INVESTIGATING THE INFLUENCE OF AROUSAL ON BELIEF UPDATING UNDER UNCERTAINTY

Masters Thesis

submitted to

Indian Institute of Science Education and Research, Pune in partial fulfilment of the requirements for the BS-MS Dual Degree Programme

by

Saanchi Thawani

Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,

Pashan, Pune 411008, INDIA.

Date: April, 2023

Under the guidance of

Prof. Dr. Tobias H. Donner,

Department of Neurophysiology & Pathophysiology,

University Medical Center Hamburg

From May 2022 to Mar 2023

CERTIFICATE

This is to certify that this dissertation entitled 'Investigating The Influence Of Arousal On Belief Updating Under Uncertainty' towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Saanchi Thawani at University Medical Center-Hamburg under the supervision of Prof. Dr. Tobias H. Donner, Professor of Integrative Neuroscience, Department of Neurophysiology & Pathophysiology, during the academic year 2022-2023.

Prof. Dr. Tobias H. Donner

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

DECLARATION

I hereby declare that the matter embodied in the report entitled Investigating The Influence Of Arousal On Belief Updating Under Uncertainty are the results of the work carried out by me at the Department of Neurophysiology & Pathophysiology, University Medical Center- Hamburg, and Department of Biology, Indian Institute of Science Education and Research, Pune, under the supervision of Prof. Dr. Tobias H. Donner, Department of Neurophysiology & Pathophysiology, University Medical Center-Hamburg and the same has not been submitted elsewhere for any other degree.

Saanchi Thawani
April, 2023

ABSTRACT

For decision making in environments with hidden changes in context and consequent uncertainty, it is suboptimal to perfectly accumulate evidence. Instead, an ideal strategy involves a non-linear belief updating process that is sensitive to these unexpected changes and adapts according to the changing statistics of the environment. Brain states described by fluctuating levels of arousal dictate variability in information processing. Neuromodulators such as phasic norepinephrine can reconfigure large-scale neural circuits of the decision machinery, but are they involved in mediating flexible decision-making? Building on what we know about neural mechanisms underlying behaviour that also cause changes in pupil size, what can we infer about these flexible mechanisms that are typically inaccessible, from measurements of the pupil? We use pupillometry to obtain pupil dilation as a proxy for arousal and employ a non-linear normative model to infer latent parameters of the belief updating process in healthy human participants performing a hierarchical decision task involving uncertainty. We find that behavioural measures of change point detection and uncertainty are encoded in the dynamics of pupil diameter on fast and slow time scales, respectively. Our findings implicate the involvement of arousal systems of the brainstem in the belief-updating process, whereby arousal serves to facilitate flexible decision-making in uncertain environments.

ACKNOWLEDGMENTS

I thank my research supervisors, Prof Dr Tobias Donner, Dr Rudy van den Brink and Dr Joshua Calder-Travis. Without their constant assistance and dedicated guidance at every step, this research would have never been accomplished. I am deeply grateful to them for giving me this opportunity and sharing their knowledge with me.

I would also like to show gratitude to my TAC co-supervisor, Raghav Rajan and, the Committee for acknowledging my work. All professors who have taught me and interacted with in the last four years at IISER and fueled my curiosity and exploration.

Thanks to Frau Hilke Peterson for handling all administrative details and arranging the possibility for my work in the UKE. Getting through my thesis required more than academic support, and I have a huge support system to thank for their friendship, motivation, and welcoming me into the Lab as family. Rudy van den Brink, Joshua Calder-Travis, Hame Park, Ayelet Arizi, Gina Monov, Josefine Hebisch, Alessandro Toso made the lab a positive place to be in. I also have to thank my friends, who made my stay in Hamburg memorable and like home.

Most importantly, my family, whose hard work and support made this journey feel less rocky.

Thanks to Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—SFB 936—178316478—A7, and Z3 and Prof Dr Tobias Donner for funding this project.

CONTRIBUTIONS

[In Alphabetical Order]

Contributor name	Contributor role
Joshua Calder-Travis, Rudy van den Brink, Tobias Donner,	Conceptualization Ideas
Joshua Calder-Travis, Rudy van den Brink, Tobias Donner,	Methodology
MATLAB, Python	Software
Rudy van den Brink	Validation
Joshua Calder-Travis, Rudy van den Brink, Saanchi Thawani	Formal analysis
Joshua Calder-Travis, Rudy van den Brink, Saanchi Thawani	Investigation
Joshua Calder-Travis, Rudy van den Brink, Saanchi Thawani	Resources
Joshua Calder-Travis, Rudy van den Brink, Saanchi Thawani	Data Curation
Saanchi Thawani	Writing - original draft preparation
Rudy van den Brink	Writing - review and editing
Saanchi Thawani	Visualization
Rudy van den Brink, Joshua Calder-Travis	Supervision
Joshua Calder-Travis, Rudy van den Brink, Tobias Donner	Project administration
Tobias Donner, Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—SFB 936—178316478—A7, and Z3	Funding acquisition

INTRODUCTION

Fundamentally, the world around us is uncertain, and we accumulate noisy sensory input in order to form a belief about the state of the world. To make the best estimates about the external states of this uncertain environment in its internal model, our brains employ probabilistic reasoning [1]. Decision-making is thus a temporally unfolding process of statistical inference, where if the sensory evidence that is coming in deviates from our currently held belief estimate, the belief is updated. This serves a normative purpose: reducing uncertainty to improve the accuracy of decisions.

In cognitive science, decision-making is typically conceptualised with sequential sampling models wherein discrete quantities representing sensory evidence are accrued to produce a decision variable which results in a choice depending on the rule.
In normative models of decision-making, Bayes' rule defines the fundamental basis belief formation in terms of Log Posterior Ratio as the Belief:

$$P(h|e) = \frac{P(e|h)\,P(h)}{P(e)}$$

The prior *P(h)* is the probability that hypothesis *h* about the state of the world is true before obtaining any evidence *(e)* about it. *P(e|h)* is the "likelihood" of obtaining an evidence e when h is true [2]. The conditional probability: posterior belief *P(h|e)* for each observed sample is obtained by re-weighting of each hypothesis by likelihood *P(e|h)* subject to how well it predicts the observed data.

Simple heuristic models include the sequential probability ratio test (SPRT), wherein the logarithm of the likelihood $log(LR)s$ associated with each piece of evidence is summed as $log(LR_{12}) \equiv log\,\frac{P(e_1, e_2, e_3 \ldots e_n \mid h_1)}{P(e_1, e_2, e_3 \ldots e_n \mid h_2)} = \sum_{i=1}^{n} log\frac{P(e_i \mid h_1)}{P(e_i \mid h_2)}$ until it reaches a defined bound Z_1 *or* Z_2 supporting either hypothesis [3–5]. Similarly, the

drift-diffusion models including different rates of accumulation of evidence over time to reach the decision bound, and related sequential-sampling models are forms of 'belief-updating' rules formulated on perfect integration over time (that is, each evidence belief is given equal weight) [5,6]. However, these models are optimal only under specific conditions: (I) source of incoming evidence remains constant during decision formation to be able to give equal weight to each belief, i.e., uncertainty or unexpected changes in environment are absent; (II) the noise corrupting the evidence is the only source of uncertainty for the observer. However, this is unlike the real world where uncertainty also originates from natural environments undergoing hidden unexpected changes in their state [7–9].

Other models like the leaky integrator allow the downweighing of evidence ("leak"), but require knowledge of the temporal structure of evidence and a stationary evidence source [10,11] . This temporal uncertainty about the change is usually unknown to the observer. Secondly, recurrent and non-linear cortical circuit dynamics contrast the linear dynamics of the perfect integration process [12,13].

Therefore, an optimal strategy involves a non-linear evidence accumulation process that is sensitive to these unexpected changes and adapts to the changing statistics[7,14]. There must be a balance in the operations such that when perfect stability is expected, evidence is accumulated perfectly, however, if instability is expected, evidence is accumulated with a leak to a non-absorbing boundary [4,14]. This expectation-dependent dynamics of the leak and decision boundary, proposed by Glaze et al. 2015 facilitates changes in the accumulation process for the identification of unpredictable changes in the environment [14] .

Cognitively, decisions transform sensory signals into motor actions in a flexible, context-dependent manner [15–17]. Human and non-human primates have been shown to be able to perform hierarchical decision-making tasks in volatile environments. In a study by Sarafyazd and Jazayeri 2019, participants have been shown to be able to infer changes in the sensory-motor mapping rule from negative feedback after judgement on a lower-order decision [18,19]. This behaviour is explained by models which integrate previous choice outcomes (from feedback after lower order decision is made) about

unreliable stimuli with an expected accuracy of judgements (choice confidence) to infer rule switches when the decision variable reaches a bound [18].

van den Brink et al. 2022, employ a variant of this task where uncertainty is restricted to the inference of active rule (higher order decision) instead, and trial-by-trial feedback is not provided [20]. Consider the following hierarchical perceptual choice task: a visual orientation discrimination task governed by a mapping rule such that vertical orientation of a visual grating is reported by button press with the left hand and horizontal orientation with the right hand, or the converse for the second mapping rule (Figure 1). This sensory-motor task required participants to continuously infer the currently active rule by accumulating noisy evidence (higher-order decision), which could underdo hidden state changes, and then apply the selected rule to report their orientation judgment (lower-order decision) (Figure 4).

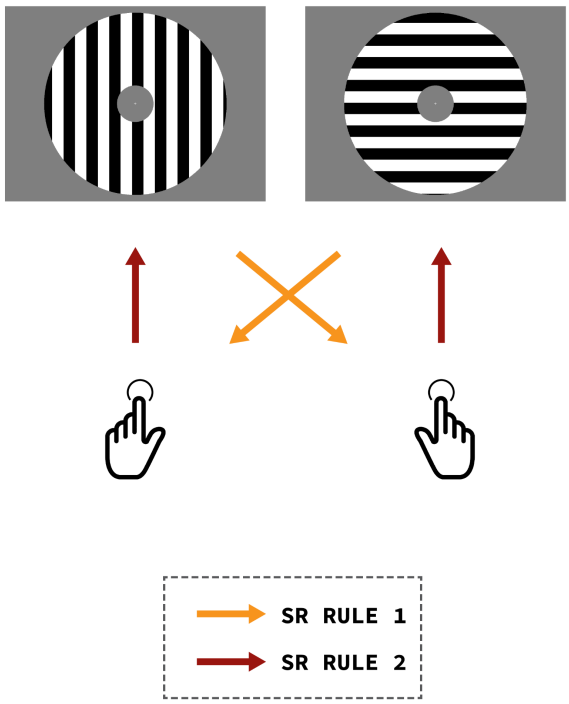

Figure 1. Representation of Stimulus to Response (SR) Mapping Rules

When rule one (SR Rule 1) is active, the vertical orientation grating is indicated with a right hand button press, and

the horizontal grating with a left hand button press. This mapping is reversed when rule two (SR Rule 2) is active. Orange, SR-Rule 1; Red, SR- Rule 2.

Because the higher-order stimulus involved unpredictable hidden changes, perfect accumulation of all the evidence is suboptimal and an ideal strategy requires a balance between stable evidence accumulation and sensitivity of belief to adapt to the volatile environment. The authors find evidence that populations of neurons encoding the stimulus features are spontaneously coupled to the neurons that encode the appropriate response, thus mapping response to stimulus in a context-dependent fashion, to facilitate the flexible functioning of the executive control system in decision making. A reversal in rule/context governing the sensory-motor map is reflected in a flip in sign of correlational activity between the visual-motor regions.

But how are networks flexibly reconfigured to achieve this context dependant belief-updating [21,22] ? One idea is that neuromodulators such as phasic norepinephrine reconfigure, i.e. "interrupt" or "reset" large-scale neural circuits of the decision machinery to establish flexible information flow to update beliefs in unexpected changes within a task [23–25]. Overall brain states described by fluctuating levels of alertness or excitability, referred to by the term arousal are thought to dictate variability in information processing and decision-making. These fluctuating arousal levels are predominantly shaped by neuromodulators from the brain-stem nucleus locus coeruleus (LC), which supplies norepinephrine (NE)[26,27]. Dampened arousal leads to drowsiness and heightened arousal due to events like unexpected stimuli can increase alertness. Here, we investigate if neuromodulators shape this coupling in such a hierarchical task.

There are two different functional modes of LC-NE system influencing arousal (through the release of NE) - phasic and tonic which operate on distinct timescales [28–30]. In the phasic mode, bursts of LC activity are associated with task-relevant decisions and behavioral responses induced by unexpected or surprising events [31]. While in the tonic mode along longer timescales, baseline activity is elevated and behavior is more exploratory or distractible. These fluctuations are hypothesized to co-vary with

unexpected uncertainty, like unsignaled context switches [32]. How can we track these neuromodulator-induced arousal responses to uncertain environmental changes?

The Pupil is known to dilate or constrict with varying ambient light conditions through a process known as the pupillary light reflex (PLR), however, the pupil diameter is modulated also in non-luminance mediated conditions, namely cognitive events like attention, uncertainty, working memory, and other arousal-related instances [33–35]. Brain circuits that drive pupil diameter modulation in response to changes in luminance also receive inputs that drive cognitive modulations of pupil size. This modulation occurs through three nuclear pathways:

1) Pretectal Olivary Nucleus (PON): It receives direct, ascending input for high-order visual-saccadic processing. Indirect evidence from lesion studies suggests that high-order information about visual perception and attention arrives at the PON from higher cortical visual pathways (frontal eye field (FEF) in the prefrontal cortex and the lateral intraparietal area (LIP) in the parietal cortex) to affect the PLR [36].

2) Superior Colliculus (SC): The SC, a specialized motor nucleus controls pupil size with both direct and indirect pathways by integrating a wide range of both sensory and associative signals. In a study by Wang et al. 2012, those layers of the SC which receive sensory, motor, and cognitive information from cortex modulated pupil size in low-light conditions possibly enhancing visual sensitivity or attention [37]. Wang et al. 2014 showed that audio and visual stimuli evoked pupil responses reflecting the involvement of SC [38]. Microstimulation studies indicate the potential involvement of multiple pathways influencing the effects of SC and either dilations/increase or constrictions/decrease of the pupil size in different conditions and at different timescales [33,37,38].

3) Locus Coeruleus (LC): The LC, the primary source of central NE which regulates global arousal has been more extensively researched in its involvement of pupil size modulation. Evidence in the form of covariance of pupil size, LC activity and global

arousal from electroencephalography (EEG), event-related potentials (ERPs), blood-oxygen-level dependent (BOLD) activity and results from pharmacological manipulations and microstimulation of the LC-NE system have led to pupil measurements being interpreted in terms of LC neural activation [26,39–41]. These studies conducted in various species: mice, monkeys and humans, in a variety of behavioural tasks using different measures of brain activity highlight the strong reliability of LC–pupil relationships [42].

Due to its sensitivity to stimulus probabilities, such that it dilates during the presentation of unexpected stimuli to capture critical computations to learn from, and adapt to changing environmental conditions, pupil size is associated with brain-state arousal [43].

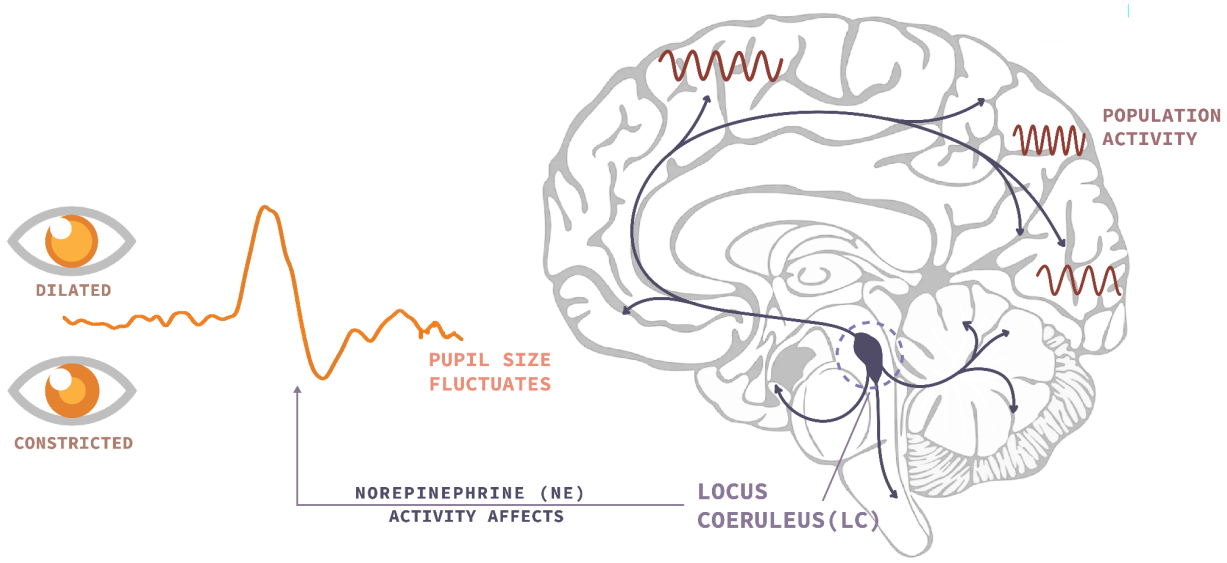

Figure 2. Using Pupil Size as a proxy for Brainstem modulation of cortical activity: Arousal

Pupil size fluctuates spontaneously even at rest . The locus coeruleus (LC) in the brainstem releases the neuromodulator Norepinephrine (NE) at cortical sites causing fluctuations in arousal. Since the LC-NE system also affects pupil response, we use pupil size as a proxy for arousal-related activity.

[Representation inspired by Pfeffer, Keitel, et al. 2022]

By virtue of this correlation between LC activity and pupil size, and what we know about the involvement of the LC-NE system in modulating arousal, we use pupil as a proxy to infer arousal induced behavioural effects in a heirarchical task governed by unexpected context switches.

To isolate cognitive effects from light effects on the pupil, we conduct Pupillometry to detect and track changes in pupil diameter and gaze positions, in isoluminant conditions. We combine pupillometry, and computational modelling to understand the interplay between uncertainty-related arousal and decision computations underlying flexible sensory-motor choice tasks [24,44,45]. We investigate the usability of the non-linear model of evidence accumulation by Glaze et al. 2015 which employs a non-linear transformation of the posterior belief from each encounter with evidence into the prior belief for the next step to understand behaviour in volatile contexts, and probe the encoding of behavioural variables generated from fitting this model to data, in fluctuations of pupil data using linear regression . We find that terms encoding changes in environmental states covary with changes in pupil diameter, supporting the possible involvement of phasic arousal shaped by norepinephrine(NE) in belief-updating in volatile environments.

METHODS

The main analyses of this thesis were preregistered [63]. The preregistration will be made public upon publication of the results.

Experimental Design

Sample, in- and exclusion criteria

For well-balanced randomization and based on sample sizes used in related experiments, we set a target of 24 participants for our study [20]. However, due to technical issues with the MEG and COVID events, a sample size of 19 participants was established, which is usual in typical and successful (MEG) studies [20,46,47]. To minimise inter-individual variability of dependant variables in our paradigm: MEG variables, pupil data parameters, and behavioural model variables, all experimental manipulations: stress induction and pharmacological interventions were performed within-subject. Furthermore, within each condition of the experiment, extensive data were collected from each participant (with data from the tasks of 2.5 hours/day over four days) to yield precise individual estimates of the above-mentioned dependant variables.

Each participant was paid approximately 411 EUR for approximately 14 hours of participation. The participation was conditional to an extensive set of exclusion and inclusion criteria approved by the ethics committee of the Hamburg Medical Association (Ärztekammer Hamburg). This included that participants are between 18 and 40 years old, have no physical, neurological, or psychiatric illnesses, have no relevant allergies, are not pregnant, and have not taken medication in the previous two months. Contraindications for the use of Prazosin, the alpha-adrenergic blocking drug used for pharmacological manipulation were also a part of the exclusion criteria. Participants were also excluded if they had MRI-incompatible objects in their body, or if they had insufficient visual acuity (without glasses) at a 60 cm distance.

Overview of study

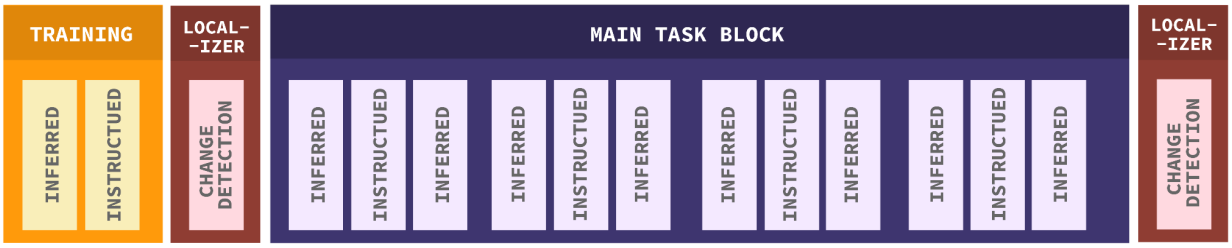

Figure 3. Task Structure for Main Behavioural Sessions

The tasks in the study spanned five sessions per participant. The first session, served to give the participants information about what the study would entail with the task instructions and let them practice the task in a behavioral psychophysics laboratory. This session also helped the participant decide if they would be able to handle the stress manipulations to continue with the main experimental sessions in the following days.

Each participant then underwent four main experimental sessions within the MEG laboratory while their pupil data was recorded with an infrared pupillometer (Figure 3). At the beginning of each experimental session either 0.5 mg Prazosin (α1-receptor antagonist to reduce availability of norepinephrine in the frontal cortex) or a placebo (visually identical to the drug) is administered orally twice in a randomized and double-blind manner. The first non-training trial of the main behavioural task is started only after at least 55 minutes following administration of the drug, to consider the average time for the drug to metabolise and come into effect. Within the main task block, between each set of inferred-instructed block, the participants were made to immerse their hand into a bucket of water, uptil their wrist. The water was either of cold temperature between 0℃ and 0.5℃ which is known to induce tonic availability of norepinephrine(NE), or room temperature, as placebo. The temperature level was constant per session, such that there were two sessions involving stress, and two with placebo. Blood pressure was monitored throughout to ensure safety of the participant after drug administration. This was done following the rationale that if by using the drug we can 'rescue' behaviour (that is, reverse the effects of the stress manipulation), then that is strong evidence for the selective role of NE in causing behavioural effects involved in belief updating. [The drug and stress manipulations were beyond the scope of investigation for the thesis]

Finally, each participant was invited for a separate interview for clearance and consent for MRI measurements, followed by structural MRI measurements (T1-weighted

MPRAGE scan) on another day, to be used for MEG source reconstruction, which was also beyond the scope of the current thesis.

Apparatus

The task is presented using MATLAB and Psychtoolbox-3 running on Linux machines at 60Hz (120Hz during training session; Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). During the training, session stimuli are presented on a VIEWPixx monitor, while during the MEG sessions stimuli are presented via projection onto the back of a transparent screen (PROPixx VPX-PRO-5050B projector), which is approximately 60 cm from participants (70 cm for the first four participants). MEG recordings (at 1200 Hz) take place in a room with magnetic shielding, using a CTF device with 275 axial gradiometers. During MEG sessions eye tracking data is acquired using an SR Research EyeLink 1000 device (at 1000 Hz), a non-invasive infrared pupillometry device. This pupillometer emits a beam of infrared light onto the eye which is partially absorbed by the iris and partially reflected back by the pupil; the reflected light hits the detector of the pupillometer which converts it into an electrical signal which is then processed into size measurements.

Behavioural Task

19 participants performed two different versions of a hierarchical decision-making choice task. The main task of this study being a variant of that used by van den Brink et al. 2022 [20]. In this hierarchical task, the selection of a changing sensory-motor (SR) mapping rule (higher-order decision) was to be used to make basic visual orientation discrimination judgment (lower-order decision) as in Figure 4.

The sensory-motor mapping rules are two different stimulus-response mapping rules, which vary throughout the course of each experimental session and block. Under "SR-rule 1", the horizontal orientation (visual stimulus) is to be reported with a left-hand button press (motor response) , and vertical with the right-hand button press. Under "SR-rule 2" the appropriate stimulus-response mappings are reversed. Only one of the two stimulus-response mapping rules is "active" at a given moment. Trial-by-trial

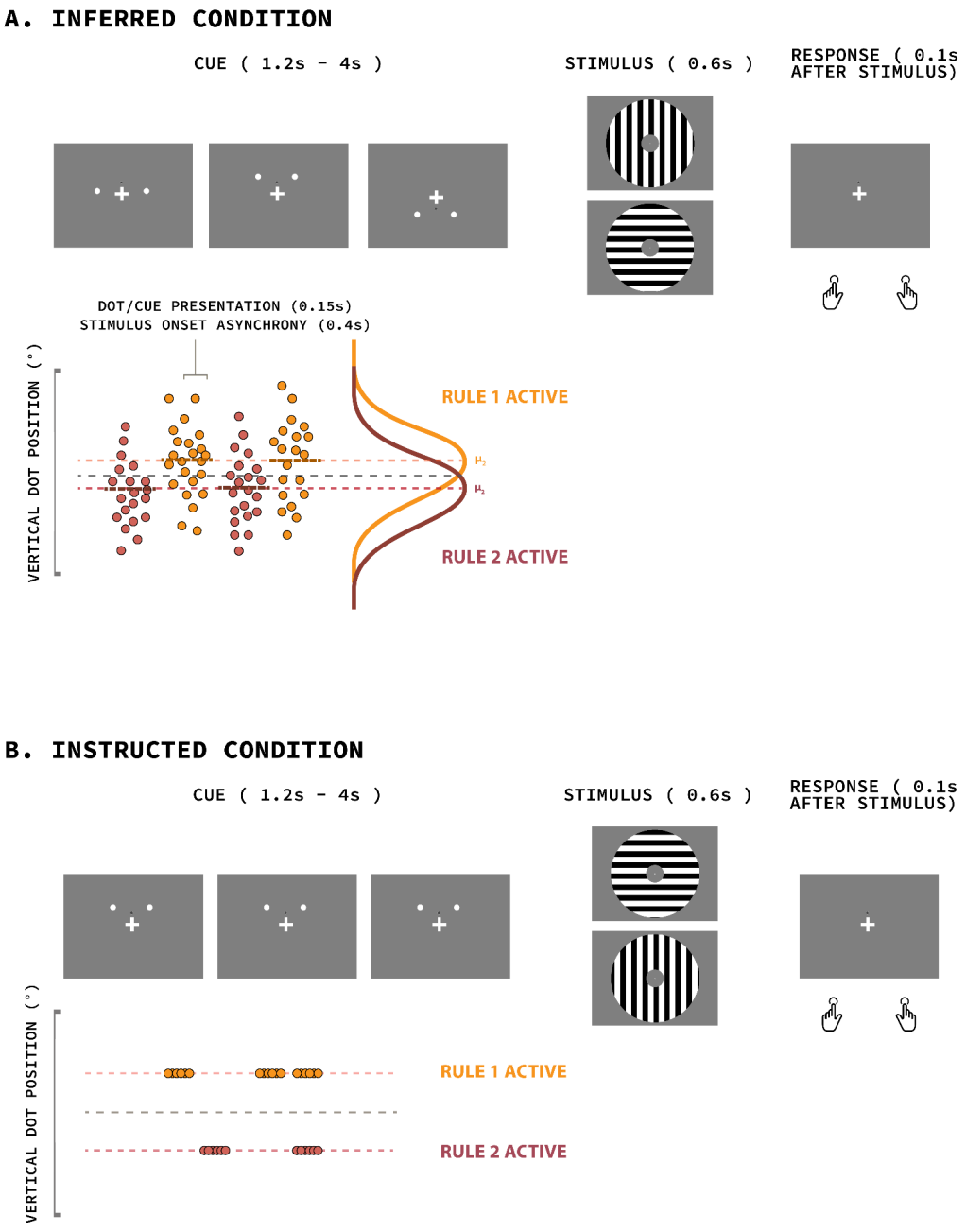

Figure 4. Example sequence of trial events

[Orange: SR-Rule 1, Red: SR- Rule 2, μ_1: Mean of generative gaussian distribution for cues indiciating active rule 1 centered above horizontal axis, μ_2: Mean of generative gaussian distribution for cues indiciating active rule 2 centered below horizontal axis]. + : Fixation Cross

(A) Inferred Rule Task. Top, an example of evidence samples displayed on a screen to participants for higher order decision during the inter trial interval preceding the visual discrimination task for lower order decision submitted with button press. Bottom, representation of the two truncated gaussian distributions from which dot positions are drawn to produce a noisy evidence stream

(B) Instructed Rule Task. Top, example of evidence samples for higher order decision displayed to participants during inter trial interval with active rule 1 as all dots are above the horizontal axis, followed by visual discrimination task for lower order decision to be submitted with button press. Below, Dot positions are unambiguous and either above or below the horizontal axis. Here, Rule 1 is active.

feedback is not provided, but at the end of each block participants are informed of the correct reponse percentage for the completed trials.

The amount of uncertainty that is associated with the higher-order decision regarding the currently active rule was varied as follows: a) The "Instructed rule" condition in which the participant is told explicitly what the active rule is on a given trial, the selection of rule was unambiguous and predictable (Figure 4(B)) and b) The "Inferred rule" condition where the active rule is uncertain and needs to be inferred from from the noisy sensory evidence samples which could undergo unexpected changes, thus making it ambiguous and more challenging (Figure 4(A)).

In the higher-order decision, visual cues are presented in the form of flashing dots sequences in rapid succession during the inter-trial-intervals (multiple of 0.4s from 1.2 - 4s inclusive, uniformly distributed). The cues are the locations (polar angles) of pairs of small dots presented simultaneously (duration: 0.15s, stimulus onset asynchrony: 0.4s) in the left and right visual hemifields at fixed eccentricity and polar angles that vary with respect to the horizontal axis in the left and right visual hemifields. The positions of

both dots in a pair are identical about the vertical meridian so that all information is contained in the polar angle of one dot and the other provides redundant information. Thus, each pair of dots refers to one cue.

The participants were instructed at the beginning of each block, which distribution corresponded to which rule and this relationship stayed constant throughout all experiment sessions. To summarise, the higher-order decision involved monitoring a stream of dots flashing around the vertical meridian, and inferring which distribution they belonged to, either a gaussian distribution centred above the horizontal axis or below it, to determine the active rule for the second-order decision. In the instructed condition, this was signaled without uncertainty, and in the inferred version, under uncertainty. The stimuli for the lower-order decision (the choice gratings) were large and full-contrast, and thus they themselves had no uncertainty associated with them. Therefore, they probably didn't limit performance on the task as seen in the results of the instructed version of the task in van den Brink et al. 2022, where participants had almost 100% accuracy. So the only form of uncertainty that contributes to formation of the final decision is the uncertainty that is contained in the cues that signal the rule in the inferred rule task[20]. Thus, the correctness of the response depended predominantly on selection of the currently active rule.

Formulation of the Task

All stimuli were created using Matlab and the Psychophysics Toolbox Version 3 and presented on a medium grey background. The fixation mark at the centre of the screen was a white symmetric cross with a length of 0.51° of visual angle and a thickness of 0.05°. The polar angle (measured clockwise from horizontal) of the right-hemifield dot, for the *n*th cue in the block, is referred to with $x_{n.}$ The condition which applies: Instructed or Inferred, changes with each block of 34 trials.

Each trial involved a sequence of cues ranging between 3-10, each presented for 150ms interspaced with 250ms intervals (between the end and start of presentation of

successive cues). At random intervals during ongoing cue monitoring, participants were asked to make a simple visual orientation judgment using the inferred rule.

In the instructed condition (Figure 4(B)) : If the active rule is "SR-rule 1", then the pair of dots is repeatedly flashed at a fixed polar angle, 12 degrees below horizontal in the lower visual hemifield, and if the currently active rule is "SR-rule 2" then the pair of dots is repeatedly flashed at 12 degrees above horizontal in the upper visual hemifield. Because of the fixed cue positions, there is no uncertainty about the active rule to be applied in each trial in this condition.

In the "inferred" condition (Figure 4(A)), where the positions vary stochastically from cue to cue, the active rule is determined from the (truncated) Gaussian distribution from which the cues are drawn. Specifically, if the state (i.e., active rule) at the nth cue, S_n, is j (where $j = 1$ indicates "SR-rule 1" and $j = 2$ indicates "SR-rule 2") then the untruncated distribution is given by

$$p(x_n|S_n = j) = N(x_n; \mu_n, \sigma^2)$$

where $N(a; b, c)$ indicates a normal distribution with mean b and variance c truncated at the vertical meridian. The standard deviation is fixed (20 degrees), but the mean varies between two states i.e., ±12 degrees from the horizontal meridian (matches the positions used in the "instructed" condition). Participants need to infer the underlying mean in order to apply the correct rule in any given trial, where $\Delta\mu$ parameterizes the difference between the two means.

$$\mu_j = \begin{cases} \frac{\Delta\mu}{2} \; if \; j = 1 \\ -\frac{\Delta\mu}{2} \; if \; j = 2 \end{cases}$$

At the beginning of a block, each of the two stimulus-response mapping rules are equally likely to be the active rule. With each successive cue, there is an 8% probability that the currently active rule will switch to the alternative rule (i.e., a "hazard rate", H, of 0.08). Thus, any single cue can be misleading and therefore participants need to accumulate information over all cues and not just the immediately preceding cue to maximize the accuracy of their inference of the generative mean.

The grating stimuli for the lower-order decision were circular, achromatic Gabor patches with full contrast and truncated at an inner eccentricity of 2.5° and an outer eccentricity of 13.85°. The spatial frequency was fixed at 1.2 cycles/°, whereas orientation (the discriminant) varied randomly from trial to trial and was either vertical or horizontal. The gratings are either along the vertical axis or perpendicular to the vertical axis.

Procedures for Measurements and Manipulations

Each MEG sessions starts with two training blocks : one instructed block, followed by one inferred block and trial-by-trial feedback is provided. This is followed by 12 blocks of the main behavioural task alternating between inferred and instructed with each block where trial-by-trial feedback is not provided.
Following the training blocks, but prior to the main experiment blocks, and again after all main experiment blocks are complete, participants complete a "localizer" block (2 localizer blocks in total). Throughout the localizer blocks, participants perform a change detection task centered around the central fixation cross. Specifically, upon the rotation of the cross, they are prompted to indicate the direction of rotation via button press, with a leftward response signifying anticlockwise rotation and a rightward response indicating clockwise rotation. Horizontal and vertical gratings with a central aperture are also periodically presented, mirroring the stimuli utilized in the primary experiment. This localizer block served the purpose of finding orientation selective vertices in the MEG data, but this information is not used in the analysis for the thesis.

Throughout MEG sessions, head position is recorded using three fiducial coils (two placed using ear plugs in the ears, and one adhered to the skin at the nasion). At the beginning of each session, once the participant is seated comfortably, the position of the head is recorded. Following this, prior to every block (if required) the participant is given feedback on how to adjust, such that their head is as close as possible to the original position. An electrocardiogram (ECG) is recorded using two electrodes (placed just below the collarbone on the participant's right side, and just below the ribs on the participant's left side). Additionally, horizontal and vertical electrooculography (EOG) recordings are taken using a pair of electrodes above and below the participant's left eye, and a pair of electrodes to the left of the left eye and right of the right eye. Finally, a ground electrode is applied above the participant's left wrist.

The "cold pressor test" (hand immersion in ice or moderately warm water) was applied repeatedly: prior to the first main experiment block, and then repeated every two blocks, giving a total of 6 immersions per session. Within two sessions ice water was used and within two sessions moderately warm water was used (randomized). Participants were asked to keep their hands in the water for as long as they could, up to 3 minutes. Within each of the 4 main experimental sessions measurements of blood 5 pressure, heart rate, and self-reported stress were performed. Because these manipulations or measurements do not form part of our preregistered tests they are not discussed further here.

Exclusions and missing data

Participants performing below the minimum average of 55% accuracy in the inferred condition, or an average of 60% accuracy in the instructed condition are excluded from the analysis. Data from only those participants who complete all four MEG sessions will be used. Individual blocks for which the pupil data is artifactual for >50% of the recordings will be excluded from further analysis. Data from the training session and from the training blocks in the MEG sessions are not used for behavioural modelling.

Behavioural Modelling

We are interested in the conditional distribution, $P(h \mid e)$, the posterior probability, which specifies the degree of belief in hypothesis h conditioned on the observation of evidence e. The model we use describes the The Log Posterior Ratio : Belief L_n expressed in log odds, for one possible state (mapping rule 1) versus the alternative possible state (mapping rule 2) given all the information collected until encountering evidence sample X_n [14] . L_n is represented by the log-likelihood ratio LLR_n reflecting the relative evidence for each alternative associated with that sample, and with the prior ψ_n before encountering X_n. H (the 'hazard rate') is the expected probability in change of the generative distribution, ie, probability that at each step that the rule will switch from one alternative to the other.

$$Ln = \psi(Ln - 1; H) + LLRn \qquad [I]$$

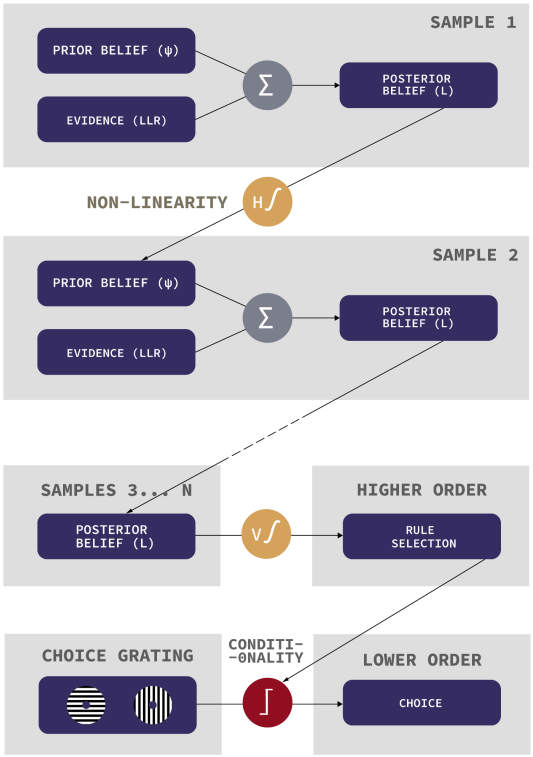

Figure 5. Schematic of normative model used for belief updating with non-linearity for rule inference and selection, followed by orientation judgement.

LLR_n can also be expressed as the difference of the log of likelihoods of states which correspond to the generative distributions of the dot presented given the fact that x_n are normally distributed (ignoring minor truncation effects [48].

$$LLRn = log(normal(X_n \mid \mu_1, \sigma_{12})) - log(normal(X_n \mid \mu_2, \sigma_{22})) \qquad [II]$$

$$LLRn = log\left(\frac{p(Xn \mid Sn = 1)}{p(Xn \mid Sn = 2)}\right)$$

$$= \frac{X_n \, \Delta\mu}{\sigma^2}$$

$$= \beta X_n$$

where β is an abbreviation. The cue values (i.e., polar angles of the dot in the right-hemifield) are scaled by β (= 3.437) to convert them to log-likelihood ratios, and we therefore refer to β as the evidence scaling.

The non-linear transformation of the posterior from each updating step into the prior for the next step renders the model adaptive in volatile environments. This possibility is indicated by hidden state changes, as a function of the subjective hazard rate H:

$$\psi_n = \psi(Ln - 1; H) = Ln - 1 + log\left(\frac{1-H}{H} + exp(-Ln - 1)\right) \qquad [III]$$

$$- log\left(\frac{1-H}{H} + exp(Ln - 1)\right)$$

At the onset of each trial of the orientation discrimination task, participants are forced to make a choice about the state they are currently in by selecting a stimulus-response mapping rule they are going to use for the second order decision. An ideal observer is deterministic as would pick the state based on the sign of the Log-posterior ratio (Ln). However, to allow for the possibility that the observer's internal rule selection decisions based on Ln are noisy, i.e. These decisions follow:

$$P(R = 1 \mid Lm) = \int_0^{\infty} N(y;\, L_m, \nu^2)\, dy \qquad [IV]$$

Such that the observer chooses the generative distribution S = 1, and uses the corresponding rule, by R = 1 and chooses S = 2 denoted by R = 2. Here m is the number of the final cue before the trial onset, and ν parameterizes the standard deviation of the decision noise. The probability of both decisions must sum to 1.

Fitting the model to observed data

The behavioural model is fit individually for each participant using three free parameters:

i) H: To allow the possibility that the observer misestimates the true hazard rate of the task, H, and instead uses $\widehat{H}$.

iii) β: The observer may not use the evidence scaling, β inherent to the task design, but a deviant value $\widehat{\beta}$, to convert dot polar angles into decision evidence *LLR*. (*Eq. II*).

iii) The standard deviation of the decision noise, ν (*Eq. IV*)

The model is fit to data by minimising the negative log-likelihood of the participants reponses (maximising the log-liklihood) using MATLAB's "fmincon" optimization algorithm (Matlab Optimization Toolbox, 2017). The upper and lower bounds for the parameter values during fitting are shown in Table 1.

We first draw 250 sets of candidate starting points of the free parameters for the optimizer. Candidate values for these free parameters are drawn from uniform distributions on the interval between the "starting point lower bound" and "starting point upper bound" shown in Table 1. The candidate starting point with the greatest log-likelihood will be used as the optimization start point. This model fitting procedure is run 10 times for each participant to minimize the risk of only discovering a local

maximum of the log-likelihood function. For each individual participant, the maximum log-likelihood found over all runs, and corresponding parameter values, are used for further analysis.

Parameter	Notation	Lower bound during fit	Upper bound during fit	Starting point lower bound	Starting point upper bound
Observer estimated hazard rate	$\widehat{H}$	0	1	0.001	0.5
Observer estimated evidence scaling	$\widehat{\beta}$	0.01	2000	0.1	300
Decision noise	ν	0	20	0.05	5

Table 1. Model parameters for fitting.

Model-based latent variable estimation

We produce estimates for various latent variables, using the parameters obtained from fitting the model as described above:

i) log-posterior ratio using *Eq. I*

ii) log-likelihood ratio after each cue using *Eq. II*

iii) observer's uncertainty prior to observing cue n using the following function of the log-prior ratio:

$$-|\psi(L_{n-1}, \widehat{H})| = -|\psi_n|$$

iv) "change point probability" or *CPP:* probability that the state changed between cue n and cue $n - 1$, given the previous belief state L_{n-1} and the new sample of evidence [46] :

$$CPP = \frac{1}{1+\Omega}$$

$$\Omega = \frac{1-\widehat{H}}{\widehat{H}} \frac{cosh(\frac{1}{2}(LLR_n + L_{n-1}))}{cosh(\frac{1}{2}(LLR_n - L_{n-1}))}$$

We will compute the accuracy of participants' choices with respect to the true states and compared this to the accuracy yielded by three idealized decision models: i) when H = 0, ψ = *L*n-1, i.e. , perfect accumulation ii) H = 0.5, selecting the rule based on only the last evidence sample, and iii) the 'ideal observer', which used the normative belief updating process in equations I-III with the knowledge of the true generative H (i.e.; H = 0.08).

We fit the model separately to each participant's data, by maximising the log-likelihood, according to the model, of the choices made by the participant to derive two computational quantities: CPP and Uncertainty as well as the expected H for each participant. *CPP* is the surprise of seeing a sample X_n given the observer's belief L_{n-1} before encountering the sample, i.e., the posterior probability that a change in the active rule has occurred, given the expected *H*. Uncertainty will be a function of the negative absolute value of the prior, before observing $X_n = -|\psi_n|$.

Pupil data

Recording and preprocessing

Recordings of the pupil diameter and gaze position (x-axis and y-axis) were sampled at 1000 Hz and calibrated and validated with a 9-point fixation routine. At the start of each MEG session, a reference recording was done with a surrogate pupil of a known size. Both pupil diameter and gaze position were preprocessed following the methods

described in van den Brink et al. (2016) and van den Brink et al. (2022). To clean the pupil data, we run an iterative algorithm through the diameter data, and sections with blinks and missing data are identified (periods where the eye-tracker lost track of the pupil or the participant blinked are already set to zero by Eyelink), and then interpolated across linearly (Knapen et al., 2016; de Gee et al., 2017). Other artifacts are identified where the derivative of the pupil diameter exceeds a threshold of 25 pixels and are also interpolated across. This process is iterated repeatedly 100 times to ensure all artefacts are identified. The sections identified to contain artifacts are interpolated across in the gaze x and y position data too. Following interpolation, deconvolution is used to identify artifactual variance in a 1000 ms period preceding the onset, and following the offset of blinks / missing data, and then regressed out.

The diameter data is converted to mm (instead of pixels) using the average of the reference recording for each block separately, to account for across-session variability in the position of the eye tracker. If the reference recording on a given session is missing, the average of the remaining reference recordings for that participant is used. Prior to running the regression, we low-pass filter the pupil data at 10Hz to filter out any high frequency noise and create a time series of z-scored pupil data referred to as Pupil(τ). When the derivative of pupil diameter is used, z-scoring is conducted after the derivative is taken.

Relationship between pupil diameter and behavior

We assessed the relationship between the dynamics of the pupil and behavior by conducting linear regression of the variables resulting from the computational model fit to individual participants' behaviour onto (the derivative of) pupil diameter. The derivative of pupil diameter is used to capture the rate of change of pupil size changes to capture the temporal dynamics of arousal-related pupillary responses [49,5051,52].

LC - NE neuromodulation can be involved in shaping decision-making via both tonic and phasic fluctuations in arousal [31] . We utilise temporally precise, task-evoked

measures of pupil derivative to relate phasic arousal with CPP associated with each cue, and slower measures of pupil diameter to relate tonic fluctuations with uncertainty formulated over longer timescales.

We time match the pupil data and behavioural data obtained from the behavioural session. For each cue presentation, we cut the pupil data into a 3001ms time window surrounding the event of interest (-1s relative to the sample, and +2s after the sample). Similarly, we include the event locked gaze positions which act as nuisance regressors.

To extract the baseline pupil values, for every sample of evidence, we time-average the first 1s of data prior to cue presentation. We subtract this baseline mean from the entire cue data to factor out any variation due to changes in any lighting conditions before we z score the time series for regression formulated as follows:

$$\frac{dPupil\,(\tau)}{dt}_{t,c} = \beta_o + \sum_{l=-1}^{0} (\beta_{1,l,t} \times CPP_{c+l} + \beta_{2,l,t} \times (-|\psi_{c+l}|) + \beta_{3,l,t} \times |LLR_{c+l}|) \quad [V]$$

$$+ \beta_4 \times gaze\,x_{t,c} + \beta_5 \times gaze\,y_{t,c} + \beta_6 \times baseline_c$$

Where,

t : time relative to cue onset

c : index cues

$\frac{dPupil\,(\tau)}{dt}_{t,c}$: pupil derivative time series at time t relative to the onset of cue c

CPP : change point probability

$-|\psi|$: uncertainty

$|LLR|$: the sensory evidence, absolute value to capture variance in pupil data independent of sign of belief state

$gazex$ and $gaze\,y$: gaze positions on screen

baseline : average pupil diameter in a 1000 ms window preceding cue onset.

Terms relating to previous cues ($l =- 1$): CPP, $- |\psi|$ and LLR were included to adjust for auto-correlation in pupil response with previous cues to isolate current cue response.

We will fit two additional variants of this regression model:

1) A variant of the regression model that includes one additional term: $|LLR_c - LLR_{c-1}|$ to exclude any relationship between pupil and CPP that is driven by low-level visual differences between consecutive cues.

2) A variant of the regression model to baseline corrected pupil diameter rather than the derivative. This regression model does not include baseline diameter or terms for preceding cues.

3) A variant of the regression model to baseline corrected pupil diameter rather than the derivative. This regression model does not include baseline diameter or terms for preceding cues, as well as terms for CPP and LLR to dissocaite slower responses from variabled that encode rapid changes in the task structure.

Statistics

Regression coefficients are compared to chance (zero) without relying on assumptions about the underlying distribution of the data, using non-parametric permutation testing (within-participant shuffling across time points simultaneously with 1000 iterations). To control for the increased risk of false positives, significance is assessed after correction for multiple comparisons across time points using the Benjamini/Hochberg (BH) false discovery rate (FDR) test.

Power analysis

A G*Power analysis was carried out to determine if the sample size (N=20) was enough to test the primary hypotheses [53]. To conduct the power analysis, we used data from van den Brink et al., (2022) (N=19) where the largest effect size for the relationship between CPP and the derivative of pupil diameter was d=0.123. With a significance level of α=0.05 and power of 0.99, a one-tailed t-test requires a minimum sample size of N=12. It's important to note that this estimate doesn't account for multiple comparison correction, but it provides an approximate indication that the sample size we obtained (N=19) is sufficient to test the main study hypotheses.

RESULTS

We fit a non-linear normative model to data from healthy human participants performing a hierarchical sensory-motor decision task involving unexpected changes in rules governing the decision. We then utilise the preprocessed pupil data around samples of interest, as a proxy for arousal, where arousal was hypothesized to facilitate belief updating in volatile environments. Specifically, we examine if two key model-inferred parameters of the belief updating process - change point probability and uncertainty - co-vary with the dynamics of the pupil.

Participant Task Performance

Investigating Signatures of Non-Linear Belief Updating in Participant Behaviour

In the instructed version of the task, previous studies show that participants were able

to switch rapidly and consistently between the rules and their performance was close to the ceiling [20]. In various real-life scenarios, however, an agent's internal belief of contextual rules can be uncertain and must be learned from incomplete and noisy information.

To track the evolution of internal belief about the sensory-motor mapping under uncertainty we model the participant's behaviour in the inferred version of the task. Here, the sensory-motor mapping rule could undergo hidden changes and had to be continuously inferred from noisy sensory evidence presented in the inter-trial intervals. The evidence was a pair of small dots presented at a high rate around the fixation mark where each position was drawn from one of two overlapping Gaussian distributions, where each distribution corresponded to the rule which was active at that time. The active rule (and thus the generative distribution from which the cues/dots were drawn from) could switch with each sample presentation with a low probability or hazard rate of 0.08. The participant then applies the selected rule to report their orientation judgment (lower-order decision).

The free parameters in the model $\widehat{H}$ and $\widehat{\beta}$, participants beliefs of statistical features of the task, which inform the the non-linear accumulation strategy employed. A mean difference of 0.002 between the median observer estimate of $\widehat{H}$ (= 0.078) and the inherent task H (= 0.08) was obtained with a p-value = 0.365 from a permutation test with 1000 permutations, thus the participants' estimate of the changing statistics of the uncertain states in the environment under the non-linear normative model is not significantly different from the model design (Figure 6(A)). Similarly, their estimate of $\widehat{\beta}$ (= 2.254) was not significantly different (p-value from permutation test with 1000 permutations = 0.633) from the scaling value for cue positions inherent in the task design as discussed in Equation. II (Task β = 3.4372) (Figure 6(A)) . This indicated that, overall, participants were relatively good at estimating the generative hazard rate and inherent evidence scaling parameters for the cues.

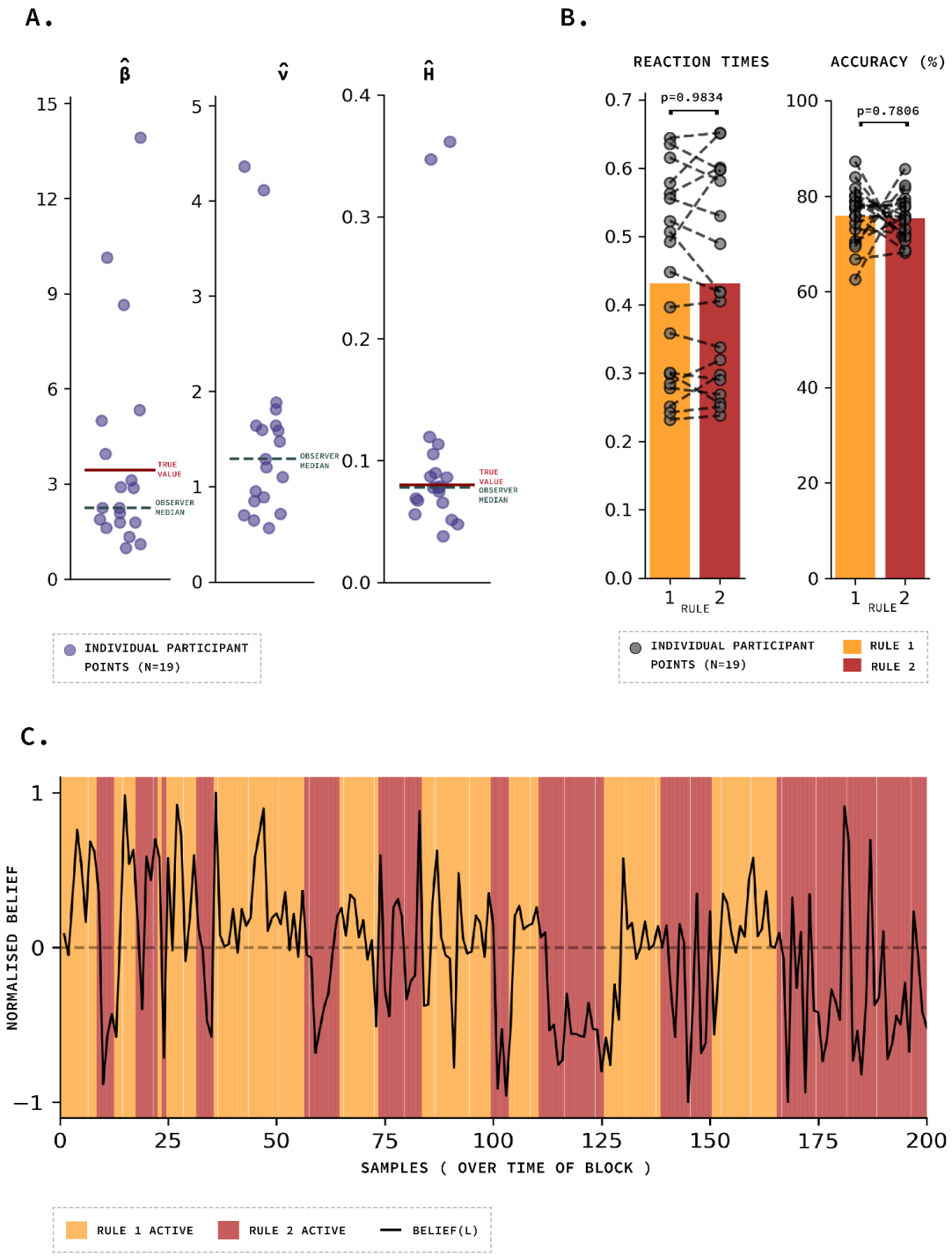

Figure 6. Features of Observer Behaviour from the Non-Linear model fitted to participant Data

(A) Parameters of best-fitting model in order: Evidence Scaling, Decision Noise, Hazard Rate. Purple dots, Individual participant values; Dark Blue Dashed Line, group average; Red Line, True H for task design.

(B) Reaction Times, and Accuracy for participants in both mapping rules. Grey Dots, Individual participant data; Shaded bars, group average. Orange: Rule 1; Red: Rule 2.

(C) Evolution of LLR of a single participant over 208 sample points in one block for change in rules with sample cue presentation. Orange: Rule 1 Active; Red: Rule 2 Active, Black Line: Tracking Normalized Belief.

Participants are able to reliably switch between the rules and their mean accuracy was 75.96% ± 5.93% (mean ± SD) for the mapping rule one and 75.35% ± 4.79% for the mapping rule two. The difference in performance between the two conditions is non significant (p-value= 0.798, two tail permutation test) (Figure 6(B)). The reaction times for rule one was 0.4320s ± 0.1430 (mean ± SD) and 0.4318s ± 0.1459 (mean ± SD) for rule two, such that it was not significantly different (p-value = 0.98, two-tail permutation test). Thus, we see that the rules were inherently symmetric in terms of difficulty, without any obvious benefit on using one rather than the other rule. This eliminates the presence of the influence of the Simon effect which often results in a difference in accuracy or reaction time between trials in which stimulus and response map on the symmetric rule vs anti-symmetric rule, in which one condition requires an inhibition of a prepotent response while the other condition does not [54]. The time-varying posterior belief (L) of the fitted model tracked the active rule in participants well (sample representation of n=1 participant data in Figure 6(C)). Yet, we see that the accuracy in tracking the change in rule is never 100% which we expect since the participants never had true knowledge of the active rule. Even an ideal observer who knows the true hazard rate of context rule switch, does not have a 100% accuracy (Figure 6B).

To summarize, the behavior of the fitted model well approximated the behavior of our participants.

Normative Models of Decision Making under Uncertainty

We compare the accuracy of n = 19 independent participant responses, data simulated with the non-linear normative model (by Glaze et al. 2015), ideal observer with knowledge of the true hazard rate dictating unexpected changes in our task (H = 0.08),

strategy perfect accumulation of evidence (H=0), strategy of deciding based on only the previous evidence sample (H=0.5).

Participants (75.65% ± 2.80%, Violet dots for individual data, and violet bar for mean in Figure 7(B)) perform better than simulated from two simple strategies: the Last Evidence sampling strategy (62.99% ± 3.71%, p<0.001, two tailed permutation test, Orange bar in Figure 7(B)) and Perfect Accumulation (51.85% ± 1.19%, p<0.001, two tailed permutation test, Blue bar in Figure 7(B)). Since the performance is not significantly different from an ideal observer (Green bar in Figure 7(B)), and the simulated model performance (Pink dots for individual data, and pink bar for mean in Figure 7(B)), we can confidently conclude that the Non-Linear belief updation model accurately captures behavioural statistics of the task.

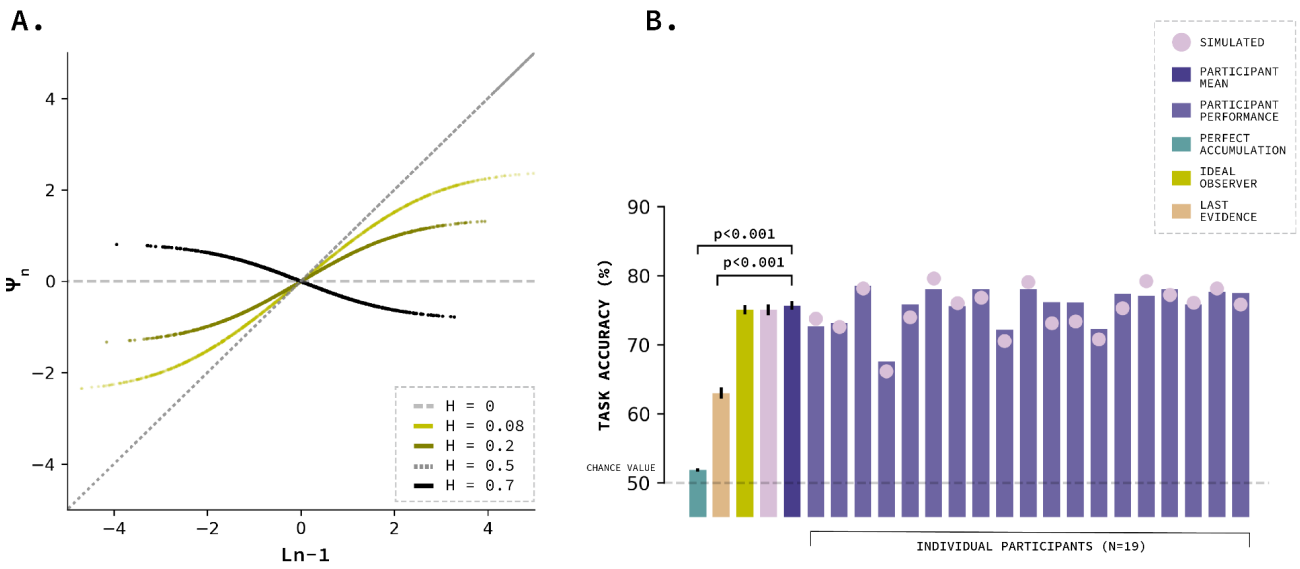

Figure 7. Different Normative Models of Decision-Making under Uncertainty

(A) Effect of Non-linearity in the normative model implicated by the hazard rate used in the choice task (H = 0.08), extremes of perfectly stable (H = 0), unpredictable (H = 0.5) environments and adjustments to intermediate values (H = 0.2, 0.7).

(B) Comparison of participant performance accuracy with best fitting normative model, ideal observer and simple heuristic strategies. Error bars, SEM.

Perfect (i.e., lossless) accumulation of all the evidence is a suboptimal strategy when the environment is undergoing hidden changes in its statistics. The normative model to maximise accuracy probes the accumulation of evidence samples non-linearly with information/incoming evidence represented in the form of log-likelihood ratios, LLRs about the two possible environmental States. Thus, updating prior belief ψ with new evidence LLR to yield a posterior belief L. The updated belief (L_n) is passed through a non-linear function which depending on the estimated hazard rate H, can saturate (slope $\approx$ 0) for strong L_n leading to perfect accumulation (equal weight for each belief) or entail information loss/leak of information ($0 \ll$ slope < 1) for weak L_n to formulate the prior for the next sample ($\psi n+1$) rendering the model adaptive (Figure 7(A)).

In our task, this non-linearity is introduced into the sensitivity of evidence accumulation to two quantities: (1) uncertainty ($-|\psi|$): negative of the modulus of prior belief representing uncertainty about environmental state prior to encountering new evidence sample, (2) the change-point probability (CPP), defined as posterior probability that an environmental state change has just occurred, given an existing belief and presentation of a new evidence sample .

Pupil Responses

The brainstem arousal system has been shown to influence levels of excitability and synchrony in cortical circuit dynamics, implying it plays a crucial role in cognitive processing. Phasic, transient signals from the LC-NE system which track event-related changes are hypothesized to adjust belief updating mechanisms when unexpected new evidence is presented, or old expected evidence is absent. To assess the involvement of brain stem arousal in belief updating in our task structure, we utilise sample evoked pupil dilations: Specifically derivative of pupil responses to increase specificity for norepinephrine (rather than ACh) transients for its involvement in the detection and

reaction to state uncertainty within a task, and improve temporal precision with respect to cue onsets [26]. This aids in decorrelating the measures of phasic arousal, i.e. the change in pupil size following event onset from tonic arousal, i.e, slower pupil response.

We extracted two measures from the time series data: 1) the average pupil diameter in a 1000ms time window before cue onset: 'baseline diameter'; and 2) the temporal derivative of diameter, to quantify the dilation or constriction of pupil in the 3001ms time points around the cue onset, 1000ms before, and 2000ms after.

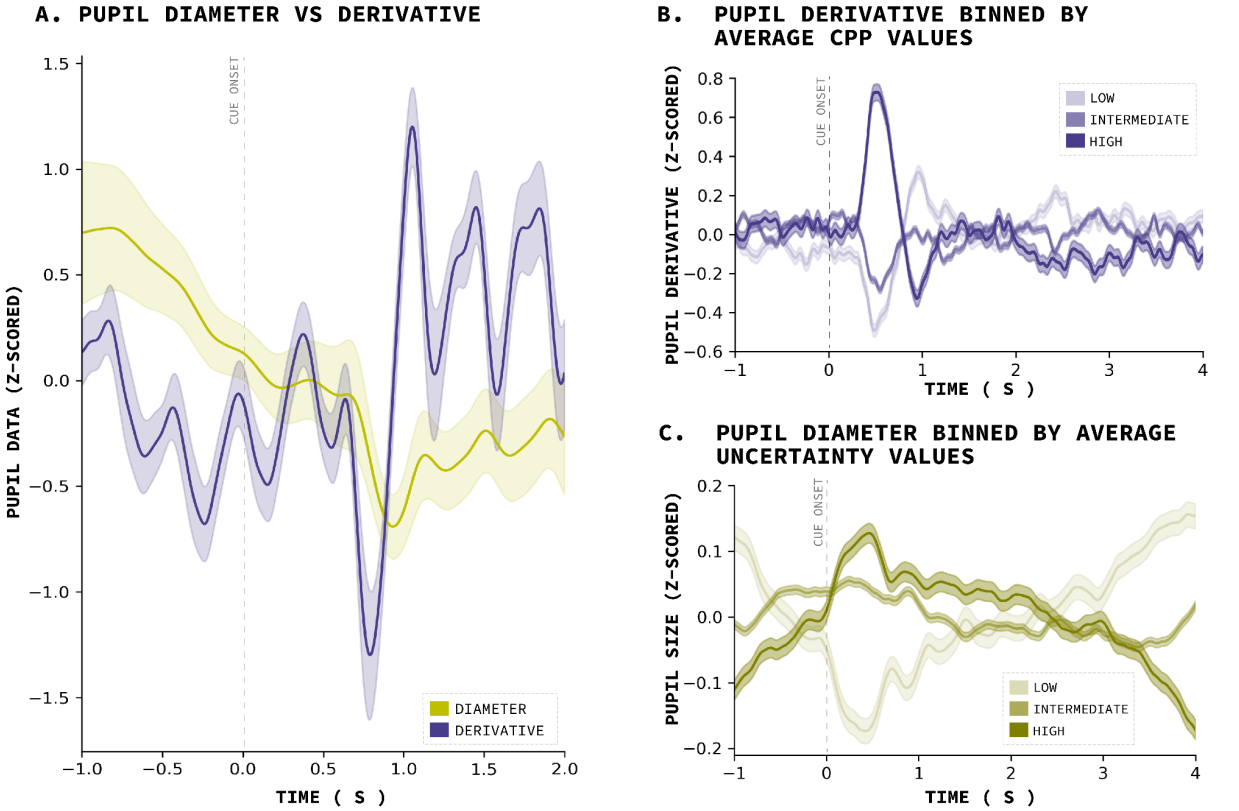

Figure 8. Features of Pupil Data

(A) Variation of pupil derivative and pupil diameter in a 3s time window around cue onset

(B) Pupil diameter discretised by CPP values averaged over all 19 participants

(C) Pupil derivative discretised by Uncertainty values averaged over all 19 participants

In Figure 8(A), we see where the dilation following a single sample of interest is only a minor fluctuation in the time series, there is a larger time specific jump in the derivative. Therefore, the derivative is ideally suited to pull out variance related to sample-evoked dilation on short time scales. Conversely, slower but sustained changes are more easily picked up by diameter, since these are less influenced by moment-to-moment fluctuations.

Pupil derivative discretized by CPP (Figure 8(B)), and diameter discretized by uncertainty (Figure 8(C)), indicate a potential relationship between the behavioural variables and pupil data worth exploring. However this method does not rule out the influence of other behavioral variables that could be confounded (uncertainty, visual saliency from previous sample , because samples follow each other closely on a temporal scale). To isolate the variance of current sample onset, we conduct linear regression of pupil response with previous samples as nuisance regressors as described in the Methods.

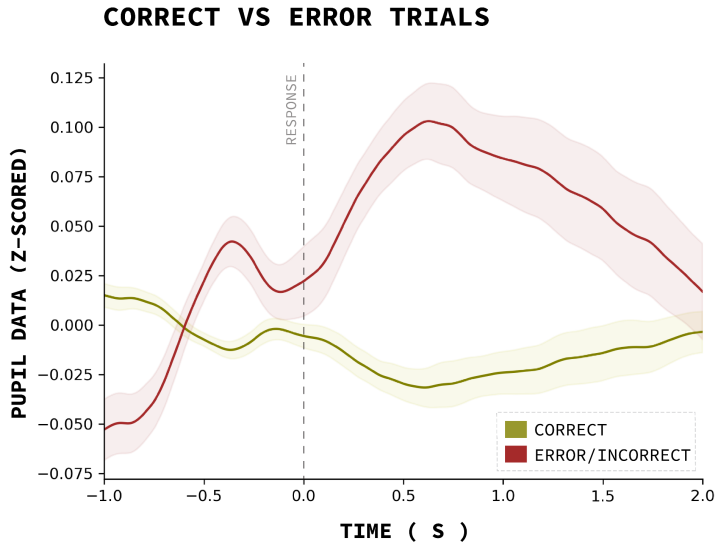

Figure 9. Evoked pupil response for decision response compares for correct vs incorrect trials averaged over all participants.

We compared the pupil responses for trials split by performance: correct trials and incorrect trials (Figure 9) . Here, since in our task no trial to trial feedback was given, incorrect trials are those where the incorrect mapping rule was applied for the lower order decision. We attribute the variance in pupil response to either higher uncertainty in estimating the active rule or the realisation of giving the wrong response, or unintended response, after submission. We subtracted the baseline pupil response from the time series so as to specifically quantify the response-evoked patterns. The pupil responses were significantly larger for erroneous trials (p-value < 0.001). Errors generally invoke alertness so that performance can be adjusted and this cognitive feature is encoded in pupil data.

Encoding of behavioural variables in pupil response

We collated the behavioural and pupil data by matching the nearest time relative to trial start time in both datasets, for further analysis to check temporal encoding of behavioural features in pupil responses during the task. We used multiple regression to examine linear relationships between Z-scored time series of pupillary responses and each of the Z-scored behavioral measures for a 3 second time window around sample of interest (cue onset) over all 19 participants.

The CPP, i.e, the posterior probability that the state changed between current cue n and previous cue n − 1, given the prior belief state L_{n-1} and the new sample of evidence, covaries with change in pupil diameter following cue onset. Thus, in our task-setting , phasic arousal was recruited by the computational quantity, CPP to modulate belief updating for 825 timepoints out of 3001 surroundung the cue in a window of -1s to 2s with a p value < 0.05 out of which 674 survived FDR correction (Figure 10(A)). This

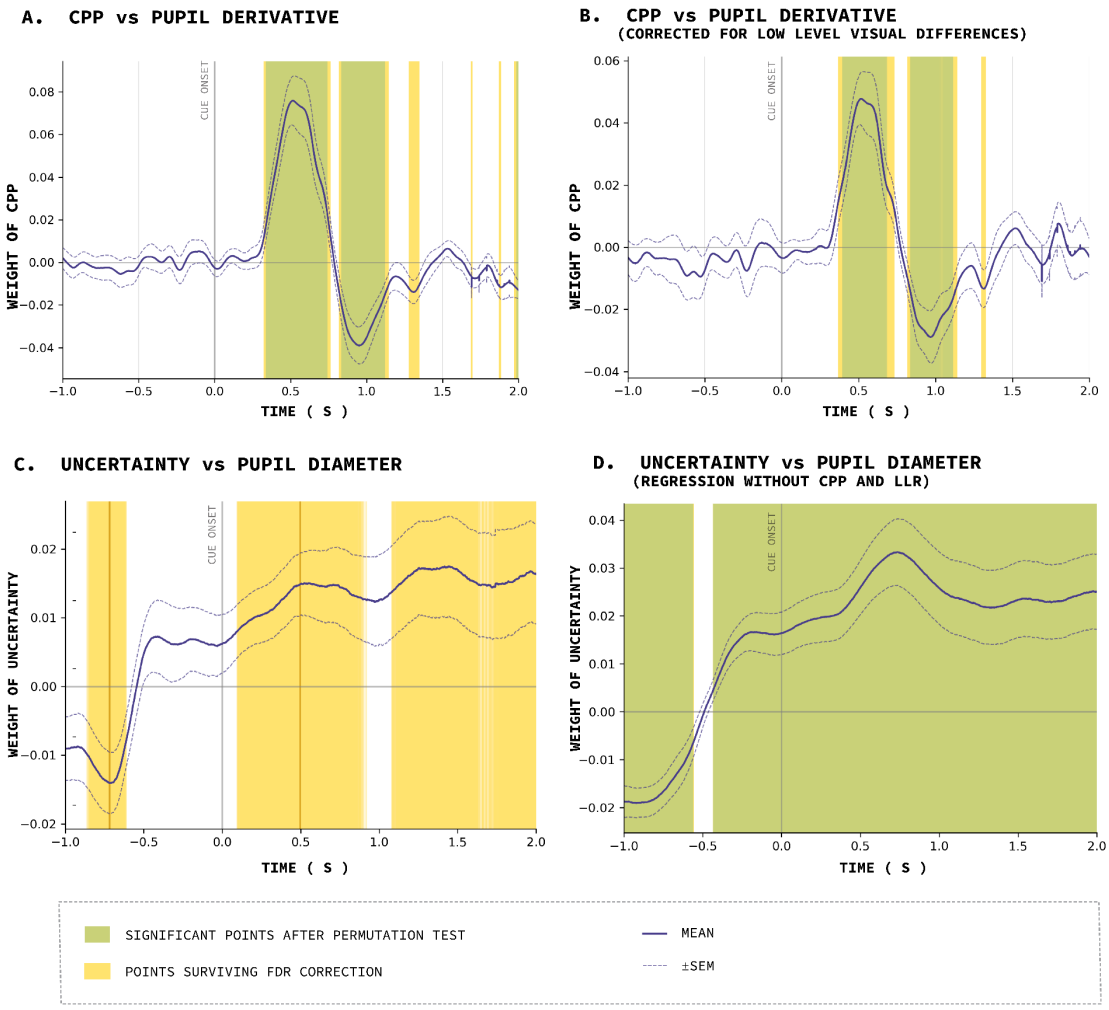

Figure 10. Encoding of Model-Derived Variables to Pupil Data averaged over all 19 participants in a 3 second time window around cue onset (X = 0 is the time of cue presentation)

(A) Encoding of CPP in pupil derivative from linear regression

(B) Encoding of CPP in pupil derivative from linear regression after correction for low-level visual saliency from previous cue

(C) Encoding of Uncertainty is pupil diameter from linear regression

(D) Encoding of Uncertainty is pupil diameter from linear regression model that excludes moment-to-moment fluctuations of CPP and LLR encoding

encoding survived even when accounting for low-level visual differences between consecutive cues in our regression model for 681 timepoints (p values < 0.05) out of which 491 timepoints survived FDR correction (Figure 10(B)).

Derivative of pupil diameter has been hypothesized to be a potential marker of attentional performance, this fits into our results reflecting a possibility of increased attention due to arousal for the detection and adaptation to these changes in context rules with cues to update beliefs [55].

Uncertainty in belief, is a behavioural feature that unfolds over longer timescales shaped by tonic NE which is associated with exploration, and detection of changes in context [32,56]. To investigate the involvement of tonic fluctuations in arousal, we checked for the encoding of uncertainty in slower fluctuations of pupil diameter, instead of moment-to-moment fluctuations in pupil derivative.

Uncertainty in participant's belief before the onset of new cue evidence encoded in tonic pupil response, starts rising right before cue onset, and increases following the onset on each cue for 1770 timepoints but only one 1 timepoint survives the FDR correction for false positives (Figure 10(C)).

Since we hypothesise that both CPP and LLR for each cue, are associated with fluctuations on a shorter timescale captured by the pupil derivative, we exclude them for the investigation of encoding of uncertainty as they add noise to the long-term uncertainty encoding in slower baseline pupil response. Uncertainty covasries with baseline pupil response for 2864 timepoints around each cue onset and all timepoints survive the FDR correction for false positives (Figure 10(D)).

DISCUSSION

There have been significant developments in the evolution of abstract, normative solutions for decision-making problems, however, there has been a gap in

understanding the processes that govern belief updating and adaptation to changing environmental properties a feature commonly found in natural environments [57]. Accurate decisions about the sensory environment thus depend not only on momentary sensory input but also on behavioural context [58].

In our task which presents unexpected changes in contexts for decision rules, behavior observed in healthy participants is well approximated with a non linear normative model of decision making in uncertain environments. This behavior is symmetric across the two rules, and performance is better than other heuristic-based decision strategies: Linear Accumulation and Last Sample strategies. We find variation of pupil responses with different cognitive events, broadly correct and incorrect responses for decision-making. Critically, we found that key computational parameters of the decision process, specifically CPP, are encoded in the dynamics of pupil diameter, specifically phasic pupil responses. These findings implicate the involvement of arousal systems of the brainstem in belief updating in uncertain environments.

Previous studies have identified the cortical encoding of normative decision variables for changing environments in the build-up activity of cortical regions involved in action planning, recurrent interactions in cortical microcircuits leading to the emergence of and adaptive computations for decision making [59]. It has been proposed that in the association cortex, neural populations function as temporary switches that respond to the combination of a stimulus and rule, directing sensory signals from sensory regions to action-related regions in a flexible manner that depends on the active rule in play [60] [21]. Our findings suggest that the brainstem, via neuromodulation may 'flip the switch' for context reversal in directing sensory signals to action-regions.

van den Brink et el. 2022 investigated the relationship between patterns of intrinsic correlations in population activity and participants' belief states during flexible sensory-motor decision making utilizing the assessment of feature-specific population codes expressed in fine-grained activity patterns within each brain region [20]. Their results show that intrinsic correlations in stimulus- and action-selective population activity reliably track belief states, but exhibit selective breakdowns at the time of behavioral errors. This finding suggests brain-wide activity encodes cognitive

computation in the structure of correlated variability which plays a significant role in flexible decision making. Context reversal, ie, change in sensory-motor mapping rule governing the decision, flips signs in the correlation population activity in the sensory and response populations in the brain [20]. Our results point towards a plausible ***cause*** of this change in signs of correlated variability: Brainstem modulated arousal.

This tight link between existing circuit models and normative evidence accumulation highlights the importance of incorporating neuromodulatory input in models for evidence accumulation by multiple processing stages.

The activity of the LC, regulating central arousal state has widespread modulatory projections to the cortex, leading to fluctuations in pupil size with the shortest latency, among the various structures studied. A growing body of literature supports the involvement of phasic norepinephrine in the flexible updating of beliefs in response to unexpected changes within a task. Sara and Bouret (2012) propose that this phasic release functions as a "reset signal" for reconfiguration of large-scale neural circuits involved in decision-making allowing for flexible information flow and the updating of beliefs in response to changes in the environment [61] . Aston-Jones and Cohen (2005) propose that norepinephrine acts as "attentional gain control" mechanisms to enhance the processing of task-relevant information by enhancing the signal-to-noise ratio of neurons, allowing for the selective amplification of task-relevant information and the suppression of irrelevant information[31]. Yu and Dayan (2006) support that this may play a crucial role in regulating the balance between exploration and exploitation during decision-making by regulating the influence of prior beliefs to cause a shift toward more ‘bottom–up’ signalling relative to ‘top–down’ signalling and thus a greater impact of new evidence on the evolving belief [25,32].

The model by Glaze et al. 2015 recruits a non-linear variable for regulation of this balance in the face of uncertainty. To further dwell into the interplay of this neuromodulatory arousal with flexible decision making we employ a task structure where the active rule governing the stimulus to response is volatile. We check for the encoding of statistical information employed in this decision-making process in arousal-driven pupil responses in two quantities: (1) the uncertainty before encountering

a new sample in a changing environment (denoted as $-|\psi|$) and (2) the change-point probability (CPP) that a state change just occurred in the task structure which has been observed to modulate normative belief updating and phasic arousal responses more strongly than alternate metrics of surprise used in literature (Shannon surprise model, Dayan & Yu model of surprise)[46,62] .

We find that phasic arousal, as measured by changes in pupil diameter, is recruited to modulate belief updating in response to changes in the task environment specifically, the posterior probability of a state change, given prior beliefs and new evidence, covaried with a clear measurable response in the derivative of pupil diameter following cue onset. We also observe greater pupil response for error encoding to adjust performance. This supports the view that the arousal response may play a role in detecting and adapting to changes in cues and updating beliefs accordingly. Moreover, since this encoding survived even when accounting for low-level visual differences between consecutive cues suggests that it is not simply a result of differences in visual salience detection between cues, but rather a potential diagnostic of modulation of belief updating. This finding is particularly relevant in the context of the idea that changes in pupil diameter may be a useful marker of attentional performance, as it reflects the increased attention required to detect and adapt to changes in cues [55].

The model we employed defines uncertainty in a formulation which dissociates it from error detection which has been linked to pupil dilation before. We also found that the uncertainty in a participant's beliefs following the onset of each sample of evidence is encoded in the baseline pupil response. However, this did not survive the FDR correction which could be either because the effect size was too small to detect with the current sample size, or that tonic pupil response does not actively track uncertainty dictated by the applied behavioural model. This points towards the caveat that applying one statistical model to all participant's strategies, may undermine the presence of individual probabilistic inference models for each brain, and its dynamic evolution throughout the task through learning. A successful fit indicates that the model is sufficient to explain the observed data, but it does not necessarily imply that it is the only possible explanation. We also hypothesize that moment-to-moment fluctuations in CPP and LLR add noise to the encoding of uncertainty in slower baseline pupil

response. After exclusion of these two terms in the regression model, we find a clearer signature of uncertainty response in baseline pupil response around cues. Thus we find two behavioural markers, one of change point detection and one of uncertainty in distinguishable physiological pupil response formulated over different timescales. This associated with the two different functional modes of the LC-NE mediated arousal, Phasic and Tonic modes respectively.

The behavioural effect of pupil-linked arousal might promote choice alternation at the level of the motor system, during response preparation, or might modulate the decision stage for dynamic updating of beliefs about the upcoming evidence,or it may affect the overall correlational variability between these two systems in dynamic belief updating. This must be investigated by selectively controlling arousal activity using pharmacological manipulations. During the data collection, we employed the use of a pharmacological drug in a double-blind setup: Prazosin, which blocks the alpha1-adrenoreceptors that are primarily located in the frontal cortex. If by using the drug we can 'rescue' behaviour (that is, reverse the effects of the stress manipulation as mentioned in the Methods, which increases noradrenaline (=norepinephrine) levels) then that is strong evidence for the selective role of NE in causing behavioural effects. Its involvement in the context of motor activity, visual response, and correlational variability by utilising the MEG data collected is next in the pipeline for the larger main study, which these discussed results are a part of. Pupil dilation is also influenced by multiple factors such as emotions, tiredness, and cognitive load, therefore, it not entirely reliable in understanding task-related behaviour in isolated analysis, and must be coupled with cortical population activity analysis. Thus, relating these indirect pupil measures to other physiological measures, such as EEG or MEG, may provide a more robust picture of ongoing computations.

Overall, we find a strong encoding of computationally derived behviouiral variables in physiological pupil responses as a marker for arousal recruitment. These findings contribute to building our understanding of how arousal shapes the detection of surprising events and updating of beliefs with changing environmental statistics, and suggest that changes in pupil diameter may be a useful marker of these processes.

Brainstem modulated arousal is a plausible *cause* of the change in signs of correlated variability.

REFERENCES

1. Shadlen, M.N., and Kiani, R. (2013). Decision Making as a Window on Cognition. Neuron *80*, 791–806. 10.1016/j.neuron.2013.10.047.

2. Sato, Y., and Kording, K.P. (2014). How much to trust the senses: likelihood learning. J. Vis. *14*, 13.

3. Green, D.M., and Swets, J.A. (1974). Signal Detection Theory and Psychophysics.

4. Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. Annu. Rev. Neurosci. *30*, 535–574.

5. Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J.D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychol. Rev. *113*, 700–765.

6. Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput. *20*, 873–922.

7. Clifford, C.W.G., and Ibbotson, M.R. (2002). Fundamental mechanisms of visual motion detection: models, cells and functions. Prog. Neurobiol. *68*, 409–437.

8. Macmillan, N.A., and Douglas Creelman, C. (2021). Detection Theory: A User's Guide, 3rd Edition (Routledge).

9. Lasley, D.J., and Cohn, T. (1981). Detection of a luminance increment: effect of temporal uncertainty. J. Opt. Soc. Am. *71*, 845–850.

10. Verghese, P., Watamaniuk, S.N., McKee, S.P., and Grzywacz, N.M. (1999). Local motion detectors cannot account for the detectability of an extended trajectory in noise. Vision Res. *39*, 19–30.

11. Usher, M., and McClelland, J.L. (2001). The time course of perceptual choice: the leaky,

competing accumulator model. Psychol. Rev. *108*, 550–592.

12. Nienborg, H., and Cumming, B.G. (2009). Decision-related activity in sensory neurons reflects more than a neuron's causal effect. Nature *459*, 89–92.

13. Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. Neuron *36*, 955–968.

14. Glaze, C.M., Kable, J.W., and Gold, J.I. (2015). Normative evidence accumulation in unpredictable environments. Elife *4*. 10.7554/eLife.08825.

15. Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature *503*, 78–84.

16. Gold, J.I., and Shadlen, M.N. (2003). The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. J. Neurosci. *23*, 632–651.

17. Tsetsos, K., Pfeffer, T., Jentgens, P., and Donner, T.H. (2015). Action Planning and the Timescale of Evidence Accumulation. PLoS One *10*, e0129473.

18. Sarafyazd, M., and Jazayeri, M. (2019). Hierarchical reasoning by neural circuits in the frontal cortex. Science *364*. 10.1126/science.aav8911.

19. Purcell, B.A., and Kiani, R. (2016). Hierarchical decision processes that operate over distinct timescales underlie choice and changes in strategy. Proc. Natl. Acad. Sci. U. S. A. *113*, E4531–E4540.

20. van den Brink, R.L., Hagena, K., Wilming, N., Murphy, P.R., Büchel, C., and Donner, T.H. (2022). Flexible Sensory-Motor Mapping Rules Manifest in Correlated Variability of Stimulus and Action Codes Across the Brain. bioRxiv, 2022.03.10.483758. 10.1101/2022.03.10.483758.

21. Cocuzza, C.V., Ito, T., Schultz, D., Bassett, D.S., and Cole, M.W. (2020). Flexible Coordinator and Switcher Hubs for Adaptive Task Control. J. Neurosci. *40*, 6949–6968.

22. Durstewitz, D., Vittoz, N.M., Floresco, S.B., and Seamans, J.K. (2010). Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. Neuron *66*, 438–448.

23. de Gee, J.W., Colizoli, O., Kloosterman, N.A., Knapen, T., Nieuwenhuis, S., and Donner, T.H. (2017). Dynamic modulation of decision biases by brainstem arousal systems. Elife *6*. 10.7554/eLife.23232.

24. Nassar, M.R., Rumsey, K.M., Wilson, R.C., Parikh, K., Heasly, B., and Gold, J.I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. Nat. Neurosci. *15*, 1040–1046.

25. Dayan, P., and Yu, A.J. (2006). Phasic norepinephrine: a neural interrupt signal for unexpected events. Network *17*, 335–350.

26. Reimer, J., McGinley, M.J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D.A., and Tolias, A.S. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nat. Commun. *7*, 13289.

27. Breton-Provencher, V., and Sur, M. (2019). Active control of arousal by a locus coeruleus GABAergic circuit. Nat. Neurosci. *22*, 218–228.

28. Rajkowski, J., Majczynski, H., Clayton, E., and Aston-Jones, G. (2004). Activation of monkey locus coeruleus neurons varies with difficulty and performance in a target detection task. J. Neurophysiol. *92*, 361–371.

29. Clayton, E.C., Rajkowski, J., Cohen, J.D., and Aston-Jones, G. (2004). Phasic activation of monkey locus ceruleus neurons by simple decisions in a forced-choice task. J. Neurosci. *24*, 9914–9920.

30. Aston-Jones, G., Rajkowski, J., and Cohen, J. (1999). Role of locus coeruleus in attention and behavioral flexibility. Biol. Psychiatry *46*, 1309–1320.

31. Aston-Jones, G., and Cohen, J.D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu. Rev. Neurosci. *28*, 403–450.

32. Yu, A.J., and Dayan, P. (2005). Uncertainty, neuromodulation, and attention. Neuron *46*, 681–692.

33. Joshi, S., Li, Y., Kalwani, R.M., and Gold, J.I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. Neuron *89*, 221–234.

34. de Gee, J.W., Knapen, T., and Donner, T.H. (2014). Decision-related pupil dilation reflects upcoming choice and individual bias. Proc. Natl. Acad. Sci. U. S. A. *111*, E618–E625.

35. Joshi, S., and Gold, J.I. (2020). Pupil Size as a Window on Neural Substrates of Cognition. Trends Cogn. Sci. *24*, 466–480.

36. Papageorgiou, E., Ticini, L.F., Hardiess, G., Schaeffel, F., Wiethoelter, H., Mallot, H.A., Bahlo, S., Wilhelm, B., Vonthein, R., Schiefer, U., et al. (2008). The pupillary light reflex

pathway. Neurology *70*, 956–963.

37. Wang, C.-A., Boehnke, S.E., White, B.J., and Munoz, D.P. (2012). Microstimulation of the Monkey Superior Colliculus Induces Pupil Dilation Without Evoking Saccades. The Journal of Neuroscience *32*, 3629–3636. 10.1523/jneurosci.5512-11.2012.

38. Wang, C.-A., Boehnke, S.E., Itti, L., and Munoz, D.P. (2014). Transient Pupil Response Is Modulated by Contrast-Based Saliency. The Journal of Neuroscience *34*, 408–417. 10.1523/jneurosci.3550-13.2014.

39. Nieuwenhuis, S., Aston-Jones, G., and Cohen, J.D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. Psychol. Bull. *131*, 510–532.

40. Liu, Y., Rodenkirch, C., Moskowitz, N., Schriver, B., and Wang, Q. (2017). Dynamic Lateralization of Pupil Dilation Evoked by Locus Coeruleus Activation Results from Sympathetic, Not Parasympathetic, Contributions. Cell Reports *20*, 3099–3112. 10.1016/j.celrep.2017.08.094.

41. Murphy, P.R., O'Connell, R.G., O'Sullivan, M., Robertson, I.H., and Balsters, J.H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. Hum. Brain Mapp. *35*, 4140–4154.

42. Gilzenrat, M.S., Nieuwenhuis, S., Jepma, M., and Cohen, J.D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. Cogn. Affect. Behav. Neurosci. *10*, 252–269.

43. Pfeffer, T., Keitel, C., Kluger, D.S., Keitel, A., Russmann, A., Thut, G., Donner, T.H., and Gross, J. (2022). Coupling of pupil- and neuronal population dynamics reveals diverse influences of arousal on cortical processing. Elife *11*. 10.7554/eLife.71890.

44. Vincent, P., Parr, T., Benrimoh, D., and Friston, K.J. (2019). With an eye on uncertainty: Modelling pupillary responses to environmental volatility. PLoS Comput. Biol. *15*, e1007126.

45. Krishnamurthy, K., Nassar, M.R., Sarode, S., and Gold, J.I. (2017). Arousal-related adjustments of perceptual biases optimize perception in dynamic environments. Nature Human Behaviour *1*. 10.1038/s41562-017-0107.

46. Murphy, P.R., Wilming, N., Hernandez-Bocanegra, D.C., Prat-Ortega, G., and Donner, T.H. (2021). Adaptive circuit dynamics across human cortex during evidence accumulation in changing environments. Nat. Neurosci. *24*, 987–997.

47. Wilming, N., Murphy, P.R., Meyniel, F., and Donner, T.H. (2020). Large-scale dynamics of perceptual decision information across human cortex. Nat. Commun. *11*, 5109.

48. Gold, J.I., and Shadlen, M.N. (2001). Neural computations that underlie decisions about sensory stimuli. Trends Cogn. Sci. *5*, 10–16.

49. Gee, J.W. de, de Gee, J.W., Correa, C.M.C., Weaver, M., Donner, T.H., and van Gaal, S. Pupil dilation and the slow wave ERP reflect surprise about choice outcome resulting from intrinsic variability in decision confidence. 10.1101/2020.06.25.164962.

50. Preuschoff, K. (2011). Pupil dilation signals surprise: evidence for noradrenaline's role in decision making. Frontiers in Neuroscience *5*. 10.3389/fnins.2011.00115.

51. Colizoli, O., de Gee, J.W., Urai, A.E., and Donner, T.H. (2018). Task-evoked pupil responses reflect internal belief states. Sci. Rep. *8*, 13702.

52. Urai, A.E., Braun, A., and Donner, T.H. (2017). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. Nat. Commun. *8*, 14637.

53. Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods *39*, 175–191.

54. Simon, J.R., Richard Simon, J., and Wolf, J.D. (1963). CHOICE REACTION TIME AS A FUNCTION OF ANGULAR STIMULUS-RESPONSE CORRESPONDENCE AND AGE. Ergonomics *6*, 99–105. 10.1080/00140136308930679.

55. van den Brink, R.L., Murphy, P.R., and Nieuwenhuis, S. (2016). Pupil Diameter Tracks Lapses of Attention. PLoS One *11*, e0165274.

56. Usher, M., Cohen, J.D., Servan-Schreiber, D., Rajkowski, J., and Aston-Jones, G. (1999). The Role of Locus Coeruleus in the Regulation of Cognitive Performance. Science *283*, 549–554. 10.1126/science.283.5401.549.

57. Behrens, T.E.J., Woolrich, M.W., Walton, M.E., and Rushworth, M.F.S. (2007). Learning the value of information in an uncertain world. Nature Neuroscience *10*, 1214–1221. 10.1038/nn1954.

58. Wang, X.-J., and Yang, G.R. (2018). A disinhibitory circuit motif and flexible information routing in the brain. Current Opinion in Neurobiology *49*, 75–83. 10.1016/j.conb.2018.01.002.

59. Ito, T., Yang, G.R., Laurent, P., Schultz, D.H., and Cole, M.W. (2022). Constructing neural network models from brain data reveals representational transformations linked to adaptive behavior. Nat. Commun. *13*, 1–16.

60. Miller, E.K., and Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. Annu. Rev. Neurosci. *24*, 167–202.

61. Network reset: a simplified overarching theory of locus coeruleus noradrenaline function (2005). Trends Neurosci. *28*, 574–582.

62. Filipowicz, A.L., Glaze, C.M., Kable, J.W., and Gold, J.I. (2020). Pupil diameter encodes the idiosyncratic, cognitive complexity of belief updating. Elife *9*. 10.7554/eLife.57872.

63. van den Brink, R. L., Calder-Travis, J. M., Thawani, S., Schwabe, L., & Donner, T. H. (2023). Preregistration: The relationship between computational parameters of decision making under uncertainty and the dynamics of pupil diameter [Retrieved from osf.io/hnwfr].