Noradrenergic modulation of learning noise in value-based decision making

Published: 25-02-2021
Last updated: 08-04-2024

The goal of the proposed research is to investigate the behavioral, computational, and neural mechanisms by which a pharmacological manipulation of noradrenergic activity impacts value-based decision making and reinforcement learning. Based on the...

Ethical review	Approved WMO
Status	Recruitment stopped
Health condition type	Other condition
Study type	Observational non invasive

Summary

ID

NL-OMON51201

Source
ToetsingOnline

Brief title
Norepinephrine and decision-making

Condition

- Other condition

Synonym
healthy

Health condition

scientific investigation of healthy subjects

Research involving
Human

Sponsors and support

Primary sponsor: Universiteit Leiden
Source(s) of monetary or material Support: NWO

Intervention

Keyword: decision-making, norepinephrine, reinforcement learning

Outcome measures

Primary outcome

• Behavioural measures: choice history, outcome history, learning performance, choice reaction time, and choice variability

• Computational measures: model fit, estimated model parameters and variables

• Psychophysiological measures: task-related and task-unrelated changes in pupil size

• Brain activity measures: effect of atomoxetine on task-related evoked BOLD or EEG signal

Secondary outcome

•Psychometric measures: state anxiety, trait anxiety, emotional arousal (to be included as regressors for behavioural, fMRI and EEG data)

•Physiological measures: salivary cortisol and alpha amylase (to be included as regressors for behavioural, fMRI and EEG data)

Study description

Background summary

1.1 Value-based decision making and the explore-exploit dilemma

When faced with a choice between multiple options (actions), humans usually

base their decision on the expected values of the available options. The outcome of their decisions (i.e., a reward or punishment) can then be used to update the expected values of the options, thus enabling learning and the improvement of subsequent decision making. This iterative procedure of action selection, outcome perception and internal updating of expected values has been described and formalized in the reinforcement learning framework (Sutton & Barto, 2018). In the past few decades, a large amount of evidence has accumulated which suggests that animals and humans show hallmarks of reinforcement learning principles both on the behavioral as well as the neural level (Daw & O*Doherty, 2014; Dayan & Daw, 2008; Niv, 2009).
However, although reinforcement learning offers a promising approach to investigate decision making and learning, the specific mechanisms and principles that underly learning and value-based decision making in the brain are still far from being fully understood. For example, it is assumed that the goal of decision making is to maximize outcomes, whereas humans reportedly fail to achieve this goal and often make seemingly irrational (i.e. non-greedy) decisions (Lee, Zhang, Munro, & Steyvers, 2011). Especially in volatile environments, in which the objective value of the available options can rapidly change over time, this suboptimal choice behavior is pronounced. Prominent theories suggest that, despite being detrimental for short-term payoff, these non-greedy decisions serve the long-term purpose of outcome maximation. The trade-off between short-term and long-term payoffs has previously been described in the exploration-exploitation dilemma. Exploitation refers to behavior which achieves short-term maximation of outcome by greedily selecting actions with the highest expected value, whereas exploration refers to behavior with non-greedy action selection to search for options with the highest objective value. Following this idea, the basis of non-greedy action selection is information seeking, as options with low objective values are chosen less often and therefore are associated with high uncertainty about their current value. Therefore, explorative, and non-greedy action selection can increase long-term payoffs, suggesting that behavioral variability is a necessary corollary of optimal decision making in volatile environments.

1.2 The role of NE in value-based decision-making and learning

On the neural level, converging evidence suggests that exploration behavior and choice variability is accompanied by modulations of the locus coeruleus-norepinephrine (LC-NE) neuromodulatory system. Besides its role in the modulation of arousal (for a review, see Thiele and Bellgrove, 2018), an increase in noradrenergic baseline activity significantly contributes to choice variability, as suggested by the adaptive gain theory, an influential theory of LC-NE function (Aston-Jones & Cohen, 2005). However, there is little direct empirical evidence for this assumption. So far, studies using a pharmacological manipulation of the LC-NE activity have only yielded conflicting evidence. For example, in one study, an increase in tonic NE levels via the administration of reboxetine, an NE reuptake inhibitor, had no effect on choice stochasticity

(Jepma, te Beek, Wagenmakers, van Gerven, & Nieuwenhuis, 2010). In contrast, a different study reported pronounced modulations in exploration behavior, when healthy subjects were administered atomoxetine (Warren et al., 2017). However, these NE-driven modulations deviated from the a-priori hypothesis derived from the adaptive gain theory, calling this prominent model of NE mechanisms into question. Taken together, more research is needed to investigate and reevaluate the links between the LC-NE system and choice variability during value-based decision-making and learning.

1.3 A new computational model of learning noise and NE function

On the computational level, almost all models of value-based decision-making and reinforcement learning assume that behavioral variability arises solely from adjustments in the action selection process (i.e., during the translation from expected values to action probabilities). Recent evidence from our research team, however, suggests that while such randomness during action selection may account for a proportion of non-greedy decisions, a non-negligible proportion of decisions is better explained by a different mechanism. This additional source of behavioral variability stems from imprecisions in the sequential updating of expected values (Findling, Skvortsova, Dromnelle, Palminteri, & Wyart, 2019). In the model, this imprecision during value updating is realized via so-called learning noise which internally accrues with every (neural) computation. Based on previous evidence for internal noise in perceptual inferences (Drugowitsch, Wyart, Devauchelle, & Koechlin, 2016), the authors suggest the possibility of internal noise also during value updating, where it corrupts action values. On the behavioral level, the effects of learning noise in value updating closely resembles exploratory action selection, although both they arise from distinct mechanisms. Interestingly, a (noisy) reinforcement learning model which incorporates both sources of choice variability (stochasticity during action selection and learning noise during value updating) outperforms classical (i.e. exact) reinforcement learning models without learning noise in a modified version of the widely-used restless bandit task (Findling et al., 2019). Moreover, by identifying the neurophysiological correlates of learning noise, this work also provided preliminary evidence for a putative connection between noradrenergic activity and the precision of learning. It was shown that learning noise correlated with both pupil size (i.e. a well-validated indicator of LC-NE activity) and BOLD fluctuations in brain regions with bidirectional interaction with the LC (Findling et al., 2019).
In the present project, we seek to test this idea that the LC-NE system controls for learning precision. To investigate this novel hypothesis, we have already reanalyzed existing pharmacological dataset from our group (Jepma et al., 2010). Again, the newly developed noisy reinforcement learning model outperforms the exact model when fit to the behavioral data. Furthermore, computational modelling suggests that reboxetine, an NE reuptake inhibitor, leads to an increase in learning noise but a decrease in stochasticity. This

finding is especially interesting as the computational results obtained from an exact instantiation of the reinforcement learning model yielded inconclusive results about the effect of reboxetine on behavior. In sum, this reanalysis strengthens our assumption that the brain is subject to learning noise and that this learning noise is conveyed by noradrenergic modulations. However, due to the design of the task, some questions could not be answered sufficiently. In the present research we are therefore interested in extending the findings from the reanalysis, and directly investigate the effects of atomoxetine on the behavioral and neural level.
According to the idea of noisy reinforcement learning, learning noise is not explicitly represented anywhere in the brain. However, by employing a careful experimental design, its putative contribution to behavior and impact on neural processing can be delineated. Therefore, we seek to employ the same behavioral task used in the study of Findling and colleagues (2019) and extend that study by using a double-blind, placebo-controlled, within-subject pharmacological manipulation: a single oral dose of the selective NE transporter blocker atomoxetine. In two separate experiments, we will examine fMRI BOLD and EEG activity in combination with pupillometry to investigate if our noisy reinforcement learning model for different levels of learning precision captures and predicts the patterns evoked by our pharmacological manipulation. Combining the subtlety of the behavioral task design, the power of a pharmacological manipulation and the rigor of computational modelling puts us in a good position to obtain a much richer picture of the causes and effects underlying behavioral variability and the role of the LC-NE systems therein.

Study objective

The goal of the proposed research is to investigate the behavioral, computational, and neural mechanisms by which a pharmacological manipulation of noradrenergic activity impacts value-based decision making and reinforcement learning. Based on the idea that LC-NE system regulates the precision of value-based learning, we set three specific objectives:
1. replicate and extend recent findings on the impact of a pharmacologically increased NE level regarding value-based learning and decision making.
2. fit computational models of reinforcement learning and decision-making to the behavioral data and extract the central model parameters (e.g., learning noise and choice variability) that account for the observed behavioral differences across treatment conditions.
3. Identify brain regions and EEG components implicated in the modulation of pharmacologically manipulated NE levels and investigate their role in the regulation of learning precision and choice variability using model-based analyses.

Study design

The proposed research consists of two separate studies, which are identical in

study design but differ regarding their neural measure.

Both studies will use a double-blind, placebo-controlled, cross-over design. Each study consists of a pre-screening interview and two testing sessions that are set one week apart from each other. The participants receive one pill (placebo or atomoxetine) per session. We will look for the within-subject effects of atomoxetine on brain and pupil signatures during sustained value-based learning and decision-making. Therefore, participants will be scanned during the well-validated canonical, restless two-armed bandit task, in which the participants* goal is to maximize their monetary payoff by sequentially sampling from one of two independent reward sources. Two separate conditions of the task will be employed: In the full outcome condition outcomes for both the chosen and the forgone action are presented. In the partial outcome condition only the outcome for the chosen but not the forgone action is presented. Task conditions will be counterbalanced across participants.

Drug intervention
Participants will receive on one occasion 40 mg of the selective NE transporter blocker atomoxetine (Navarra et al., 2008), orally administered. The 40 mg dose is a typical starting dose used in clinical practice that avoids reported side effects of increased heart rate at high atomoxetine doses (Heil, et al., 2002). In the other session, either one week earlier or one week later, participants will receive a placebo pill (125 mg of lactose monohydrate with 1% magnesium stearate), visually identical to the drug.

General procedure
The proposed studies will consist of two sessions of fMRI/EEG and behavioral data collection during the bandit tasks. The sessions are scheduled one week apart at the same time of day. Each subject will perform the tasks in the MRI scanner/EEG chamber under the influence of atomoxetine in one session, and under the influence of a placebo in the other. Study 1 will start in the LUMC (Radiology department) in a behavioral testing room and move to the fMRI room about 65 minutes after the subject first arrives. Study 2 will start at the Pieter de la Cour building in the EEG lab. With breaks, time for task training, for the drug to take effect and for moving between locations, each session will last approximately three and a half hours. Participants will be administered the drug 90 minutes before the first set of task blocks to ensure that tasks are performed during peak blood levels (Chamberlain, Müller, Blackwell, Robbins, et al., 2006). The final functional scan or electrophysiological recording will be completed about 3 hours after taking the drug, within the window of time when the drug should still be having an effect on cognition (Sauer, Ring, & Witcher, 2005). Total scanning time will constitute approximately 95 minutes in each session.

Study burden and risks

Atomoxetine

A single dose of atomoxetine has not been reported to have long-lasting effects, either adverse or beneficial. Previous studies using single dosages of 40-60 mg, including two 40-mg studies conducted in our group (P13.026 and P13.282), show that this was well tolerated by healthy volunteers. Short-term side effects of the drug in a dosage of 40 mg in healthy volunteers are mild and typically include fatigue, increased heart rate and dry mouth, which have been shown to disappear around 2 hours after drug ingestion (Chamberlain, Müller, Blackwell, Clark, et al., 2006; Chamberlain, Müller, Blackwell, Robbins, & Sahakian, 2006). For some groups, use of atomoxetine does carry risk for more serious side-effects: individuals with glaucoma, with heart disease, or taking monoamine oxidase inhibitors (MAO inhibitors). We will only include subjects in excellent physical health who are not using psychotropics.

fMRI
There are no known risks associated with participating in an fMRI study. This is a noninvasive technique involving no catheterizations or introduction of exogenous tracers. Numerous human subjects have undergone magnetic resonance studies without apparent harmful consequences. Radiofrequency power levels and gradient switching times used in these studies are within the FDA-approved ranges. Some people become claustrophobic while inside the magnet and in these cases the study will be terminated immediately at the subject's request.

EEG
There are no known risks associated with participating in an EEG study. This is a noninvasive technique involving no catheterizations or introduction of exogenous tracers. Numerous human subjects have undergone electrophysiological studies without apparent harmful consequences. Some people become claustrophobic while inside the EEG chamber and in these cases the study will be terminated immediately at the subject's request.

Pupillometry
The eye-tracker system uses detailed analysis of high-definition video to record pupil diameter at any given time during the experiment. The subjects do not have to wear any special apparatus for the eye-tracker to work, and are at no significant risk of any type of injury or discomfort due to this aspect of the experiment.

Contacts

Public
Universiteit Leiden

Wassenaarseweg 52

Leiden 2333AK
NL
Scientific
Universiteit Leiden

Wassenaarseweg 52
Leiden 2333AK
NL

Trial sites

Listed location countries

Netherlands

Eligibility criteria

Age
Adults (18-64 years)
Elderly (65 years and older)

Inclusion criteria

Healthy adult subjects with no history of neurological disorder/disease and no counter-indications to 3 Tesla MRI, EEG or to atomoxetine, and no personal relationship with the researchers will be included in this study. All participants will be right-handed with normal vision or contact lenses.

Exclusion criteria

Significant history of head trauma, premature birth, learning disabilities, neurological or psychiatric illness. Heart arrhythmia, glaucoma, congenital eye diseases, hyperopia, myopia, hypertension and use of antidepressants or psychotropic medication and possible pregnancy (in adult females). MRI contra-indications, including metal implants and claustrophobia. Smoking more than five cigarettes a day - to avoid nicotine withdrawal effects during the study. Alcohol consumption < 24 hours before study, caffeine consumption < 3 hours before study.
These criteria will be assessed by a self-report questionnaire administered during pre-screening.

Study design

Design

Study type:	Observational non invasive
Intervention model:	Crossover
Masking:	Double blinded (masking used)
Control:	Uncontrolled
Primary purpose:	Treatment

Recruitment

NL

Recruitment status:	Recruitment stopped
Start date (anticipated):	10-05-2021
Enrollment:	60
Type:	Actual

Ethics review

Approved WMO

Date:	25-02-2021
Application type:	First submission
Review commission:	METC Leiden-Den Haag-Delft (Leiden) metc-ldd@lumc.nl

Approved WMO

Date:	18-05-2021
Application type:	Amendment
Review commission:	METC Leiden-Den Haag-Delft (Leiden) metc-ldd@lumc.nl

Study registrations

Followed up by the following (possibly more current) registration

No registrations found.

Other (possibly less up-to-date) registrations in this register

No registrations found.

In other registers

Register	ID
CCMO	NL75588.058.20