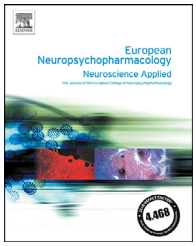
European Neuropsychopharmacology (2020) 35, 17–29

www.elsevier.com/locate/euroneuro



Stimulation of the vagus nerve reduces learning in a go/no-go reinforcement learning task

Anne Kühnel[a,*], **Vanessa Teckentrup**[b], **Monja P. Neuser**[b], **Quentin J.M. Huys**[c,d], **Caroline Burrasch**[b], **Martin Walter**[b,e,f,g], **Nils B. Kroemer**[b,*]

[a] *Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry and International Max Planck Research School for Translational Psychiatry (IMPRS-TP), Munich, Germany*
[b] *Department of Psychiatry and Psychotherapy, University of Tübingen, Calwerstr. 14, 72076 Tübingen, Germany*
[c] *Division of Psychiatry, University College London, London, United Kingdom*
[d] *Max Planck UCL Centre for Computational Psychiatry and Ageing Research, London, United Kingdom*
[e] *Otto-von-Guericke University Magdeburg, Department of Psychiatry and Psychotherapy, Magdeburg, Germany*
[f] *Department of Psychiatry and Psychotherapy, Jena University Hospital, Jena, Germany*
[g] *Leibniz Institute for Neurobiology, Magdeburg, Germany*

Received 22 September 2019; received in revised form 6 February 2020; accepted 27 March 2020

KEYWORDS
tVNS;
Reinforcement learning;
Computational modeling;
Metabolic state;
Instrumental action

Abstract
When facing decisions to approach rewards or to avoid punishments, we often figuratively go with our gut, and the impact of metabolic states such as hunger on motivation are well documented. However, whether and how vagal feedback signals from the gut influence instrumental actions is unknown. Here, we investigated the effect of non-invasive transcutaneous auricular vagus nerve stimulation (taVNS) vs. sham (randomized cross-over design) on approach and avoidance behavior using an established go/no-go reinforcement learning paradigm in 39 healthy human participants (23 female) after an overnight fast. First, mixed-effects logistic regression analysis of choice accuracy showed that taVNS acutely impaired decision-making, $p = .041$. Computational reinforcement learning models identified the cause of this as a reduction in the learning rate through taVNS ($\Delta\alpha = -0.092$, $p_{boot} = .002$), particularly after punishment ($\Delta\alpha_{Pun} = -0.081$, $p_{boot} = .012$ vs. $\Delta\alpha_{Rew} = -0.031$, $p_{boot} = .22$). However, taVNS

* Corresponding authors.
E-mail addresses: anne_kuehnel@psych.mpg.de (A. Kühnel), nils.kroemer@uni-tuebingen.de (N.B. Kroemer).

https://doi.org/10.1016/j.euroneuro.2020.03.023

had no effect on go biases, Pavlovian response biases or response time. Hence, taVNS appeared to influence learning rather than action execution. These results highlight a novel role of vagal afferent input in modulating reinforcement learning by tuning the learning rate according to homeostatic needs.

1. Introduction

To survive, organisms must procure energy by approaching options that pay off while avoiding costly options potentially incurring punishments. Fundamental learning mechanisms have evolved to support this vital optimization of instrumental actions (Beeler, 2012; Keramati and Gutkin, 2014; Korn and Bach, 2015; Kroemer and Small, 2016). One key challenge is to balance short-term and long-term goals of reward-related behavior. For example, patiently resisting temptation to receive bigger returns later is often beneficial in the longer term. However, in certain bodily states, forfeiting immediate rewards can have negative long-term consequences as missing out on food in a hungry state (Skrynka and Vincent, 2017) can increase the risk of starvation (Keramati and Gutkin, 2014). Despite its evolutionary importance, little is known about how homeostatic needs shape decision-making in humans. One plausible candidate for modulatory input onto circuits involved in reward learning would be a caloric feedback signal (Veldhuizen et al., 2017) originating from the gut.

Signals about bodily states are largely transmitted via the vagus nerve which connects peripheral organs such as the gut and heart with the brain. For example, the vagal nerve is an important part in the autonomous nervous system and responds to stress (Porges, 1995; Zagon, 2001; Lee et al., 2018; Gurel et al., 2020). Regarding metabolic state, circulating hormones such as cholecystokinin (CCK), GLP-1, and ghrelin provide feedback on food intake via stimulation of vagal afferents (Simonian et al., 2005; Dockray, 2009; Date, 2012; Tellez et al., 2013; Breit et al., 2018). Moreover, the vagus nerve also transmits other feedback signals from the gut and intestines such as stomach dilation and changes in microbiota (Waise et al., 2018). Notably, gut microbiota can also modulate neurotransmission contributing to interoceptive communication via the gut-brain axis (Strandwitz, 2018). Vagal afferents terminate in the nucleus tractus solitarii, NTS (Lartigue, 2016), a hub further relaying metabolic information to the mid- and forebrain (Grill and Hayes, 2012; Lartigue, 2016) including to dopaminergic neurons in the substantia nigra. Along that pathway, vagal afferents have been shown to modulate dopaminergic (Tellez et al., 2013; Han et al., 2018), noradrenergic (Roosevelt et al., 2006; Raedt et al., 2011), GABAergic (Ben-Menachem et al., 1995; Capone et al., 2015), and cholinergic signaling (Hulsey et al., 2016). Accordingly, endogenous stimulation of the gut with nutrients evokes dopamine responses in the dorsal striatum tracking energy (de Araujo et al., 2012; Ferreira et al., 2012). These dopamine signals are critical for appetitive conditioned learning (Davis et al., 2008; Tellez et al., 2013; de Lartigue et al., 2014) and motivated behavior (Palmiter, 2007, 2008). Additional cognitive functions such as memory (Peña et al., 2013; Suarez et al., 2018) that are highly relevant for reward seeking are affected by vagal signaling, but are primarily modulated by other neurotransmitter systems. Collectively, these results suggest that vagal signals may shape reward seeking according to bodily states, via alterations in multiple neurotransmitter systems including dopamine.

Whereas a dopaminergic modulation by vagal input has been conclusively shown in animals, research in humans has been limited by the invasive nature of cervical vagus nerve stimulation (VNS). Lately, non-invasive transcutaneous auricular VNS (taVNS) has become feasible targeting the auricular branch of the vagus nerve at the ear. This has been shown to affect projections to the NTS in preclinical studies (He et al., 2013). Studies using taVNS with concurrent fMRI have revealed enhanced activity in the NTS and other interconnected brain regions including the dopaminergic midbrain and nucleus accumbens (Kraus et al., 2013; Frangos et al., 2015). Previous studies showed behavioral effects of taVNS on memory retention (Jacobs et al., 2015; Burger et al., 2016), cognitive performance (Sellaro et al., 2015; Steenbergen et al., 2015), and response inhibition (Beste et al., 2016) that are predominantly associated with the noradrenergic or GABAergic (Quetscher et al., 2015) system. Recently, we have shown that taVNS elicits efferent effects on energy metabolism leading to a reduced gastric frequency (Teckentrup et al., 2020). Since vagal signals also modulate the dopaminergic system (Tellez et al., 2013; Han et al., 2018), taVNS may provide a promising approach to investigate the link between interoceptive signals transmitted via the vagus nerve and reward-related behavior in humans.

To test the effects of taVNS on reward learning, we applied taVNS (vs. sham) mimicking interoceptive signaling via vagal afferents in a sample of overnight fasted participants. Increases in dopamine tone would be expected to increase vigor (Niv et al., 2007) while learning via reward prediction errors (RPE) would be attenuated as the signal-to-noise ratio is reduced (Hamid et al., 2016; Kroemer et al., 2019). We probed reward learning with an established valence-dependent go/no-go learning paradigm (Guitart-Masip et al., 2012). To investigate which specific reinforcement learning process is altered by taVNS, we used computational modeling. In addition, we explored effects of taVNS on go response rates or response time, which would indicate heightened vigor.

2. Experimental procedure

2.1. Participants

In total, 44 individuals participated in the study. Initially, we estimated that about 40 participants would be necessary to assess

medium-sized effects (Cohen's $f = .20$, dz˜.40) with sufficiently high power (1-$\beta = .79$), given a moderate test-retest reliability of behavioral measures ($r_{12} = .60$), and tested more participants to account for dropouts or exclusion (10%). All participants were physically and mentally healthy, German speaking, and right-handed, as determined by a telephone interview. For the current analysis, five participants had to be excluded (n=4: did not complete both task sessions, n=1: did not make any go response). Thus, we included 39 participants (23 female, M_{age}= 25.5 ±4.0 years; M_{BMI}= 23.0 ±3.0 kg/m^2). The institutional review boards of the University of Tübingen approved the study and we obtained informed consent from all participants prior to taking part in the experiment.

2.2. Experimental procedure

Participants were required to fast overnight (i.e., >8h) before both experimental sessions. Sessions were conducted in a randomized, single-blind manner as the experimenter was not blind to the stimulation condition. Participants were close to chance in guessing the correct condition (60%; $p_{\text{binomial}} = .041$) and their subjective belief about which stimulation they received was not associated with the effects induced by the stimulation (all ps > .1) suggesting that blinding was effective. Sessions started between 7.00 am and 11:00 am and lasted about 2.5h each. After participants arrived for the first session, they provided written informed consent. Next, we collected anthropometric and state-related information before the taVNS electrode was placed on the left ear targeting the auricular branch of the vagus nerve. In line with the stimulation procedure by Frangos et al. (2015), the electrode was located at the left cymba conchae for taVNS and at the left earlobe for sham stimulation. Stripes of surgical tape served to secure the electrode in place. We determined individual stimulation strength for every session separately using pain VAS ratings ("How intensely do you feel pain induced by the stimulation?" ranging from 0 ("no sensation") to 10 ("strongest sensation imaginable")). Stimulation was initiated at an amplitude of 0.1 mA and increased by the experimenter in 0.1-0.2 mA steps. Participants rated the sensation after every increment until ratings settled around 5 corresponding to "mild prickling". Then, the stimulation continued throughout the task block according to the stimulation protocol of the device (i.e., alternating blocks of stimulation on and off for 30 s each). Within this block, participants completed a food-cue reactivity task (˜20 min) and an effort allocation task (˜40 min) before the learning task. As participants had received stimulation for ˜1 h before the learning task and stimulation-induced changes in brain activation last for minutes after turning stimulation off (Frangos et al., 2015), we did not expect differences due to the ON and OFF cycles of the stimulation protocol.

After completing state-related questions, participants received rewards according to task performance and compensation (either as 32€ or partial course credit). Both visits followed the same standardized protocol.

2.3. taVNS device

To stimulate the auricular branch of the vagus nerve, we used the *NEMOS*® stimulation device (cerbomed GmbH, Erlangen, Germany). These devices have been previously employed in clinical trials (Kreuzer et al., 2012; Bauer et al., 2016) and proof-of-principle neuroimaging studies (Frangos et al., 2015). The stimulation protocol for the *NEMOS*® is preset to a biphasic impulse frequency of 25 Hz with a stimulation duration of 30 s, followed by a 30 s stimulation pause. The electrical current is transmitted by a titanium electrode placed at the cymba conchae (taVNS; Badran et al., 2018; Burger and Verkuil, 2018) or earlobe (sham; Fig. 1A) of the left ear (Frangos et al., 2015). To match the subjective experience of the stimulation, intensity was determined for each participant and each condition individually to correspond to mild prickling (taVNS: $M_{taVNS} = 1.21$ ±0.43; 0.2-1.9 mA; sham: $M_{sham} = 1.92$ ±0.67; 0.5-3.1 mA). We decided to match the subjective experience of the stimulation to reduce the risk of confounding behavior due to individual differences in sensitivity. To verify that taVNS-induced changes were not dependent on different objective amplitudes of the stimulation, we tested if taVNS-induced changes were associated with the amplitude.

2.4. Paradigm

Due to the reported behavioral and neuromodulatory effects of taVNS, we hypothesized that taVNS affects reinforcement learning. More precisely, we expected reward learning to be reduced in the active stimulation condition as taVNS might translate to an increased dopamine tone that, in turn, decreases the signal-to-noise ratio of RPEs and thus impairs learning (Hamid et al., 2016). Due to the well-known characteristics of the dopaminergic circuit (Frank et al., 2004; Cox et al., 2015), we sought to disentangle effects of taVNS on action- or valence-dependent learning (Guitart-Masip et al., 2012, 2014; Mkrtchian et al., 2017). Specifically, we expected that taVNS-induced changes in the go bias or differential learning from rewards versus punishments would be indicative of differential modulation of direct (D1 receptor dependent) versus indirect pathways (D2 receptor dependent) (Frank et al., 2004, 2007; Frank and O'Reilly, 2006). Moreover, a potential GABAergic modulation would likely be reflected in an alteration of no-go learning due to the well-established role in behavioral inhibition (Quetscher et al., 2015; Cheng et al., 2017). Instead, non-specific taVNS-induced changes could be indicative of differences in dopamine tone within the dopaminergic circuit. Relatedly, Guitart-Masip et al. (2014) reported a decrease in Pavlovian bias by increased dopamine levels after L-DOPA administration, leading to improved performance in incongruent action-valence (e.g. go-to-avoid-punishment) combinations while reducing performance in congruent action-valence combinations (e.g. go-to-win). Alternatively, and not mutually exclusive, increases in noradrenalinergic tone might also have non-specific effects on behavior by altering learning via unsigned prediction errors tracking surprise (Dayan and Yu, 2006; Payzan-LeNestour et al., 2013).

In this task, participants learn state-action contingencies and receive rewards or punishments. Each trial consisted of three stages (Fig. 1B). First, participants saw a fractal cue (state) out of a set of four different fractals per session. These fractals were initially randomized to one of the four possible combinations of the go × win two-factorial design of the task. Second, participants had to complete a target detection task and either respond by pressing a button (go) or withhold their response (no-go). Third, they saw the outcome of the state-action combination, which was either a win (5 cents), punishment (-5 cents), or an omission (no win/punishment, 0 cents). Using trial and error, participants had to learn which action following each fractal was best in terms of maximizing wins or minimizing losses.

Outcomes were presented probabilistically. Thus, participants had 80% chances to win after correct state-action sequences, 20% chances to win after incorrect sequences for rewarded trials as well as 80% chances to avoid losses after correct, and 20% chances to avoid losses after incorrect sequences for punished trials. Participants were instructed about the probabilistic nature of the task and that either go or no-go responses could be correct for a given fractal. There was no change in the contingencies over time. To ensure that participants understood the task, they were queried before starting the task. In total, the task included 240 trials, 120 go trials and 120 no-go trials with 60 trials for each condition (e.g win or avoid loss), and took 15min to be completed.

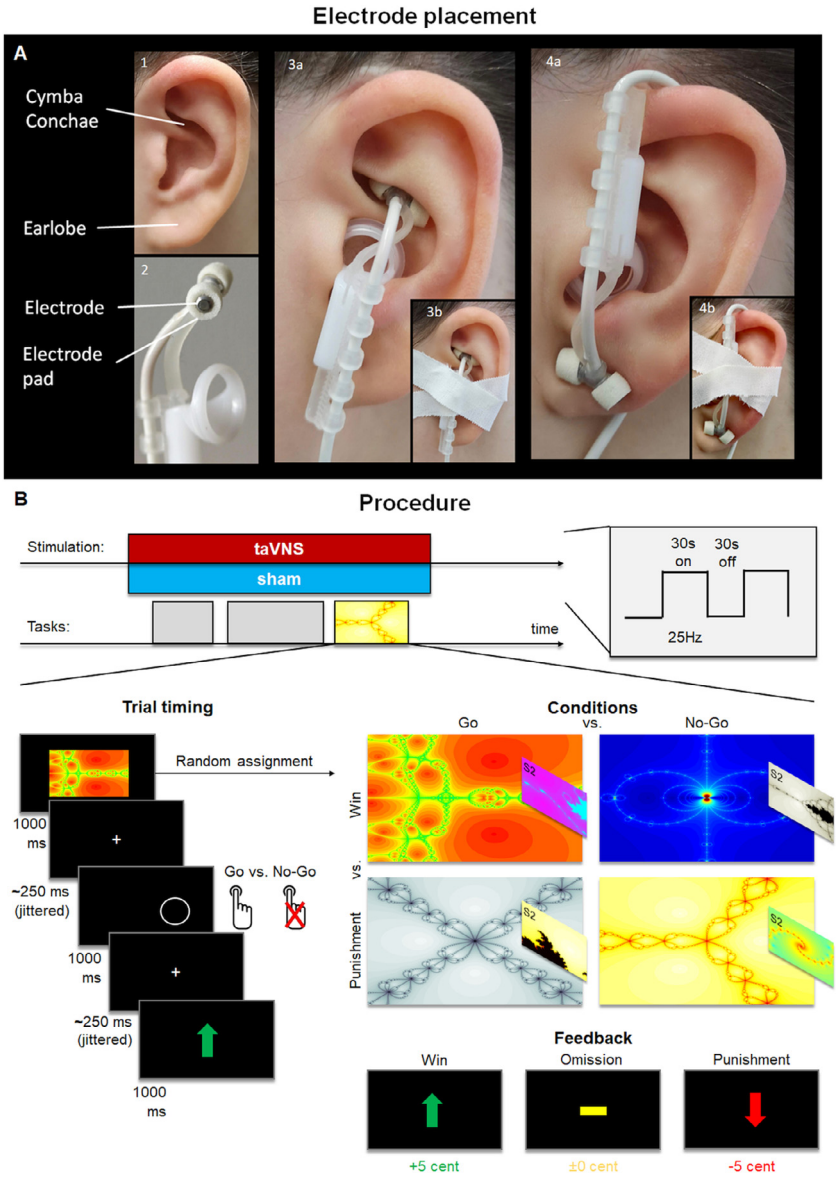


Fig. 1 Schematic summary of the experimental design. A: Placement of the electrodes for taVNS (image 3) and sham conditions (image 4). B: Participants completed taVNS or sham sessions in a randomized crossover design. The stimulation lasted for the duration of the experiment and included alternating blocks of 30s on vs. 30s off stimulation. To maximize the total payoff in the go/no-go reinforcement learning task, participants had to learn which action (go vs. no-go) during the target detection stage following a given fractal was associated with the best possible outcome (i.e., receiving reward or avoiding impending punishment). These contingencies were randomly assigned and had to be learned by trial and error. Insets illustrate that different fractals were used in Session 2 (S2).

2.5. Data analysis

2.5.1. Full mixed-effects analysis of the go/no-go reinforcement learning task

To estimate the effects of taVNS on choice accuracy (go vs. no-go), we defined a full mixed-effects analysis as implemented in HLM (Raudenbush et al., 2011). Effects of the conditions were modeled by predicting if a given choice (Bernoulli distribution) was correct based on the regressors go (dummy coded), win (dummy coded), and the interaction term go × win in a generalized linear model. To assess taVNS effects, the model included terms for the stimulation condition (dummy coded, 0=sham, 1=taVNS) and interactions of the stimulation term with the condition regressors (i.e., stimulation × win, stimulation × go, stimulation × go × win). Furthermore, we included a log-transformed trial regressor capturing improvements in accuracy across trials. At the participant level, we calculated two models that included random effects for all intercepts and slopes: model 1 controlled only for order, whereas model 2 additionally controlled for sex (coded as -0.5 and 0.5) and body mass index (BMI, grand-mean centered). We also tested an additional interaction term stimulation × trial but found that the coefficient estimate was highly correlated with the stimulation main effect. Thus, we excluded this term to avoid redundancy. All other random effects were complementary and showed significant between-subject variance ($p<.001$). Analogous to using expectation maximization in the computational model, we obtained em-

pirical Bayes estimates, which take group-level distributions into account, as individual estimates of taVNS effects.

2.5.2. Reinforcement learning model

To dissociate which facet of instrumental action learning was altered by taVNS, we fit reinforcement learning models to participant's behavior starting with the winning model detailed in Guitart-Masip et al. (2012, 2014) as standard model.

Participants learn stimulus (*s*) specific action (*a*) values (*Q*) that are updated at each trial *t* according to the Rescorla-Wagner rule as follows:

$$Q_t(s_t, a_t) = Q_{t-1}(s_t, a_t) + \alpha(\rho r_t - Q_{t-1}(s_t, a_t)),$$

with learning rate alpha ($\alpha \in [0,1]$), reward sensitivity ρ, a positive free parameter quantifying the individual importance of reward and obtained rewards r_t coded as -1 in case of punishment, 1 in case of reward and 0 if participants received neither reward nor punishment. Further, agents learn action-independent values (V) of each state updated after the same rule indicating if a stimulus is associated with punishments or rewards.

$$V(s_t) = V_{t-1}(s_t) + \alpha(\rho r_t - V_{t-1}(s_t)),$$

Action values (*Q*) and stimulus values (*V*) at each trial are used to compute action weights as follows:

$$W_t(a, s) = \begin{cases} Q_t(a, s) + b + \pi V_t(s), & a = go \\ Q_t(a, s), & else \end{cases}$$

where *b* is a free parameter that reflects a constant bias to choose the go option. The influence of Pavlovian tendencies (e.g. increased go behavior in potentially rewarding situations and avoidance in aversive situations) is parameterized by π, a positive free parameter. The Pavlovian parameter inhibits the go tendency in conditions that are associated with punishments and thus have negative learned state-values (*V*), while it increases go tendencies in conditions associated with reward and positive state-values. Consequently, this leads to impaired learning in incongruent (e.g. go-to-avoid punishment) trials.

The action at each trial is selected based on action probabilities that are estimated by passing action weights (*W*) through a softmax function and adding a noise parameter (*lapse*, $\xi \in [0, 1]$) modulating the influence of learned expectations on subsequent decisions.

$$p(a_t|s_t) = \left[\frac{\exp(W(a_t|s_t))}{\sum_a \exp(W(a'|s_t}\right](1-\xi) + \frac{\xi}{2}$$

Subsequently, we fit three further models to disentangle possible effects depending on reward valence by estimating either learning rate, learning rate and reward sensitivity, or learning rate, reward sensitivity and Pavlovian bias for reward and punishment conditions separately.

Models were fit using hierarchical expectation maximization as described by Huys et al. (2011). To fit models with expectation maximization, individual parameters as well as the underlying group distribution parameters are estimated iteratively. The current group distributions are used as priors to estimate individual level parameters using Laplace approximation in the E-step. Consequently, in the M-step, group-level distributions are updated based on the new individual parameter estimates and their uncertainty. Model fit was assessed using group-level integrated Bayesian information criterion (iBIC, Huys et al., 2011) where model fit and model complexity across all measurements are taken into account. As better group-level fit may be driven by large improvements in few participants, we additionally used likelihood-ratio tests to determine the best fitting model for each session. To ensure stability of individual estimates, we repeated the fitting procedure 10 times. In each repetition, we used the complete data, but used different initializations of the algorithm. We calculated the mean and the coefficient of variation of the parameters for each session and participant. Repeated sessions (sham and taVNS) were treated as independent measurements and one underlying distribution was fit over all participants and measurements. Reward sensitivity and Pavlovian bias parameters were log transformed and learning rate and noise parameters were transformed using the inverse sigmoid function to ensure theoretical parameter constraints. Furthermore, we assessed recovery of observed behavior based on simulations with estimated parameters. To this end, we simulated 100 runs of the experiment with individual parameter estimates from each session.

2.5.3. Statistical analysis and software

We assessed all taVNS effects using a significance threshold of $p < .05$ (two-tailed) and corrected for multiple comparisons across the five parameters in the computational model analysis using Bonferroni correction. We also planned correction across condition-specific interaction terms in the mixed-effects model, but they did not reach uncorrected significance. To account for non-normal distributions of parameters from the computational model, differences in parameter estimates between the taVNS and sham condition were tested using bootstrapping (1000 resamples). We performed data analyses with Matlab v2016a (computational model) or HLM v7 (mixed-effects models) and data visualization with R v3.5.1 (R Core Team, 2018) and Deducer (Fellows, 2012).

3. Results

3.1. taVNS reduces choice accuracy across conditions

We first analyzed the performance of participants by estimating effects of reward valence, required action, and stimulation on accuracy in a full mixed-effects model. In line with previous studies, accuracy was higher in conditions requiring a go response ($t = 5.93$, $p < .001$), whereas reward valence only influenced accuracy in interaction with the required action (valence: $t = 0.83$, $p = .41$, valence × action: $t = 7.198$, $p < .001$). In other words, participants performed worse in the *go-punishment* and *no-go-win* conditions in which Pavlovian biases (*approach reward; avoid punishment*) and instrumental behavior were incongruent.

Next, we assessed main and interaction effects of taVNS vs. sham stimulation on choice accuracy. Across conditions, taVNS reduced accuracy ($t = -2.13$, $p = .041$; model uncorrected for BMI and sex: $t = -1.98$, $p = .055$). However, we observed no interaction effects with action ($t = -0.35$, $p = .73$) or valence ($t = 1.08$, $p = .29$). As task-dependent improvement across sessions is common in reinforcement learning tasks, we controlled for stimulation order (sham/taVNS first) in the analyses. Order of stimulation modulated stimulation slopes ($t = -3.60$, $p < .001$) with stronger impairments in overall performance if taVNS was applied first. Notably, acute taVNS-induced reductions in performance did not lead to deficits in the second session with higher day-to-day improvements in the group that received taVNS first ($t = 2.05$, $p = .048$) (Fig. 2).

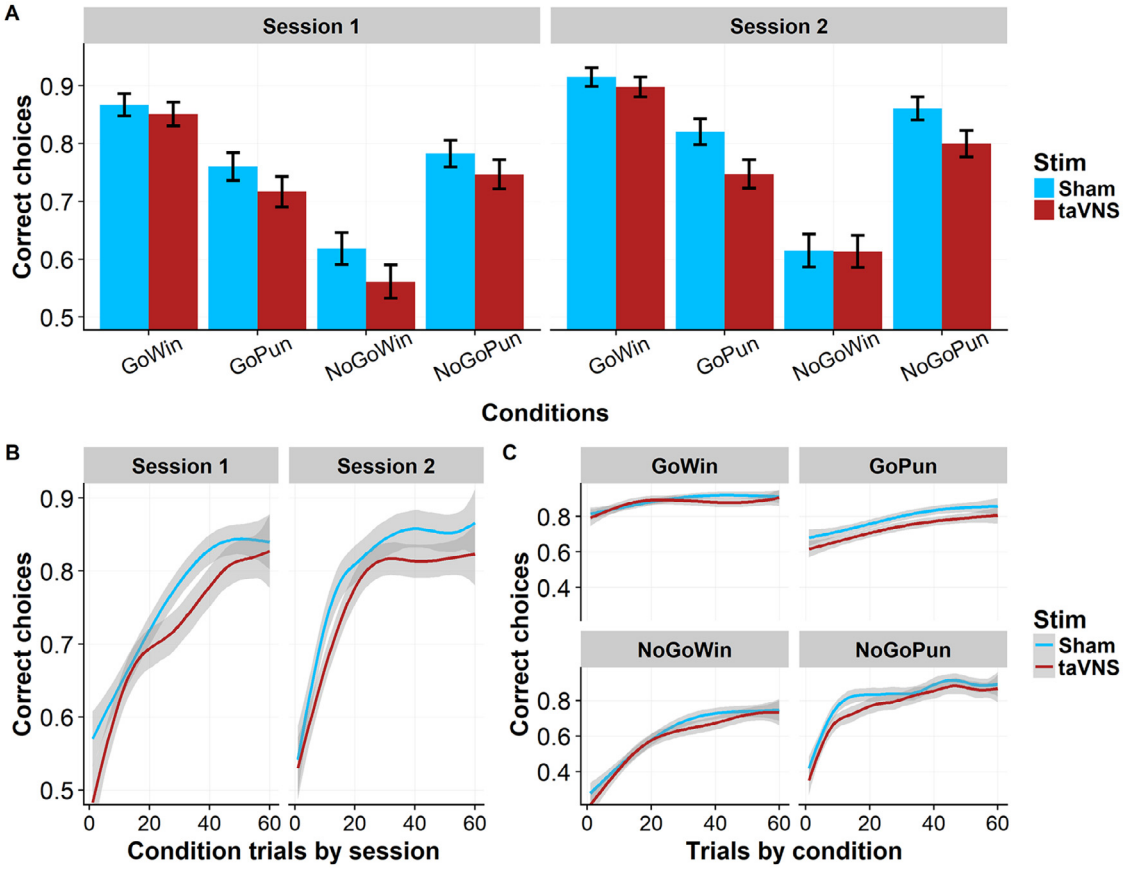


Fig. 2 Choice accuracy is reduced in the taVNS condition compared to sham stimulation. A: Mean choice accuracy for taVNS and sham stimulation in each session and condition. Error bars depict 95% confidence intervals. B: Choice accuracy for taVNS and sham stimulation over all trials separated by session averaged across conditions indicate stronger taVNS-induced reduction of choice accuracy in session 1. C: Choice accuracy for taVNS and sham stimulation over trials separated by condition do not suggest action- or valence-specific effects of taVNS.

3.2. taVNS reduces the learning rate in a computational model of behavior

To further characterize which learning processes were affected by taVNS leading to impaired performance, we fitted a computational reward-learning model (Guitart-Masip et al., 2012) using an expectation maximization algorithm to empirically regularize parameter estimates. We estimated five parameters controlling choices over time: learning rate, reward sensitivity, go bias, Pavlovian bias, and noisiness of choices for each session and calculated differences between taVNS and sham sessions. Simulated data based on individually estimated parameters corresponded well with observed data (Fig. 3A-C). Parameter estimates were sufficiently stable across the 10 repeated initializations of the expectation maximization algorithm with median coefficients of variation between .002 and .030 for the five parameters.

Impaired performance during taVNS was mainly reflected in a reduced learning rate alpha (Fig. 4A-C; $\Delta\alpha$= -0.092, p = .009, p_{boot} = .002, corrected for stimulation order: t = -2.741, p = .009; Bonferroni corrected alpha level = .01). Additionally, participant's choices in the taVNS condition were 'noisier' and less dependent on learned action values ($\Delta\xi$ = 0.035, p = .086, p_{boot} = .050), although only nominally significant before correction for multiple testing. Crucially, the taVNS-induced changes in α were independent of the stimulation type participants thought they had received in a given session (p = .80) as well as of the taVNS amplitude (r = .10, p = .56, Bayes factor (BF) = 0.23), also if it was tested separately for wins (r = .20, p = .23, BF = 0.39) and losses (r = .04, p = .79, BF = 0.21).

Valence-specific effects of taVNS may be captured by modeling separate parameters for rewards and punishments. Therefore, we also built an extended 6-parameter model assuming separate learning rates. While the 6-parameter model provided a more parsimonious account at the group level (ΔiBIC = 263), it did not improve individual model fits for 51 out of 78 sessions. Nonetheless, stability of individual parameter estimates was sufficient (median coefficient of variation in the range between 0.004 - 0.074 for the six parameters). Subsequent estimation of taVNS effects revealed that the slower learning rate during taVNS stimulation was predominantly driven by a decrease of the learning rate in the punishment condition ($\Delta\alpha_{pun}$ = -0.081, p = .019, p_{boot} = .012, corrected for order: t = -2.516, p = .016) while decreases of alpha in reward conditions were less pronounced and non-significant ($\Delta\alpha_{rew}$ = -0.031,

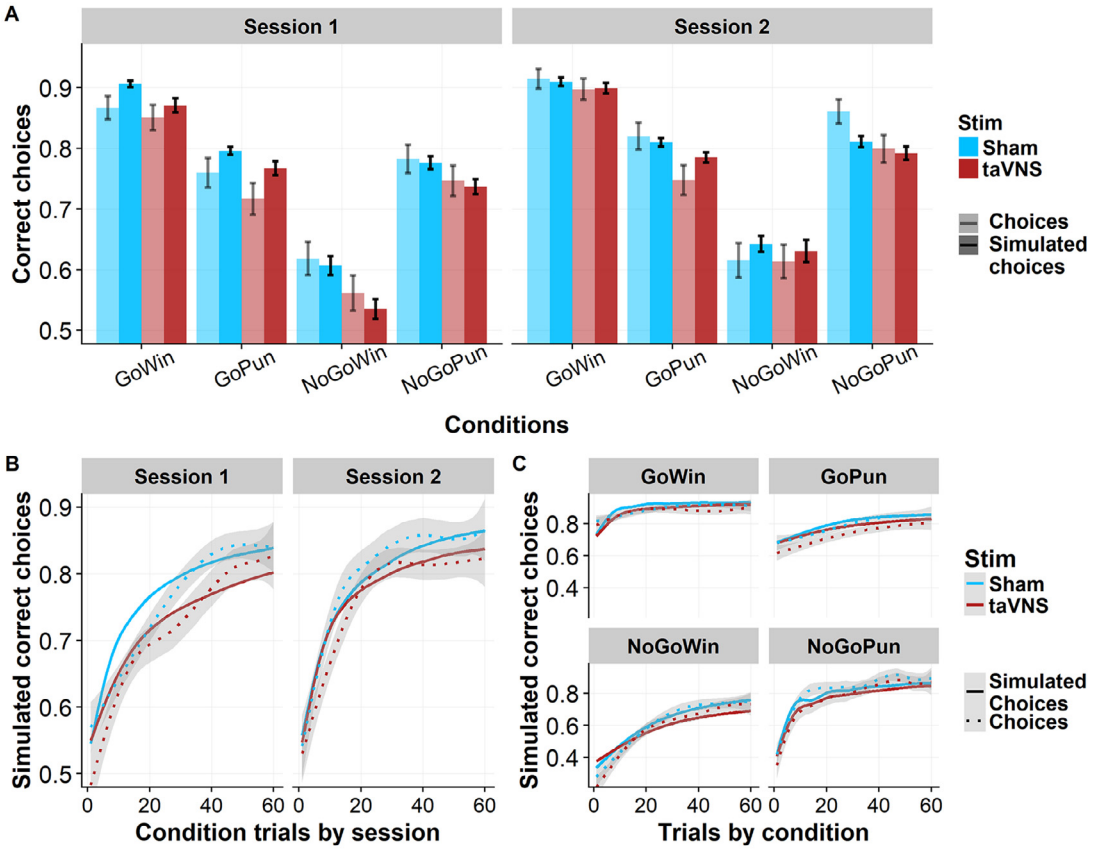


Fig. 3 A: Mean choice accuracy simulated from individual parameter estimates recover stimulation effects in each session and condition. Error bars depict 95% confidence intervals. Transparent bars depict choices, dark bars simulated choices. B: Choices simulated from individual parameter estimates recover participants' choice patterns and stimulation effects for sessions (all trials averaged across conditions). C: Recovered choice patterns indicate no difference in stimulation effects depending on valence or action.

$p = .22$ $p_{boot} = .21$, corrected for order: $t = -1.244$, $p = .21$). However, the interaction between stimulation × valence for the learning rate was not significant, $F(1,37) = 1.975$, $p = .168$, indicating only weak specificity of the taVNS effect on punishment learning. In contrast to the 5-parameter model, taVNS did not affect choice stochasticity in the extended model ($\Delta\xi = -0.0031$, $p = .86$, $p_{boot} = .94$). Again, stimulation effects on performance were recovered in the averaged simulated data.

We also explored more complex models by additionally separating reward sensitivity and/or Pavlovian bias for reward and punishment as previously described (Mkrtchian et al., 2017). However, these models did not provide a more parsimonious account at the individual level compared to the simpler models and individual estimates became increasingly unstable across iterations precluding their use to reliably estimate within-subject stimulation effects. An additional model including an exponential decay parameter for the learning rate revealed a comparable reduction in initial learning rates by taVNS, but simulated data did not recover the empirical data as well. Modeling taVNS effects as within-participant change in one specific parameter (Swart et al., 2017) showed worse model fit, indicating that taVNS leads to changes in multiple aspects of reward learning that are dominated by a decreased learning rate.

3.3. taVNS effects on the learning rate are associated with body weight

Since changes in body weight have been linked to alterations in the sensitivity to interoceptive feedback signals (Herbert and Pollatos, 2014; Simmons and DeVille, 2017), it is possible that behavioral effects of taVNS might be dependent on BMI as well. In line with this expectation, taVNS effects on general accuracy ($t = 2.000$, $p = .053$), as well as on learning rate ($t = 2.351$, $p = .024$) depended partly on participants' BMI. More specifically, taVNS reduced the speed of acquisition more strongly in participants with a low (healthy) BMI. However, as the association with accuracy was not significant in our sample, our results for a potential modulation by BMI should be regarded as preliminary.

3.4. taVNS does not affect response time

Lastly, we estimated effects of taVNS on response time as an indicator of alterations in response vigor. However, no significant changes in response time were observed ($t = 0.826$, $p = .41$, Fig. 5). This further corroborates that impaired performance was mediated by slowed learning and not by altered action selection.

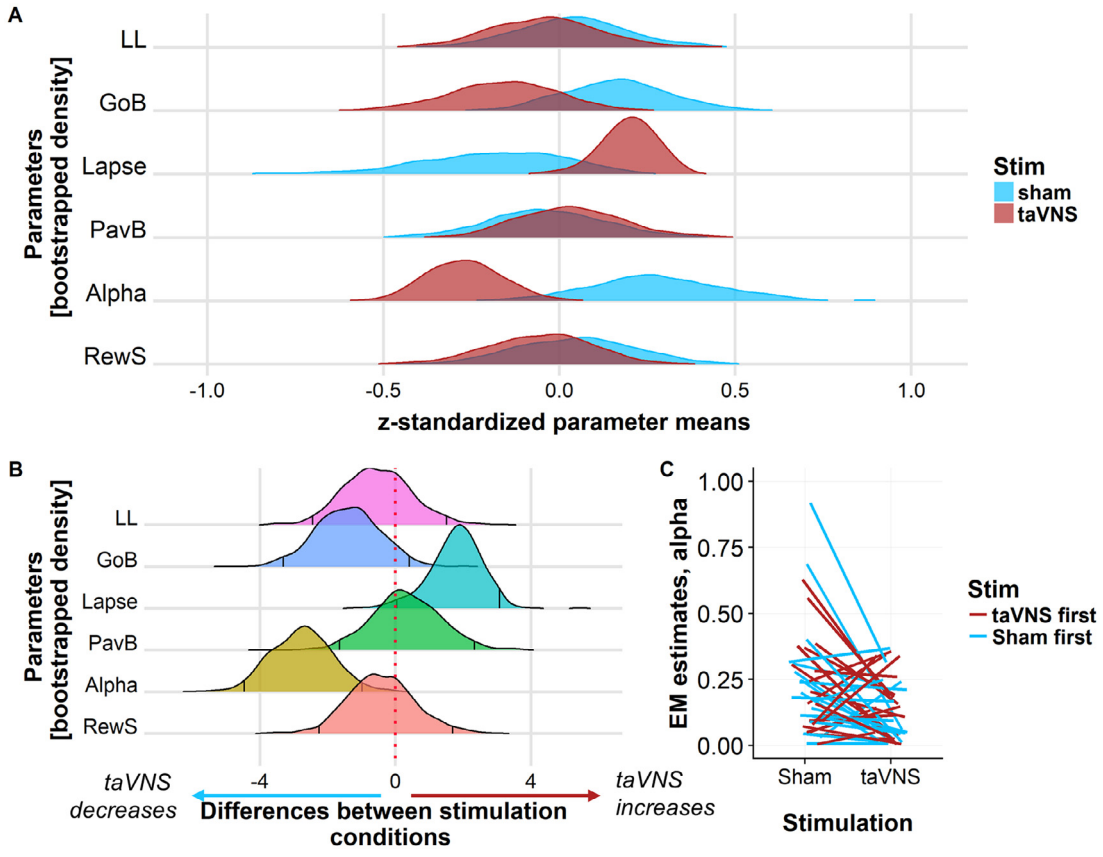


Fig. 4 Reduced choice accuracy in the taVNS condition is driven by a reduced learning rate and increased choice stochasticity in the 5-parameter computational model. A: Bootstrapped density plots of parameter means separately for sham and taVNS. Standardization of the values was calculated across all estimates to visualize the parameter differences on a comparable scale (z-score). B: Bootstrapped density plots of the differences in individual parameter estimates between taVNS and sham stimulation (in t-values). Lines indicate 95% confidence intervals. C: Individual changes in the learning rate indicate mainly a reduction of high baseline learning rates after taVNS. Stimulation in session 1: blue = sham, red = taVNS LL = Log-Likelihood, GoB = Go bias, PavB = Pavlovian bias (π), RewS = Reward sensitivity (ρ). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

The vagus nerve rapidly transmits interoceptive signals to the brain. It thereby confers interoceptive information such as metabolic state and modulates neurotransmission including within the dopamine system. Here, we investigated changes in instrumental reinforcement learning, which is critically dependent on dopamine, after emulating vagal feedback signals using taVNS. Importantly, we found that taVNS reduced overall accuracy of choices driven by a slowed acquisition of action contingencies, predominantly for punishments. In contrast, action- or valence-specific biases were unaffected by taVNS. Thus, using the novel non-invasive stimulation of the vagus nerve, our results provide evidence that interoceptive feedback signals may alter reward learning by tuning the speed of acquisition according to interoceptive signals.

Vagal feedback signals evoked by taVNS acutely impaired choice accuracy and reduced learning rates in valenced go-/no-go learning. These results provide further evidence on the crucial modulatory role of vagal efferent signals in the control of motivated behavior. For example, preclinical studies have shown that vagal sensory signaling promotes hippocampal memory function via the NTS which is important for food-seeking behavior (Suarez et al., 2018). Capitalizing on vagal afferent signals via an implanted vagal stimulation device has also been shown to provide a powerful means to restrict food intake and control body weight in rodents (Yao et al., 2018). Notably, the effects of bariatric surgery on fat appetite are also partly modulated via afferent vagal feedback signals leading to changes in dorsal striatal dopamine signaling (Hankir et al., 2017). Moreover, flavor-nutrient conditioning in humans is dependent on metabolic states and peripheral energy metabolism (Yeomans et al., 2008; de Araujo et al., 2013; Veldhuizen et al., 2017) demonstrating the coupling between homeostatic signals and reward learning. By showing a modulation of reinforcement learning for non-food rewards in humans, our study adds an important insight to the growing literature on vagally-mediated aspects of motivated behavior.

The observed changes in choice accuracy and the learning rate are in agreement with the value theory of dopamine and with studies showing that the impact of phasic RPE

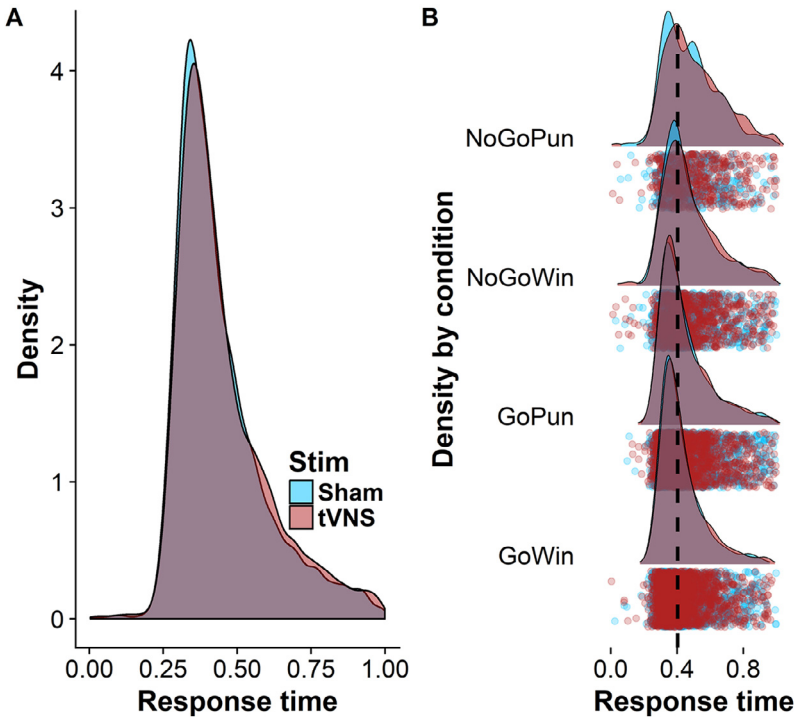


Fig. 5 Transcutaneous vagus nerve stimulation (taVNS) does not induce changes in response time as is seen across (A) and within (B) conditions.

signals on actions depends on dopamine tone (Hamid et al., 2016; Kroemer et al., 2019). In short, increased dopamine tone leads to a comparably smaller signal-to-noise ratio if phasic signals are unaffected and to reduced action control via phasic outcome signals. Accordingly, reduced learning after L-DOPA administration has been reported in patients with Parkinson's disease (Cools et al., 2007; Vo et al., 2014) as well as healthy participants (Vo et al., 2016). Impaired performance may also be caused by increased choice stochasticity (Beeler, 2012; Eisenegger et al., 2014), but taVNS-induced increases in decision noise were not significant after correction for multiple testing and not consistent across models suggesting limited effects at best. Collectively, these results suggest that taVNS primarily affects action-contingency learning and not solely noise in value-based decisions. In line with animal work showing increased dopamine signaling after endogenous stimulation of the vagus nerve (de Araujo et al., 2012; Ferreira et al., 2012) and subsequent modulations of appetitive learning (Davis et al., 2008; Lartigue, 2016) and motivated behavior (Palmiter, 2008), this may indicate that reduced learning rates during taVNS could be a consequence of increases in dopamine tone.

Alternatively, changes in noradrenergic transmission may explain general alterations in reward learning as well, possibly via a comparable decrease in the phasic signal-to-noise ratio. Administration of taVNS has also been associated with heightened noradrenergic signaling via the locus coeruleus leading to improved memory performance by increasing arousal and attention (Burger et al., 2016). Nonetheless, phasic noradrenaline signals have also been shown to track unsigned prediction errors ("surprise") (Dayan and Yu, 2006; Payzan-LeNestour et al., 2013). Surprise signals are critical for learning and, accordingly, treatment with an noradrenaline reuptake inhibitor was associated with comparable changes in learning rates (Jepma et al., 2016). Since action contingencies are fixed throughout the task, it is not possible to dissociate dopaminergic and noradrenergic processes acting via signed (reward) or unsigned (surprise) prediction errors, respectively. Moreover, as dopamine is the precursor of noradrenaline, future studies disentangling both systems are necessary. Notwithstanding, behavioral effects consistent with taVNS-induced increases in monoaminergic tone provide only indirect evidence in humans and studies further elucidating mechanisms of taVNS, for example, by PET or pharmacological imaging are necessary.

In contrast to our hypothesis, taVNS did neither affect response-specific biases such as Pavlovian or go biases nor response times in any condition. In previous studies, pharmacologically-induced increases in tonic dopamine modulated Pavlovian (Guitart-Masip et al., 2014) or motivational biases (Swart et al., 2017) and differentially affected learning from rewards versus punishments (Frank, 2005; Cools et al., 2006). Although we observed that punishment learning, but not reward learning was significantly reduced, there was no significant interaction between valence-dependent learning rates and taVNS effects. Thus, taVNS-induced effects were generally independent of valence or the required action. One plausible explanation is that taVNS affects multiple transmitter systems and their interplay might lead to mixed behavioral alterations compared to pharmacological interventions. However, many common drugs such as L-DOPA also act on other transmission systems (De Deurwaerdère et al., 2017) suggesting that this is an insufficient explanation. Another possibility is that modulatory effects of taVNS are more confined within the

motivational circuit compared to systemic drug administration. For example, it is conceivable that taVNS could alter the balance between fast reinforcement learning, primarily linked to the amygdala, and slow reinforcement learning, primarily linked to the striatum (Averbeck and Costa, 2017). It has been shown that chronic taVNS increases functional connectivity between the amygdala and the prefrontal cortex in depressed patients (Liu et al., 2016) whereas VNS acutely reduces amygdala-evoked responses in the prefrontal cortex of rats (Lyubashina and Panteleev, 2009). Thus, future research may help to resolve these questions by detailing corresponding alterations in motivational circuits as our study leads to testable predictions about shifting the balance more towards slow striatal reinforcement learning (Averbeck and Costa, 2017). Moreover, a potential modulation of GABA would have likely led to specific changes in response inhibition versus execution (Quetscher et al., 2015; Cheng et al., 2017) and our results provide little support for changes in inhibitory control. To conclude, taVNS appears to primarily reduce the speed of reinforcement learning, but more research is needed to establish corresponding changes in neurotransmission in humans.

While we initially hypothesized that behavioral effects of taVNS in the reinforcement learning task could be caused by modulations of dopamine, our study has the major limitation of lacking a direct dopaminergic readout. Although we focused on dopamine, taVNS is arguably not specifically targeting one neurotransmitter system. Additional PET or pharmacological fMRI studies investigating precisely which targets and neuromodulators are affected by taVNS in humans may shed light on the pathway driving reductions in learning. Furthermore, the within-subject crossover design offers increased statistical power to detect stimulation effects, especially considering baseline dependence. Nonetheless, repeated completion of the task may have affected performance and modulated taVNS effects. We accounted for order effects in the statistical analyses, but replication in independent groups would be preferable. Lastly, we investigated the effects of taVNS after an overnight fast only. In future studies, it would be desirable to test taVNS-induced effects also in a sated or a neither hungry, nor full condition. This would help to directly compare taVNS-induced effects to differences induced by certain homeostatic states.

To summarize, we showed that vagal signals impair choice accuracy by acutely reducing learning speed in a reinforcement learning task. Slower acquisition may be due to a reduced signal-to-noise ratio of evoked phasic dopamine according to the value theory of dopamine. Our findings are in accordance with the hypothesis that vagal afferents may modulate dopamine tone, but this remains to be confirmed in humans with markers of dopaminergic transmission. We conclude that how much we learn from rewards and punishments may depend on interoceptive signals forwarded by the vagus nerve. Thereby, rapid learning which actions in a given environment lead to future reward or punishment could be facilitated during a hungry state compared to a less deprived state. Critically, this behavioral flexibility with respect to the current metabolic state was less pronounced in overweight participants, which is in line with the reported reduced sensitivity to peripheral metabolic feedback (Klok et al., 2007). Furthermore, reported anti-depressant effects of taVNS may partly rely on reduced learning, especially from punishments, as this may compensate for the reported increased punishment sensitivity in depressed patients (Hevey et al., 2017; Mkrtchian et al., 2017). More broadly, reduced dependence on learned contingencies may also offer the possibility to prevent over-reliance on learned action–outcome combinations and encourage exploration. In turn, this could lead to greater behavioral flexibility that may be advantageous in many naturalistic environments.

Role of funding source

The study was supported by the University of Tübingen, Faculty of Medicine fortune grant #2453-0-0. VT, CB, & NBK received salary support from the University of Tübingen, Faculty of Medicine fortune grant #2453-0-0. MPN received salary support from the Else Kröner-Fresenius Stiftung, grant 2017_A67. QJMH is supported by the UCL NIHR BRC. The funding sources had no further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

Contributors

NBK was responsible for the study concept and design. VT implemented the task. CB & MPN collected data under supervision by MW & NBK. AK, QJMH, & NBK conceived the method including statistical and computational models. AK & NBK processed the data and performed the data analysis. AK & NBK wrote the manuscript. All authors contributed to the interpretation of findings, provided critical revision of the manuscript for important intellectual content and approved the final version for publication.

Conflict of Interest

All authors declare that they have no competing conflicts of interest.

Acknowledgment

We thank Anni Richter for sharing the German task instructions with us.

References

Averbeck, BB, Costa, VD, 2017. Motivational neural circuits underlying reinforcement learning. Nat. Neurosci. 20, 505–512.

Badran, BW, Brown, JC, Dowdle, LT, Mithoefer, OJ, LaBate, NT, Coatsworth, J, DeVries, WH, Austelle, CW, McTeague, LM, Yu, A, Bikson, M, Jenkins, DD, George, MS, 2018. Tragus or cymba conchae? Investigating the anatomical foundation of transcutaneous auricular vagus nerve stimulation (taVNS). Brain Stimul.: Basic Transl. Clin. Res. Neuromodul. 11, 947–948.

Bauer, S, Baier, H, Baumgartner, C, Bohlmann, K, Fauser, S, Graf, W, Hillenbrand, B, Hirsch, M, Last, C, Lerche, H, Mayer, T, Schulze-Bonhage, A, Steinhoff, BJ, Weber, Y, Hartlep, A, Rosenow, F, Hamer, HM, 2016. Transcutaneous vagus nerve stimulation (tvns) for treatment of drug-resistant epilepsy: a randomized, double-blind clinical trial (cMPsE02). Brain Stimul. 9, 356–363.

Beeler, JA, 2012. Thorndike's Law 2.0: Dopamine and the Regulation of Thrift. Front Neurosci. 6. Available at: https://www.frontiersin.org/articles/10.3389/fnins.2012.00116/full [Accessed December 26, 2018].

Ben-Menachem, E, Hamberger, A, Hedner, T, Hammond, EJ, Uthman, BM, Slater, J, Treig, T, Stefan, H, Ramsay, RE, Wernicke, JF, 1995. Effects of vagus nerve stimulation on amino acids and other metabolites in the CSF of patients with partial seizures. Epilepsy Res. 20, 221–227.

Beste, C, Steenbergen, L, Sellaro, R, Grigoriadou, S, Zhang, R, Chmielewski, W, Stock, A-K, Colzato, L, 2016. Effects of concomitant stimulation of the gabaergic and norepinephrine system on inhibitory control – a study using transcutaneous vagus nerve stimulation. Brain Stimul. 9, 811–818.

Burger, AM, Verkuil, B, 2018. Transcutaneous nerve stimulation via the tragus: are we really stimulating the vagus nerve? Brain Stimul.: Basic, Transl. Clin. Res. Neuromodul. 11, 945–946.

Burger, AM, Verkuil, B, Van Diest, I, Van der Does, W, Thayer, JF, Brosschot, JF, 2016. The effects of transcutaneous vagus nerve stimulation on conditioned fear extinction in humans. Neurobiol. Learn. Mem. 132, 49–56.

Breit, S, Kupferberg, A, Rogler, G, Hasler, G, 2018. Vagus Nerve as Modulator of the Brain–Gut Axis in Psychiatric and Inflammatory Disorders. Front Psychiatry 9. Available at: https://www.frontiersin.org/articles/10.3389/fpsyt.2018.00044/full [Accessed January 31, 2020].

Capone, F, Assenza, G, Di Pino, G, Musumeci, G, Ranieri, F, Florio, L, Barbato, C, Di Lazzaro, V, 2015. The effect of transcutaneous vagus nerve stimulation on cortical excitability. J. Neural Transm. 122, 679–685.

Cheng, C-H, Niddam, DM, Hsu, S-C, Liu, C-Y, Tsai, S-Y, 2017. Resting GABA concentration predicts inhibitory control during an auditory Go-Nogo task. Exp. Brain Res. 235, 3833–3841.

Cools, R, Altamirano, L, D'Esposito, M, 2006. Reversal learning in Parkinson's disease depends on medication status and outcome valence. Neuropsychologia 44, 1663–1673.

Cools, R, Lewis, SJG, Clark, L, Barker, RA, Robbins, TW, 2007. L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. Neuropsychopharmacology 32, 180–189.

Cox, SML, Frank, MJ, Larcher, K, Fellows, LK, Clark, CA, Leyton, M, Dagher, A, 2015. Striatal D1 and D2 signaling differentially predict learning from positive and negative outcomes. Neuroimage 109, 95–101.

Date, Y, 2012. Chapter sixteen – ghrelin and the vagus nerve. In: Kojima, M, Kangawa, K (Eds.), Methods in Enzymology. Academic Press, Ghrelin, pp. 261–269 Available at: http://www.sciencedirect.com/science/article/pii/B9780123812728000167 [Accessed January 31, 2020].

Davis, JF, Tracy, AL, Schurdak, JD, Tschöp, MH, Lipton, JW, Clegg, DJ, Benoit, SC, 2008. Exposure to elevated levels of dietary fat attenuates psychostimulant reward and mesolimbic dopamine turnover in the rat. Behav. Neurosci. 122, 1257–1263.

Dayan, P, Yu, AJ, 2006. Phasic norepinephrine: A neural interrupt signal for unexpected events. Netw. Comput. Neural Syst. 17, 335–350.

de Araujo, IE, Ferreira, JG, Tellez, LA, Ren, X, Yeckel, CW, 2012. The gut–brain dopamine axis: a regulatory system for caloric intake. Physiol. Behav. 106, 394–399.

De Deurwaerdère, P, Di Giovanni, G, Millan, MJ, 2017. Expanding the repertoire of L-DOPA's actions: a comprehensive review of its functional neurochemistry. Prog. Neurobiol. 151, 57–100.

de Lartigue, G, Ronveaux, CC, Raybould, HE, 2014. Deletion of leptin signaling in vagal afferent neurons results in hyperphagia and obesity. Mol. Metab. 3, 595–607.

de Araujo, IE, Lin, T, Veldhuizen, MG, Small, DM, 2013. Metabolic regulation of brain response to food cues. Curr. Biol. 23, 878–883.

Dockray, GJ, 2009. Cholecystokinin and gut–brain signalling. Regul. Pept. 155, 6–10.

Eisenegger, C, Naef, M, Linssen, A, Clark, L, Gandamaneni, PK, Müller, U, Robbins, TW, 2014. Role of dopamine D2 receptors in human reinforcement learning. Neuropsychopharmacology 39, 2366–2375.

Fellows, I, 2012. Deducer: a data analysis GUI for R. J. Stat. Softw. 49, 1–15.

Ferreira, JG, Tellez, LA, Ren, X, Yeckel, CW, Araujo, IE de, 2012. Regulation of fat intake in the absence of flavour signalling. J. Physiol. 590, 953–972.

Frangos, E, Ellrich, J, Komisaruk, BR, 2015. Non-invasive access to the vagus nerve central projections via electrical stimulation of the external ear: fMRI evidence in humans. Brain Stimul. 8, 624–636.

Frank, MJ, 2005. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. J. Cogn. Neurosci. 17, 51–72.

Frank, MJ, O'Reilly, RC, 2006. A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. Behav. Neurosci. 120, 497–517.

Frank, MJ, Santamaria, A, O'Reilly, RC, Willcutt, E, 2007. Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. Neuropsychopharmacology 32, 1583–1599.

Frank, MJ, Seeberger, LC, O'Reilly, RC, 2004. By carrot or by stick: cognitive reinforcement learning in Parkinsonism. Science 306, 1940–1943.

Grill, HJ, Hayes, MR, 2012. Hindbrain neurons as an essential hub in the neuroanatomically distributed control of energy balance. Cell Metab. 16, 296–309.

Guitart-Masip, M, Economides, M, Huys, QJM, Frank, MJ, Chowdhury, R, Duzel, E, Dayan, P, Dolan, RJ, 2014. Differential, but not opponent, effects of l-DOPA and citalopram on action learning with reward and punishment. Psychopharmacology 231, 955–966.

Guitart-Masip, M, Huys, QJM, Fuentemilla, L, Dayan, P, Duzel, E, Dolan, RJ, 2012. Go and no-go learning in reward and punishment: interactions between affect and effect. Neuroimage 62, 154–166.

Gurel, NZ, Huang, M, Wittbrodt, MT, Jung, H, Ladd, SL, MdMH, Shandhi, Ko, Y-A, Shallenberger, L, Nye, JA, Pearce, B, Vaccarino, V, Shah, AJ, Bremner, JD, Inan, OT, 2020. Quantifying acute physiological biomarkers of transcutaneous cervical vagal nerve stimulation in the context of psychological stress. Brain Stimul. 13, 47–59.

Hamid, AA, Pettibone, JR, Mabrouk, OS, Hetrick, VL, Schmidt, R, Vander Weele, CM, Kennedy, RT, Aragona, BJ, Berke, JD, 2016. Mesolimbic dopamine signals the value of work. Nat. Neurosci. 19, 117–126.

Han, W, Tellez, LA, Perkins, MH, Perez, IO, Qu, T, Ferreira, J, Ferreira, TL, Quinn, D, Liu, Z-W, Gao, X-B, Kaelberer, MM, Bohórquez, DV, Shammah-Lagnado, SJ, de Lartigue, G, de Araujo, IE, 2018. A neural circuit for gut-induced reward. Cell 175, 665–678 .e23.

Hankir, MK, Seyfried, F, Hintschich, CA, Diep, T-A, Kleberg, K, Kranz, M, Deuther-Conrad, W, Tellez, LA, Rullmann, M, Patt, M, Teichert, J, Hesse, S, Sabri, O, Brust, P, Hansen, HS, de Araujo, IE, Krügel, U, Fenske, WK, 2017. Gastric bypass surgery recruits a gut PPAR-α-striatal D1R pathway to reduce fat appetite in obese rats. Cell Metab. 25, 335–344.

He, W, Jing, X-H, Zhu, B, Zhu, X-L, Li, L, Bai, W-Z, Ben, H, 2013. The auriculo-vagal afferent pathway and its role in seizure suppression in rats. BMC Neurosci. 14, 85.

Herbert, BM, Pollatos, O, 2014. Attenuated interoceptive sensitivity in overweight and obese individuals. Eat. Behav. 15, 445–448.

Hevey, D, Thomas, K, Laureano-Schelten, S, Looney, K, Booth, R, 2017. Clinical depression and punishment sensitivity on the BART. Front. Psychol. 8. Available at: https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00670/full. [Accessed December 26, 2018].

Hulsey, DR, Hays, SA, Khodaparast, N, Ruiz, A, Das, P, Rennaker, RL, Kilgard, MP, 2016. Reorganization of motor cortex by vagus nerve stimulation requires cholinergic innervation. Brain Stimul. 9, 174–181.

Huys, QJM, Cools, R, Gölzer, M, Friedel, E, Heinz, A, Dolan, RJ, Dayan, P, 2011. Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. PLoS Comput. Biol. 7, e1002028.

Jacobs, HIL, Riphagen, JM, Razat, CM, Wiese, S, Sack, AT, 2015. Transcutaneous vagus nerve stimulation boosts associative memory in older individuals. Neurobiol. Aging 36, 1860–1867.

Jepma, M, Murphy, PR, Nassar, MR, Rangel-Gomez, M, Meeter, M, Nieuwenhuis, S, 2016. Catecholaminergic regulation of learning rate in a dynamic environment. PLoS Comput. Biol. 12, e1005171.

Keramati, M, Gutkin, B, 2014. Homeostatic reinforcement learning for integrating reward collection and physiological stability. eLife 3, e04811. Available at: https://elifesciences.org/articles/04811 . [Accessed December 26, 2018].

Klok, MD, Jakobsdottir, S, Drent, ML, 2007. The role of leptin and ghrelin in the regulation of food intake and body weight in humans: a review. Obes. Rev. 8, 21–34.

Korn, CW, Bach, DR, 2015. Maintaining homeostasis by decision-making. PLoS Comput. Biol. 11, e1004301.

Kraus, T, Kiess, O, Hösl, K, Terekhin, P, Kornhuber, J, Forster, C, 2013. CNS BOLD fMRI effects of sham-controlled transcutaneous electrical nerve stimulation in the left outer auditory canal – a pilot study. Brain Stimul. 6, 798–804.

Kreuzer, PMM, Landgrebe, M, Husser, O, Resch, M, Schecklmann, M, Geisreiter, F, Poeppl, TB, Prasser, SJ, Hajak, G, Langguth, B, 2012. Transcutaneous vagus nerve stimulation: retrospective assessment of cardiac safety in a pilot study. Front. Psychiatry 3, 70. Available at: https://www.frontiersin.org/articles/10.3389/fpsyt.2012.00070/full. [Accessed April 14, 2019].

Kroemer, NB, Lee, Y, Pooseh, S, Eppinger, B, Goschke, T, Smolka, MN, 2019. L-DOPA reduces model-free control of behavior by attenuating the transfer of value to action. Neuroimage 186, 113–125.

Kroemer, NB, Small, DM, 2016. Fuel not fun: Reinterpreting attenuated brain responses to reward in obesity. Physiol. Behav. 162, 37–45.

Lartigue, G de, 2016. Role of the vagus nerve in the development and treatment of diet-induced obesity. J. Physiol. 594, 5791–5815.

Lee, SW, Kulkarni, K, Annoni, EM, Libbus, I, KenKnight, BH, Tolkacheva, EG, 2018. Stochastic vagus nerve stimulation affects acute heart rate dynamics in rats. PLoS One 13, e0194910. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5874066/ . [Accessed February 6, 2020].

Liu, J, Fang, J, Wang, Z, Rong, P, Hong, Y, Fan, Y, Wang, X, Park, J, Jin, Y, Liu, C, Zhu, B, Kong, J, 2016. Transcutaneous vagus nerve stimulation modulates amygdala functional connectivity in patients with depression. J. Affect. Disord. 205, 319–326.

Lyubashina, O, Panteleev, S, 2009. Effects of cervical vagus nerve stimulation on amygdala-evoked responses of the medial prefrontal cortex neurons in rat. Neurosci. Res. 65, 122–125.

Mkrtchian, A, Aylward, J, Dayan, P, Roiser, JP, Robinson, OJ, 2017. Modeling avoidance in mood and anxiety disorders using reinforcement learning. Biol. Psychiatry 82, 532–539.

Niv, Y, Daw, ND, Joel, D, Dayan, P, 2007. Tonic dopamine: opportunity costs and the control of response vigor. Psychopharmacology 191, 507–520.

Palmiter, RD, 2007. Is dopamine a physiologically relevant mediator of feeding behavior? Trends Neurosci. 30, 375–381.

Palmiter, RD, 2008. Dopamine signaling in the dorsal striatum is essential for motivated behaviors. Ann. N. Y. Acad. Sci. 1129, 35–46.

Payzan-LeNestour, E, Dunne, S, Bossaerts, P, O'Doherty, JP, 2013. The neural representation of unexpected uncertainty during value-based decision making. Neuron 79, 191–201.

Peña, DF, Engineer, ND, McIntyre, CK, 2013. Rapid remission of conditioned fear expression with extinction training paired with vagus nerve stimulation. Biol. Psychiatry 73, 1071–1077.

Porges, SW, 1995. Cardiac vagal tone: a physiological index of stress. Neurosci. Biobehav. Rev. 19, 225–233.

Quetscher, C, Yildiz, A, Dharmadhikari, S, Glaubitz, B, Schmidt-Wilcke, T, Dydak, U, Beste, C, 2015. Striatal GABA-MRS predicts response inhibition performance and its cortical electrophysiological correlates. Brain Struct. Funct. 220, 3555–3564.

R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria Available at: https://www.R-project.org/.

Raedt, R, Clinckers, R, Mollet, L, Vonck, K, El Tahry, R, Wyckhuys, T, De Herdt, V, Carrette, E, Wadman, W, Michotte, Y, Smolders, I, Boon, P, Meurs, A, 2011. Increased hippocampal noradrenaline is a biomarker for efficacy of vagus nerve stimulation in a limbic seizure model. J. Neurochem. 117, 461–469.

Raudenbush, S, Bryk, A, Cheong, Y, Congdon, R, Du Toit, M, 2011. HLM 7 [Computer software]. Scientific Software International, Lincolnwood, IL.

Roosevelt, RW, Smith, DC, Clough, RW, Jensen, RA, Browning, RA, 2006. Increased extracellular concentrations of norepinephrine in cortex and hippocampus following vagus nerve stimulation in the rat. Brain Res. 1119, 124–132.

Sellaro, R, van Leusden, JWR, Tona, K-D, Verkuil, B, Nieuwenhuis, S, Colzato, LS, 2015. Transcutaneous vagus nerve stimulation enhances post-error slowing. J. Cogn. Neurosci. 27, 2126–2132.

Simmons, WK, DeVille, DC, 2017. Interoceptive contributions to healthy eating and obesity. Curr. Opin. Psychol. 17, 106–112.

Simonian, HP, Kresge, KM, Boden, GH, Parkman, HP, 2005. Differential effects of sham feeding and meal ingestion on ghrelin and pancreatic polypeptide levels: evidence for vagal efferent stimulation mediating ghrelin release1. Neurogastroenterol. Motility 17, 348–354.

Skrynka, J, Vincent, BT, 2017. Hunger Increases Delay Discounting of Food and Non-Food Rewards. Available at: https://psyarxiv.com/qgp54/ [Accessed October 10, 2018].

Steenbergen, L, Sellaro, R, Stock, A-K, Verkuil, B, Beste, C, Colzato, LS, 2015. Transcutaneous vagus nerve stimulation (tVNS) enhances response selection during action cascading processes. Eur. Neuropsychopharmacol. 25, 773–778.

Strandwitz, P, 2018. Neurotransmitter modulation by the gut microbiota. Brain Res. 1693, 128–133.

Suarez, AN, Hsu, TM, Liu, CM, Noble, EE, Cortella, AM, Nakamoto, EM, Hahn, JD, Lartigue, G de, Kanoski, SE, 2018. Gut vagal sensory signaling regulates hippocampus function through multi-order pathways. Nat. Commun. 9, 2181.

Swart, JC, Froböse, MI, Cook, JL, Geurts, DE, Frank, MJ, Cools, R, Ouden, HE den, 2017. Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. eLife Sci. 6, e22169.

Teckentrup, V, Neubert, S, Santiago, JCP, Hallschmid, M, Walter, M, Kroemer, NB, 2020. Non-invasive stimulation of vagal afferents reduces gastric frequency. Brain Stimul. 13, 470–473.

Tellez, LA, Medina, S, Han, W, Ferreira, JG, Licona-Limón, P, Ren, X, Lam, TT, Schwartz, GJ, Araujo, IE de, 2013. A gut lipid messenger links excess dietary fat to dopamine deficiency. Science 341, 800–802.

Veldhuizen, MG, Babbs, RK, Patel, B, Fobbs, W, Kroemer, NB, Garcia, E, Yeomans, MR, Small, DM, 2017. Integration of sweet taste and metabolism determines carbohydrate reward. Curr. Biol. 27, 2476–2485.

Vo, A, Hiebert, NM, Seergobin, KN, Solcz, S, Partridge, A, MacDonald, PA, 2014. Dopaminergic medication impairs feedback-based stimulus-response learning but not response selection in Parkinson's disease. Front. Hum. Neurosci. 8, 784. Available at: https://www.frontiersin.org/articles/10.3389/fnhum.2014.00784/full. [Accessed December 26, 2018].

Vo, A, Seergobin, KN, Morrow, SA, MacDonald, PA, 2016. Levodopa impairs probabilistic reversal learning in healthy young adults. Psychopharmacology 233, 2753–2763.

Waise, TMZ, Dranse, HJ, Lam, TKT, 2018. The metabolic role of vagal afferent innervation. Nat. Rev. Gastroenterol. Hepatol. 15, 625–636.

Yao, G, Kang, L, Li, J, Long, Y, Wei, H, Ferreira, CA, Jeffery, JJ, Lin, Y, Cai, W, Wang, X, 2018. Effective weight control via an implanted self-powered vagus nerve stimulation device. Nat. Commun. 9, 5349.

Yeomans, MR, Leitch, M, Gould, NJ, Mobini, S, 2008. Differential hedonic, sensory and behavioral changes associated with flavor–nutrient and flavor–flavor learning. Physiol. Behav. 93, 798–806.

Zagon, A, 2001. Does the vagus nerve mediate the sixth sense? Trends Neurosci. 24, 671–673.