *The Theory of Probability*. OUP Oxford.

Jongkees, B. J., M. A. Immink, A. Finisguerra, and L. S. Colzato. 2018. "Transcutaneous Vagus Nerve Stimulation (tVNS) Enhances Response Selection During Sequential Action." *Frontiers in Psychology* 9: 1159. https://doi.org/10.3389/fpsyg.2018.01159.

Kaan, E., I. de Aguiar, C. Clarke, D. G. Lamb, J. B. Williamson, and E. C. Porges. 2021. "A Transcutaneous Vagus Nerve Stimulation Study on Verbal Order Memory." *Journal of Neurolinguistics* 59: 100990. https://doi.org/10.1016/j.jneuroling.2021.100990.

Kaniusas, E., S. Kampusch, M. Tittgemeyer, et al. 2019. "Current Directions in the Auricular Vagus Nerve Stimulation I – A Physiological Perspective." *Frontiers in Neuroscience* 13: 854. https://doi.org/10.3389/fnins.2019.00854.

Kim, A. Y., A. Marduy, P. S. de Melo, et al. 2022. "Safety of Transcutaneous Auricular Vagus Nerve Stimulation (taVNS): A Systematic Review and Meta-Analysis." *Scientific Reports* 12, no. 1: 22055. https://doi.org/10.1038/s41598-022-25864-1.

Koepp, V., J. Klaus, M. Ferstl, F. K. Müller, A. Kühnel, and N. B. Kroemer. 2021. "Non-invasive Vagus Nerve Stimulation Normalizes Food Liking and Improves Liking Ratings in Depression." https://doi.org/10.31234/osf.io/3j5xy.

Körber, M., J. Radlmayr, and K. Bengler. 2016. "Bayesian highest density intervals of take-over times for highly automated driving in different traffic densities. In Proceedings of the Human Factors and Ergonomics Society annual meeting." *SAGE Publications* 60, no. 1: 2009–2013. https://doi.org/10.1177/1541931213601457.

Kozorosky, E. M., C. H. Lee, J. G. Lee, V. Nunez Martinez, L. E. Padayachee, and H. M. Stauss. 2022. "Transcutaneous Auricular Vagus Nerve Stimulation Augments Postprandial Inhibition of Ghrelin." *Physiological Reports* 10, no. 8: e15253. https://doi.org/10.14814/phy2.15253.

Kraus, T., O. Kiess, K. Hösl, P. Terekhin, J. Kornhuber, and C. Forster. 2013. "CNS BOLD fMRI Effects of Sham-Controlled Transcutaneous Electrical Nerve Stimulation in the Left Outer Auditory Canal – A Pilot Study." *Brain Stimulation* 6, no. 5: 798–804. https://doi.org/10.1016/j.brs.2013.01.011.

Kreuzer, P. M., M. Landgrebe, O. Husser, et al. 2012. "Transcutaneous Vagus Nerve Stimulation: Retrospective Assessment of Cardiac Safety in a Pilot Study." *Frontiers in Psychiatry* 3: 70. https://doi.org/10.3389/fpsyt.2012.00070.

Kruschke, J. 2014. *Doing Bayesian Data Analysis: A Tutorial With R, JAGS, and Stan*. Academic Press.

Kruschke, J. K. 2018. "Rejecting or Accepting Parameter Values in Bayesian Estimation." *Advances in Methods and Practices in Psychological Science* 1, no. 2: 270–280. https://doi.org/10.1177/2515245918771304.

Kühnel, A., V. Teckentrup, M. P. Neuser, et al. 2020. "Stimulation of the Vagus Nerve Reduces Learning in a Go/No-Go Reinforcement Learning Task." *European Neuropsychopharmacology* 35: 17–29. https://doi.org/10.1016/j.euroneuro.2020.03.023.

Kunisato, Y., Y. Okamoto, K. Ueda, et al. 2012. "Effects of Depression on Reward-Based Decision Making and Variability of Action in

Probabilistic Learning." *Journal of Behavior Therapy and Experimental Psychiatry* 43, no. 4: 1088–1094. https://doi.org/10.1016/j.jbtep.2012.05.007.

de Lartigue, G. 2016. "Role of the Vagus Nerve in the Development and Treatment of Diet-Induced Obesity." *Journal of Physiology* 594, no. 20: 5791–5815. https://doi.org/10.1113/JP271538.

Lerche, V., A. Voss, and M. Nagler. 2017. "How Many Trials Are Required for Parameter Estimation in Diffusion Modeling? A Comparison of Different Optimization Criteria." *Behavior Research Methods* 49: 513–537. https://doi.org/10.3758/s13428-016-0740-2.

Lucchi, F., B. Lloyd, and S. Nieuwenhuis. 2024. "Non-invasive Vagus Nerve Stimulation and the Motivation to Work for Rewards: A Replication of Neuser et al. (2020, Nature Communications)." *Psychophysiology* 61, no. 4: e14484. https://doi.org/10.1111/psyp.14484.

Manta, S., M. El Mansari, G. Debonnel, and P. Blier. 2013. "Electrophysiological and Neurochemical Effects of Long-Term Vagus Nerve Stimulation on the Rat Monoaminergic Systems." *International Journal of Neuropsychopharmacology* 16, no. 2: 459–470. https://doi.org/10.1017/S1461145712000387.

Marsman, M., and E. J. Wagenmakers. 2017. "Three Insights From a Bayesian Interpretation of the One-Sided P Value." *Educational and Psychological Measurement* 77, no. 3: 529–539. https://doi.org/10.1177/0013164416669201.

McIntire, L. K., R. A. McKinley, C. Goodyear, J. P. McIntire, and R. D. Brown. 2021. "Cervical Transcutaneous Vagal Nerve Stimulation (ctVNS) Improves Human Cognitive Performance Under Sleep Deprivation Stress." *Communications Biology* 4, no. 1: 1–9. https://doi.org/10.1038/s42003-021-02145-7.

Mueller, S. T., and B. J. Piper. 2014. "The Psychology Experiment Building Language (PEBL) and PEBL Test Battery." *Journal of Neuroscience Methods* 222: 250–259. https://doi.org/10.1016/j.jneumeth.2013.10.024.

Neuser, M. P., V. Teckentrup, A. Kühnel, M. Hallschmid, M. Walter, and N. B. Kroemer. 2020. "Vagus Nerve Stimulation Boosts the Drive to Work for Rewards." *Nature Communications* 11, no. 1: 3555. https://doi.org/10.1038/s41467-020-17344-9.

Öztürk, L., P. E. Büning, E. Frangos, G. de Lartigue, and M. G. Veldhuizen. 2020. "tVNS Increases Liking of Orally Sampled Low-Fat Foods: A Pilot Study." *Frontiers in Human Neuroscience* 14: 600995. https://doi.org/10.3389/fnhum.2020.600995.

Pedersen, M. L., and M. J. Frank. 2020. "Simultaneous Hierarchical Bayesian Parameter Estimation for Reinforcement Learning and Drift Diffusion Models: A Tutorial and Links to Neural Data." *Computational Brain & Behavior* 3, no. 4: 458–471. https://doi.org/10.1007/s42113-020-00084-w.

Pedersen, M. L., M. J. Frank, and G. Biele. 2017. "The Drift Diffusion Model as the Choice Rule in Reinforcement Learning." *Psychonomic Bulletin & Review* 24: 1234–1251. https://doi.org/10.3758/s13423-016-1199-y.

Peuker, E. T., and T. J. Filler. 2002. "The Nerve Supply of the Human Auricle." *Clinical Anatomy* 15, no. 1: 35–37. https://doi.org/10.1002/ca.1089.

Phillips, I., R. C. Calloway, V. P. Karuzis, N. B. Pandža, P. O'Rourke, and S. E. Kuchinsky. 2021. "Transcutaneous Auricular Vagus Nerve Stimulation Strengthens Semantic Representations of Foreign Language Tone Words During Initial Stages of Learning." *Journal of Cognitive Neuroscience* 34, no. 1: 127–152.

Python Software Foundation. 2019. "Python (Version 3.8) [Computer software]." https://www.python.org/.

R Core Team. 2022. *A language and environment for statistical computing (Version 4.2.2)*. R Foundation for Statistical Computing. https://www.R-project.org/.

Ratcliff, R., and R. Childers. 2015. "Individual Differences and Fitting Methods for the Two-Choice Diffusion Model of Decision Making." *Decision* 2, no. 4: 237–279. https://doi.org/10.1037/dec0000030.

Ratcliff, R., and F. Tuerlinckx. 2002. "Estimating Parameters of the Diffusion Model: Approaches to Dealing With Contaminant Reaction Times and Parameter Variability." *Psychonomic Bulletin & Review* 9, no. 3: 438–481. https://doi.org/10.1037/dec0000030.

Redgrave, J., D. Day, H. Leung, et al. 2018. "Safety and Tolerability of Transcutaneous Vagus Nerve Stimulation in Humans; A Systematic Review." *Brain Stimulation* 11, no. 6: 1225–1238. https://doi.org/10.1016/j.brs.2018.08.010.

Rescorla, R. A. 1972. "A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Non-reinforcement." *Classical Conditioning, Current Research and Theory* 2: 64–69.

Schultz, W. 2001. "Book Review: Reward Signaling by Dopamine Neurons." *Neuroscientist* 7, no. 4: 293–302. https://doi.org/10.1177/107385840100700406.

Seki, A., H. R. Green, T. D. Lee, et al. 2014. "Sympathetic Nerve Fibers in Human Cervical and Thoracic Vagus Nerves." *Heart Rhythm* 11, no. 8: 1411–1417. https://doi.org/10.1016/j.hrthm.2014.04.032.

Sellaro, R., B. de Gelder, A. Finisguerra, and L. S. Colzato. 2018. "Transcutaneous Vagus Nerve Stimulation (tVNS) Enhances Recognition of Emotions in Faces but Not Bodies." *Cortex* 99: 213–223. https://doi.org/10.1016/j.cortex.2017.11.007.

Sun, L., J. Peräkylä, K. Holm, et al. 2017. "Vagus Nerve Stimulation Improves Working Memory Performance." *Journal of Clinical and Experimental Neuropsychology* 39, no. 10: 954–964. https://doi.org/10.1080/13803395.2017.1285869.

Teckentrup, V., and N. B. Kroemer. 2023. "Mechanisms for Survival: Vagal Control of Goal-Directed Behavior." *Trends in Cognitive Sciences* 28, no. 3: 237–251. https://doi.org/10.1016/j.tics.2023.11.001.

Tellez, L. A., S. Medina, W. Han, et al. 2013. "A Gut Lipid Messenger Links Excess Dietary Fat to Dopamine Deficiency." *Science* 341, no. 6147: 800–802. https://doi.org/10.1126/science.1239275.

Vehtari, A., A. Gelman, and J. Gabry. 2017. "Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC." *Statistics and Computing* 27: 1413–1432. https://doi.org/10.1007/s11222-016-9696-4.

Verlinden, T. J., K. Rijkers, G. Hoogland, and A. Herrler. 2016. "Morphology of the Human Cervical Vagus Nerve: Implications for Vagus Nerve Stimulation Treatment." *Acta Neurologica Scandinavica* 133, no. 3: 173–182. https://doi.org/10.1111/ane.12462.

Voss, A., K. Rothermund, and J. Voss. 2004. "Interpreting the Parameters of the Diffusion Model: An Empirical Validation." *Memory & Cognition* 32, no. 7: 1206–1220. https://doi.org/10.3758/BF03196893.

Weber, I., H. Niehaus, K. Krause, et al. 2021. "Trust Your Gut: Vagal Nerve Stimulation in Humans Improves Reinforcement Learning." *Brain Communications* 3, no. 2: fcab039. https://doi.org/10.1093/braincomms/fcab039.

Wiecki, T. V., I. Sofer, and M. J. Frank. 2013. "HDDM: Hierarchical Bayesian Estimation of the Drift-Diffusion Model in Python." *Frontiers in Neuroinformatics* 7: 55610. https://doi.org/10.3389/fninf.2013.00014.

Wiehler, A., and J. Peters. 2024. "Decomposition of Reinforcement Learning Deficits in Disordered Gambling via Drift Diffusion Modeling and Functional Magnetic Resonance Imaging." *Computational Psychiatry* 8, no. 1: 23–45. https://doi.org/10.5334/cpsy.104.

Yakunina, N., S. S. Kim, and E.-C. Nam. 2017. "Optimization of Transcutaneous Vagus Nerve Stimulation Using Functional MRI." *Neuromodulation: Technology at the Neural Interface* 20, no. 3: 290–300. https://doi.org/10.1111/ner.12541.

Yang, M., X. Zhao, and R. R. Miselis. 1999. "The Origin of Catecholaminergic Nerve Fibers in the Subdiaphragmatic Vagus Nerve of Rat." *Journal of the Autonomic Nervous System* 76, no. 2: 108–117. https://doi.org/10.1016/S0165-1838(99)00014-4.

Yuan, H., and S. D. Silberstein. 2016. "Vagus Nerve and Vagus Nerve Stimulation, a Comprehensive Review: Part I." *Headache: The Journal of Head and Face Pain* 56, no. 1: 71–78. https://doi.org/10.1111/head.12647.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.