Focused ultrasound neuromodulation of mediodorsal thalamus disrupts decision flexibility during reward learning

Graeme Mackenzie[1,2], William Gilmour[1,2], Shen-Shen Yang[1], Szabolcs Suveges[1], Jennifer MacFarlane[3], Avinash Kanodia[4], James Manfield[5], Sadaquate Khan[1,6], Jibril Osman-Farah[8], Antonella Marcerolo[7,8,9], Douglas Steele[1,2],Tom Gilbertson[1,2*]

[1]Division of Neuroscience, Ninewells Hospital & Medical School, University of Dundee, James Arnott Drive, Dundee, DD2 1SG, UK

[2]Department of Neurology, Ninewells Hospital & Medical School, James Arnott Drive, Dundee, DD2 1SG, UK.

[3]Medical Physics, Ninewells Hospital & Medical School, James Arnott Drive, Dundee, DD2 1SG, UK.

[4]Department of Radiology, Ninewells Hospital & Medical School, James Arnott Drive, Dundee, DD2 1SG, UK.

[5]Institute of Neurological Sciences, Queen Elizabeth University Hospital, 1345 Govan Road, Glasgow, G51 4TF, UK.

[6]Department of Neurosurgery, Royal Infirmary of Edinburgh, 51 Little France Cresent, Old Dalkeith Rd, Edinburgh, EH16 4SA.

[7]University of Liverpool, Institute of Systems, Molecular & Integrative Biology, Faculty of Health and Life Sciences, Brownlow Hill, L69 7ZX, UK.

[8] The Walton Centre NHS Foundation Trust for Neurology and Neurosurgery, Liverpool, UK.

[9]Faculty of Health & Education, Manchester Metropolitan University, UK.

Display items: 6

Running Title: Human thalamus role in reinforcement learning

Keywords: Focused ultrasound, thalamotomy, decision making, mediodorsal thalamus, reward sensitivity.

[*]Corresponding Author:-

Tom Gilbertson
Division of Neuroscience & Department of Neurology
Ninewells Hospital & Medical School, Dundee, UK, DD1 9SY
t.gilbertson@dundee.ac.uk

Abstract

When learning to find the most beneficial course of action, the prefrontal cortex guides decisions by comparing estimates of the relative value of the options available. Basic neuroscience studies in animals support the view that the thalamus can regulate the activity within and across the prefrontal cortex. However, it is unknown if it can modify value-based decision making in humans. Understanding this is clinically important because disorders of value-based decision making underpin common motivational syndromes of apathy and impulsivity. Neuromodulation of the thalamus is routinely performed for motor symptoms and therefore could be a viable future target for these neuropsychiatric complications of neurodegenerative disease. Here, a group of patients (n=37) undergoing MR guided focused ultrasound for essential tremor, were tested using the restless bandit, a reward reinforcement learning task, immediately before and after thalamotomy. Performance was compared to a group of age- and gender-matched healthy controls (n = 32).

Thalamotomy significantly impaired the proportion of switch choices during the task ($P<0.001$) without affecting overall performance. Importantly, this reduction in choice flexibility could not be attributed to diminished attention, as thalamotomy did not affect response time, accuracy, or lateralisation. A reinforcement learning model fitted to the patients' choices replicated the effect of thalamotomy when the model increased exploitation of the bandits' learnt value estimate. This shift in the explore-exploit trade-off, manifesting as reduced choice flexibility, co-varied with the extent of post-operative oedema extension into mediodorsal nucleus ($R= 0.64$, $p<0.001$), but no other, including the Ventral intermediate nucleus, targeted clinically during MRgFUS thalamotomy for tremor control. Furthermore, using a normative functional connectome, resting-state fMRI connectivity between the volume of oedema encroaching the mediodorsal nucleus and the prefrontal cortex, predicted individual patients change in bandit performance. Lastly, we used probabilistic tractography to confirm a thalamo-cortical circuit between the lateral portion of mediodorsal nucleus and the frontal polar cortex, which regulates decision flexibility and accounts for the thalamotomy induced behavioural effect observed in the restless bandit task.

These findings confirm a causal role of the thalamus and specifically the mediodorsal nucleus, in regulating the extent to which value estimates are used to guide decisions and learning from reward. Neuromodulation of this target would be worthy of further investigation in patients with diminished motivation characteristic of apathy in conditions including Parkinson's disease.

Introduction

Survival relies upon adaptation to environmental changes and making decisions that lead to maximum utility. In humans, the prefrontal cortex (PFC), consists of multiple sub-regions which contribute discrete functions to both behavioural control and motivation [1–4]. These furnish the human brain with cognitive flexibility by allocating values to actions which are based upon learnt experience[3,5–7]. A feature of human decision making is the ability to dynamically alter the extent to which value estimates are incorporated and expressed into behaviour[8–10]. Neural activity in PFC encodes behavioural strategies under conditions were environmental stability requires maintenance of the *status quo* to exploit confidence in the reliability of a learnt choices value[8,9,11–13]. Equally, when the correct course of action is unclear, activity in PFC predicts adaptive, exploratory strategies[8,9,13]. Exploration aim's to minimise uncertainty, and regain confidence in the reliability of the value-estimate, by updating, through knowledge gained from experiencing an alternative course of action[11,14–17]. Evidence from the past two decades has converged on the existence of two partly distinct PFC circuits which contribute to these aspects of executive function[18]. Cognitive control, which imparts the ability to monitor and switch behaviour, inhibit actions, and detect conflict, has been associated with dorsolateral PFC (DLPFC) and anterior cingulate cortex (ACC)[19,20]. In contrast, a valuation network, centred on the ventromedial and lateral orbitofrontal regions (vmPFC/ OFC), guides decisions through learning and encoding neural estimates of decision value[21,22]. How these circuits interact, and their respective neural functions, are integrated, to promote switching between these types of behaviour remains unclear. Converging evidence suggests subcortical structures, whose respective grey matter nuclei send and receive neural projections to these regions of the PFC may contribute to this role[23–25].

Located deep within the brain, the thalamus contributes to multiple sensory, motor and cognitive processes. In contrast to ventroposterolateral (VPL) nuclei (which receives sensory afferents from the periphery), and the motor, ventral intermediate nuclei (VIM), (which receives afferents from the periphery and cerebellum), the 'higher-order' cognitive nuclei are defined by their reciprocal connections with the PFC and a separate group of subcortical nuclei known as the basal ganglia[26]. Higher-order thalamic nuclei include the mediodorsal (MD), Ventrolateral anterior (VLa), centromedian (CM), centrolateral (CL) anterior (ATN) and intralaminar parafasicular (Pf) complex[27]. Of these, the functional contribution of the MD nucleus to cognition has been most extensively characterised[28–31]. Its importance in cognition is reflected in dense connections with frontal cortical regions including DLPFC, ACC and premotor cortex[32]. Accordingly, lesions of the MD nucleus result in behaviour that show striking similarity to those of PFC lesions such as diminished behavioural flexibility, perseveration, and deficits in working memory[33,34].

The MD nucleus can be divided into sub-regions based upon their connectivity[35,36]. The lateral portion, or parvocellular (MDpc), exhibits strong reciprocal connectivity with the DLPFC (Brodmann areas 9/46) and Frontal Pole (Brodmann area 10). Consistent with the role these cortical region plays in the formation of abstract rules, patients with lesions of the lateral MD nucleus are impaired in tasks

that rely upon rapid learning and re-learning stimulus-action relationships such as the Wisconsin Card Sorting Test (WCST)[37]. In mice, enhancing the activity of the lateral MD nucleus promotes the formation of the cortical neural assembles that form in PFC during stimulus-response learning[38,39]. This has led to the idea that MD nucleus may act to amplify connectivity within the PFC and in doing so enhance the behaviourally relevant rule formation necessary for adaptive-decision making[31,40]. Lesions of the medial magnellocelluar (MDmc) subdivision, lead to an in-ability to persist in choosing the most rewarding beneficial decision after it has been identified[41]. The MDmc shares extensive reciprocal connectivity with the OFC[35,36]. The MDmc's role within this thalamo-cortical network, may be to regulate the threshold of whether to persist, or change decision strategy, based upon information about a decision's reward value it receives encoded by the PFC[42,43].

MR guided Focused ultrasound (MRgFUS) ablation of the VIM nucleus of the thalamus is routinely performed for patients with medication refractory tremor[44]. The VIM lies within a central position of the thalamus in close proximity to the higher order thalamic nuclei, including the MD, CM and VLa[45]. A characteristic of MRgFUS VIM-thalamotomy is the extensive vasogenic oedema which evolves within hours of the procedure and to varying degrees envelopes the surrounding thalamic tissue[46]. MRgFUS Thalamotomy therefore affords the unique opportunity to prospectively study the lesional effect on the higher-order thalamic nuclei function. By testing patients before and after thalamotomy using the 4-armed "restless bandit" task we aimed to elucidate the thalamic functional role in reinforcement learning, as successful performance in this task relies upon both recruitment of cognitive control and valuation functions of the PFC[8,9]. We hypothesised that if the cognitive thalamic nuclei such as the MD are critical hubs regulating the PFC's cognitive control or valuation network functions, then oedema extending into this network would produce a behavioural correlate which indexes a loss of cognitive control, value-based learning, or a combination of both.

Using lesion-behaviour mapping analysis, we show that post-operative lesional oedema extending into the lateral portion of the MD nucleus explains post-operative impairment in choice switching. A reinforcement learning model fitted to this behaviour, suggests the impairment stems from heightened confidence in learned value-estimates which is expressed behaviourally as diminished decision flexibility. These results support a role for the MD nucleus in augmenting cognitive flexibility for adaptive decision making, and indicate its potential as a neuromodulation target for disorders of motivation and cognitive control.

Methods

Participants

Thirty-seven patients with a clinical diagnosis of Essential Tremor (7 female), age 73.9±6.4), (Table 1) were recruited from MRgFUS Thalamotomy treatment centres in Dundee (n=25) and Liverpool (n=12). One patient was excluded from further analysis as they were unable to respond within the trial response window in >25% of trials in both pre- and post-operative task sessions.

Thirty-two age and sex-matched healthy controls (17 female, age 72±3.5) were recruited using the SHARE health informatics register (https://www.registerforshare.org/). Healthy controls were screened for a history of significant neurological or psychiatric conditions. Nine of the healthy controls had been tested as part of a previously published study to test a different hypothesis[47].

The study was approved by the local ethics committee (North East Scotland 21/ES/0035). Written consent was obtained from all participants in accordance with the declaration of Helsinki.

	Healthy Controls	Essential Tremor		ET versus HC p-value
		Dundee	Liverpool	
n	32	25	12	
Age	72.1±3.5	72.8±6.1	76.4±4.3	0.13
Gender M:F	20:12	19:5	12:2	0.10*
Disease Duration years	n/a	19.8±8.3	32±20	
CRST (Total)	n/a	54±11	50±11	n/a
Tremor Medication	n/a	Propanolol (4) Primidone (2) Other§± (3)	Primidone (1)	n/a
Treated hemisphere L:R	n/a	22:3	10:2	n/a
ACE/Mini-ACE	92.3±2.5	93.5±3.0	27±2.8	0.46
HADS-D	2.8±1.8	3.2±	4.1±4.2	0.26
HADS-A	2.6±1.6	3.4±	3.1±	0.21

Values expressed as mean ± standard deviation. ET, Essential Tremor, CRST, Clinical Rating Scale for Tremor, HADS, Hospital Anxiety and Depression Rating Scale, MoCA, ACE-R, Addenbrooks Cognitive Examination. All p-values calculated using Two-Tailed unpaired T-test, *Fisher's exact test. § Other medications included Clonazepam (n = 2) Zonisamide (n = 1).

Procedure

Patients performed the restless bandit task on a laptop twice. In all patients the first session was performed within 24 hours of their thalamotomy. Due to differences in the clinical workflow between the MRgFUS sites (inpatient treatment in Dundee, day-case in Liverpool) Dundee patients reperformed the task 24 hours after their procedure. Liverpool patients were tested 2-3 hours after their thalamotomy. Healthy controls performed the task once.

Experimental design

Participants performed the restless four-armed bandit task[47]. Subjects were given written instructions on how to perform the task and were informed that with each trial they could win between 0 and 100 points and agreed to maximise outcome points.

Each trial started with presentation of four different coloured squares with all four bandit's levers in the upright position representing the four choice options (Fig.1). Participants made their selection using a four key mini-keyboard (Ecarke-EU) with the colour of each button corresponding to a square presented on the computer monitor corresponding to each of the "bandits". If a button press was not made within a 1.5 second response deadline, a large red "X" was displayed for 4.2 seconds at the centre of the screen. These trials were designated as missed trials and no outcome feedback was provided. For choices made within the response deadline, the chosen bandit was highlighted with its

lever shown depressed and a checker-board pattern appeared at the centre of the bandit's square. After a 3 second waiting time, this pattern was replaced by the outcome number of points earned on that trial, in the centre of the chosen bandit's square for 1 second. Then the bandit image disappeared and was replaced by a fixation cross until 6 s after the trial onset, followed by a jittered inter-trial interval (Poisson distribution, mean: 2 s (0–5 s)) before the next trial was started. The payout (outcome) schedule of each of the four bandit choices varied according to a decaying Gaussian random walk. We used three instantiations from Daw et. al., (2006) [14] for the two experimental sessions. To prevent unlikely chance of between session learning effect, we randomly assigned different walk to each patients testing session. Healthy controls performed the same Gaussian walk on each session.

The task consisted of 300 trials. Participants were given short breaks after every 75 trials to improve concentration and task engagement.

Analysis of behavioural performance

Model free metrics of behavioural performance in the task included best bandit choice (i.e. choosing the option with the most points), decision time and the probability of missing a trial. A choice made to the same bandit on two consecutive trials was defined as a stay choice. Decision time was defined as the time between the four bandits being presented and the participants' button press on that trial. Where a measure of behavioural performance is expressed as a probability, this was achieved by dividing by the total number of trials correctly executed within response deadline, either in a 50-trial block, or by dividing this by the total number of responses made across the whole task.

Statistical analysis

We used a repeated measured design ANOVA with a fixed effect (between subject) of group with three levels, (Pre-thalamotomy, Post-thalamotomy, and healthy control) and a random effect (within subject) variable of within task block with (six, 50-trial blocks). All results are reported as mean values ± S.E.M. Bayesian distributions are expressed as mean posterior, prior and posterior difference of the mean with 95% higher density intervals. Correction for multiple comparisons was performed using the Bonferroni method.

Computational modelling of decision making

To determine whether the thalamotomy modified decision making based upon neural value estimates of the bandits, we fitted eight computational model variations of the decision process[9].

Each model variant used one of two learning rules, (Delta rule, or Bayesian learner) combined with one of four choice rules: i) Softmax (SM), ii) Softmax with exploration bonus (SME), iii) Softmax with perseveration bonus (SMP), (iv) Softmax with exploration and perseveration bonuses (SMEP). Posterior parameter distributions were estimated for each subject for each of the free parameters specific to the learning (α in the Delta learning rule) and choice rules for each model variant (β , φ, ρ).

Details of the model and fitting procedure are provided in Supplementary Material.

Lesion mapping analysis

We acquired both pre- and post-operative imaging using a 3T Siemens Prisma Fit Scanner (Siemens Healthineers, Erlangen, Germany) and proprietary 20 channel head and neck coil. Data used in lesion mapping analysis included: Pre-operative volumetric T_1-weighted images using a Magnetization Prepared Rapid Gradient Echo (MPRAGE) sequence, [Echo Time (TE)/repetition time (TR) = 2.32/2300 ms, inversion time (TI) = 900 ms, flip angle = 8°, field of view (FOV) = 240 mm, isotropic voxels 0.9x0.9x0.9mm3, acquisition time (TA) = 5min21sec]. Diffusion tensor imaging (DTI/dMRI) acquired pre- and post-operatively using a spin echo, echo planar imaging sequence (EPI), diffusion weighting b = 0 s/mm² and in 30 uniformly distributed directions (b = 1000 s/mm²) and TE/TR = 93.0/3100ms, averages = 4, FOV = 230 mm, voxels = 1.8x1.8x5.0mm3, TA = 6min35sec]. Post-operative demarcation of the thalamotomy used a T_2-weighted FLAIR-SPACE sequence [TE/ TR = 387 /5000 ms, TI = 1600 ms, FOV = 230 mm, voxel size = 0.4 x 0.4 x 0.9 mm3, TA = 5min 42sec]. Patients treated in Liverpool where MRI was performed within hours of their procedure were not included in the lesion-behaviour mapping as full maturation of the thalamic oedema does not occur and is not maximal until at least 24hours post proceedure[46]. We therefore limited our analysis mapping the relationship of this factor and changes in task behaviour, to the patients treated in Dundee, for whom post- operative MRI and behavioural data were acquired concurrently, allowing oedema definition on MRI.

Post-operative T_2-weighted images were used to manually delineate the voxels enveloped by thalamotomy oedema to create a binary mask of the lesion. The thalamotomy mask was demarcated to include the necrotic lesion core (hypointense), surrounding cytotoxic oedema (markedly hyperintense) and the outer margin by the surrounding perilesional oedema (moderately hyperintense)[48]. Masking was performed following normalisation into Montreal Neurological Institute (MNI) space using SPM12 (https://www.fil.ion.ucl.ac.uk/spm/software/spm12). A second “lesion-nucleus overlap” mask was then generated for each patient which consisted of voxels with shared MNI co-ordinates with individual thalamic nuclei defined using the THOMAS[49] thalamic atlas in Lead-DBS V2.5 (https://www.lead-dbs.org)[50]. In addition to the targeted VIM nucleus, overlap masks were created for the MD[51], CM[52] and VLa[53] as all three have previously associated with cognitive function and are in close proximity to the VIM.

The individual patient masks for each nucleus where then subjected to voxel-wise statistical mapping to define probability mask by averaging across voxels in MNI space. To define the “sweetspot” of maximal behavioural influence of thalamotomy, we weighted each patient’s mask with the measured behavioural outcome and subjected each voxel to a Wilcoxon signed rank test. Correction for false discovery rate used a *q* value of 0.05 applying the Benjamini & Hochberg method for multiple comparisons[54].

Functional connectivity behavioural analysis

Each patient's lesion oedema-nucleus overlap mask was used as seed regions to estimate individual functional connectivity profile maps using the openly available group connectome derived from resting state functional connectivity MRI data of 1000 healthy volunteers[55,56]. We then used each patient's connectivity map to perform voxel-wise correlations between the estimated connectivity profile and the change in behaviour in the restless bandit task[23,57]. The resulting "R-map" therefore represents the visualisation of brain regions where high levels of functional connectivity correlated (either positively or negatively) with the behavioural effect of thalamotomy. To test the validity of this map's predictions, we correlated each individual's connectivity profile with the R-map's correlation coefficients across every brain voxel. This produced an estimate of the spatial similarity between the patient's functional connectivity profile and the group level R-map. For the predictive value of an R-map to be meaningful, the change in behaviour should correlate linearly with the spatial similarity of the individuals connectivity to the original R-map[69]. To test for correlations that might arise by chance, we calculated spatial similarly estimates after 1000 permutations of the behavioural data and re-estimated this the correlation coefficient. This analysis was repeated for each thalamic nucleus (MD, VIM, VLa & CM). This allowed us to build a profile of how much individual variation in the effect of thalamotomy on behaviour, could be explained by the functional connectivity profile of oedema extension into each nuclei.

Probabilistic tractography analysis

Patients pre-operative T_1-weighted and post-operative T_2-weighted images were co-registered using lead DBS. The co-registered pre-operative MRI was then used to segment Pre-frontal cortical regions and individual thalamic nuclei using Freesurfer[59,60]. Transforming the same binary masks used in the functional connectivity analysis, back into native space, we then created a second mask, 'MD-thalamotomy overlap', which consisted of voxels where thalamotomy oedema overlapped with the patients MD nucleus.

Pre-operative dMRI images were subject to skull stripping, bias field correction and corrected for geometric distortion using the Brainsuite Diffusion Pipeline(https://brainsuite.org/processing/diffusion). Co-registration of the dMRI to the pre-operative T_1-weighted image, and then registration to MNI space were performed using Advanced Normalisation Tools[61]. White matter fibre orientations were estimated using FSL's *bedpostx* and streamline densities using *probtrackx*[62]. The MD-thalamotomy overlap mask and a Freesurfer segmented ROI within the prefrontal cortex (frontal pole and rostral middle frontal gyrus) were defined as the seed and waypoint regions, respectively, for probabilistic tractography. Cortical ROI's were defined using the Desikan-Killiany Atlas[63]. Individual streamline density map's were wrapped back into MNI space before voxel-wise GLM analysis with FSL's *randomise* function using a one-sample T-Test or a regression design matrix. Clusters were defined using Threshold-Free Cluster Enhancement (TFCE) after 5000 permutations[64].

Results

Behavioural & computational modelling results

Our experimental design tested two mutually exclusive hypotheses: Firstly, if the higher order thalamic nuclei are critical to maintaining cognitive flexibility, then thalamotomy associated oedema should disrupt this function by deafferenting cortical input from and output to the thalamus. Secondly, by testing patients with a task that relies upon learning from, and encoding of, rewarding feedback in the form of a neural value estimate, we could examine the extent to which the thalamus regulates reward sensitivity. This second hypothesis was further enriched by fitting reinforcement learning model variants[9] which can distinguish between a pure influence on reward sensitivity rather cognitive inflexibility mediated by perseveration.

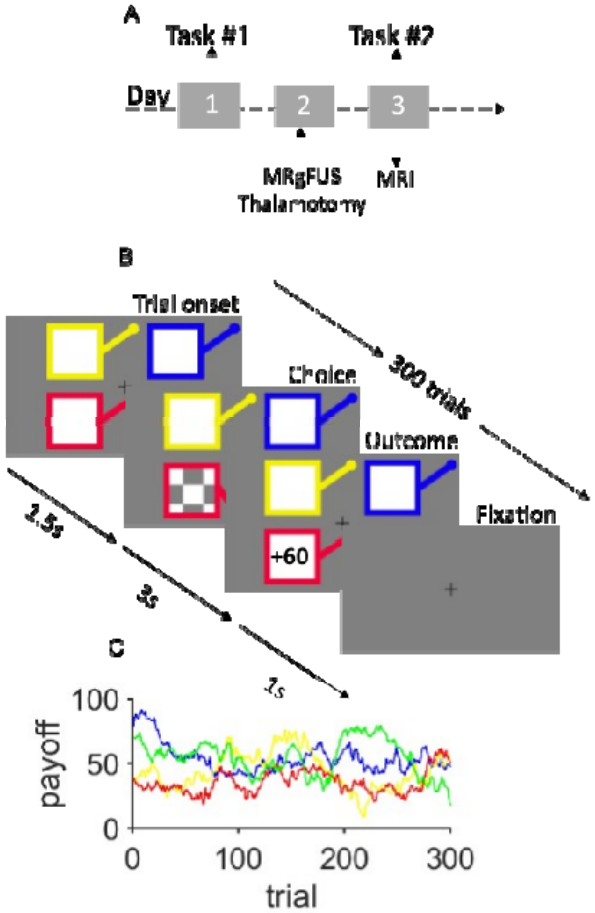

Fig. 1 Restless bandit Task (**A**) Patients underwent testing in the task before and after focused ultrasound (MRgFUS) thalamotomy.(**B**) Each trial has a fixed trial length of 6 seconds with a variable intertrial interval designated by the time between the fixation cross and the onset of the trial (mean 2 seconds). At trial onset, four coloured squares (bandits) were presented. The participant selected one bandit within 1.5 s, which was then highlighted by the bandit lever depressing and a chequer board appearing (Choice screen). After a 3 s delay the, the outcome of the choice, as number of "points" won, was displayed for 1 s. In trials where the participant failed to make a choice within 1.5 seconds the choice screen was replaced by a large red cross (not shown) signifying a missed trial. (**C**) Example of the underlying payout (reward) structure across the 300 trials of the task for each of the 4 bandits. The payout varied from one trial to the next by a gaussian walk.

With this motivation in mind, we tested thirty-seven patients with Essential Tremor (ET), in a prospective, case-control design (Fig.1A). Performance was compared to a group of thirty-two age-matched healthy controls. Each participant performed the 4-armed restless bandit task (Fig.1B), 24 hours before and within 24 hours of their thalamotomy. The restless bandit was chosen as the guassian walk which underlies feedback (Fig. 1C) for each bandit independently fluctuates on a trial-to-trial basis creating significant choice uncertainty. Choosing the best of the four bandits relies upon a binary choice on each trial, either i) staying with the same bandit, or ii) switching to one of the three alternative bandits. Human behaviour in this task has been can be robustly reproduced by reinforcement learning models[8,9,47]. These generate trial-by-trial estimates of each bandits value based upon which the PFC has been proposed to guide choices[4]. We therefore reasoned that fitting reinforcement learning models to both pre- and post-thalamotomy behaviour in the task would allow us to examine the influence of the thalamus on value-based decision making.

Patients identified the best bandit (0.64±0.03) to similar levels as an age-matched healthy control (HC) group (0.65±0.02, mANOVA, $F(1,340) = 0.08$, $p = 0.77$). Following thalamotomy, patients were also still able to identify the best bandit to similar levels (0.62±0.03, rmANOVA, $F(1,360) = 2.52$, $p=0.08$, Fig. 2A&B). Furthermore, task performance was unaffected by non-specific effects (e.g. diminished attention) of the procedure, as the decision time before (0.67±0.02 seconds) and after (0.65±0.02 seconds) the procedure were no different (rmANOVA, $F(1,360) = 1.02$, $p = 0.35$) , nor where the number of missed trials before (3.7±0.5%) or after (3.5±0.6%, rmANOVA, $F(1,360) = 0.22$, $p = 0.63$), Fig.2E&F. Performance was also unaffected by their pre-operative tremor, as HC group decision time (0.66±0.02 seconds, mANOVA, $F(1,340) = 1.02$, $p = 0.31$) and proportion of missed trials preoperatively were no different to patients pre-operatively (3.2%±0.6, $F(1,340) = 2.3$, $p = 0.1$).

We further examined the effects of thalamotomy on flexibility of decisions by classifying these into either stay or switch choices. The probability of choosing the same bandit on two consecutive trials, across the task, P(stay), increased significantly after thalamotomy (0.75±0.02) relative to pre-operative testing (0.67±0.03, rmANOVA, $F(1,340) = 32.1$, $p<0.001$, Fig2.C). Increased stay choices, post thalamotomy, could not be explained by simple lateralisation bias of a unilateral lesion, as the proportion of choices made to the two bandits on the left and right of the presentation screen (0.45±0.04) was no different to that following thalamotomy (0.42±0.03) (rmANOVA, $F(1,360) = 2.3$, $p = 0.13$). Staying with the same option, rather than switching, could be explained by two mechanisms. Firstly, a pure loss of cognitive flexibility, whereby adaptive, stimulus-response mapping, is replaced with a sub-optimal strategy of choice perseveration. Alternatively, choice repetition could be driven by exploitative decisions which aim to maximise the reward obtained by choosing the options with the highest estimated value. Computational models can be used to distinguish between these by varying how the models learnt value estimate is incorporated into the model's choice rule. For example, preservation, can be modelled by a preservation bonus which multiplies a *single* actions value estimate. In contrast, a reward sensitivity parameter, which multiplies the influences of *all* actions learnt value estimate, can delineate the extent to which choice repetition underpins value exploitation.

We fitted a series of previously validated model variants, which included both a perseveration bonus, reward sensitivity and an exploration bonus, as free parameters[13,14] to the patient's choices. By modelling the value of each of the four options throughout the task, we classified stay/switch choices further, with a trinary categorisation based upon the model's value estimate.

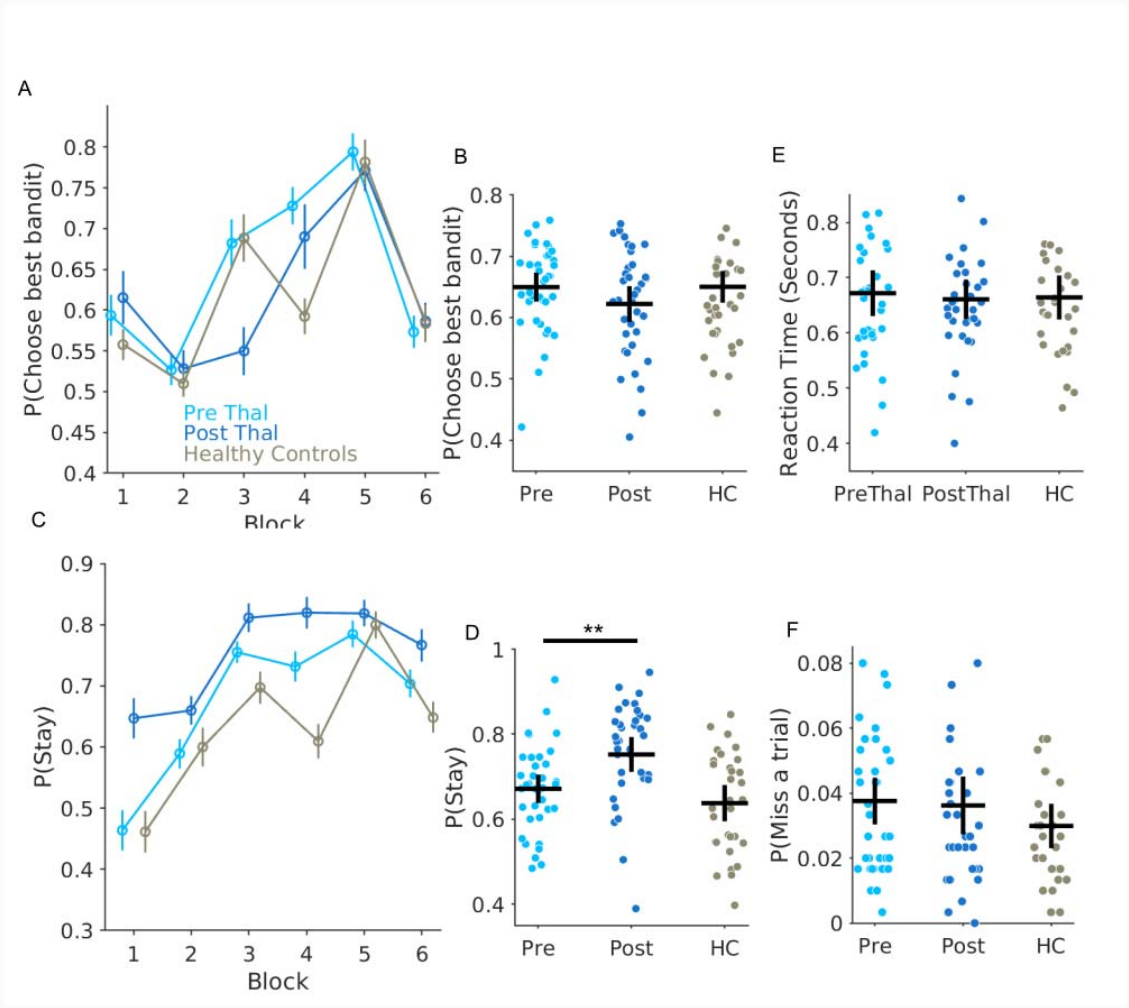

Fig 2. Thalamotomy reduces flexibility of decision making in the restless bandit task. The average probability of choosing the bandit with the highest payout, P(Choose best bandit) is plotted pre (light blue) and post-thalamotomy(dark blue) and in an healthy control group (grey). (**A**) The increase in average values of best bandit choice plotted across six, 50 trial blocks across the task. Vertical lines S.E.M. In **(B**) each circle represents average best bandit choice probability across the task for an individual subject with the horizontal and vertical bar represents the group mean and 95% confidence limits. * p <0.01. (**C**) Probability of stay choices P(stay) across six, 50 trial bins, plotted for pre-, post-thalamotomy and healthy controls. Individual average P(stay) values across the task, as a whole in (**D**), **p<0.001. Average reaction time (**E**) and proportion of trials where the response window was missed (**F**) where unaffected by the thalamotomy.

An exploitative choice refers to the selection of the option perceived to have the highest expected value. A directed exploratory choice is made towards one of the three less valued options, particularly the one with the highest level of uncertainty, and in general, has been least recently selected[15].

Directed exploration scales linearly with the modelled exploration bonus. Random exploratory choices involve the selection of a less valued option regardless of current knowledge or uncertainty about their payout[65]. In contrast to directed exploration, which actively seeks information to update the confidence of a choices value, randomly explored options reflect indifference to the value-estimate and can be linked to decision noise[66]. The effect of thalamotomy on these choice types, was analysed by estimating each subjects the proportion of each choice through the task. The probability of making an exploitative, random or directed exploratory choice were designated as P(Exploit), P(RE) and P(DE) respectively.

Leave one out (LOO) cross-validation estimates[67] were computed for each model variant and to pre-post- operative patient and healthy control groups. Consistent with prior studies[9,47], the Bayesian Learning model with terms for directed exploration and perseveration bonuses (Bayes-SMEP) was the model which most accurately represented both the pre-thalamotomy (loo-log likelihood = -0.7456) and control groups (loo-log likelihood = -0.7526), Fig3A. Following thalamotomy, the Bayesian learning model with a perseveration bonus, but no exploration bonus (Bayes-SMP) best explained the patient's choices (loo-log likelihood = -0.6909). Consistent with the winning models capturing decision making in the task robustly, model parameters could still be recovered (Fig.3B-D) from synthetic choice data generated from simulated choices. Simulated data from the winning models also overlapped with the experimental choices, replicating both the proportion of experimentally observed P(Best bandit), Fig3E&G and P(Stay) choices, Fig3F&H, across each of the 6 analysis blocks of the task and across the task as a whole.

Post-thalamotomy, the patients had significantly higher values of the reward sensitivity parameter, β, posterior difference in mean [Contrast: Pre-Thal minus Post-Thal], M_{diff} = 0.015 (HDI = 0.007,0.02), Fig 3.I&K, but no difference in the preservation bonus ρ, M_{diff} = 0.28 (HDI = -0.7, 1.3), Fig3. J&L, Supplementary Table 1.

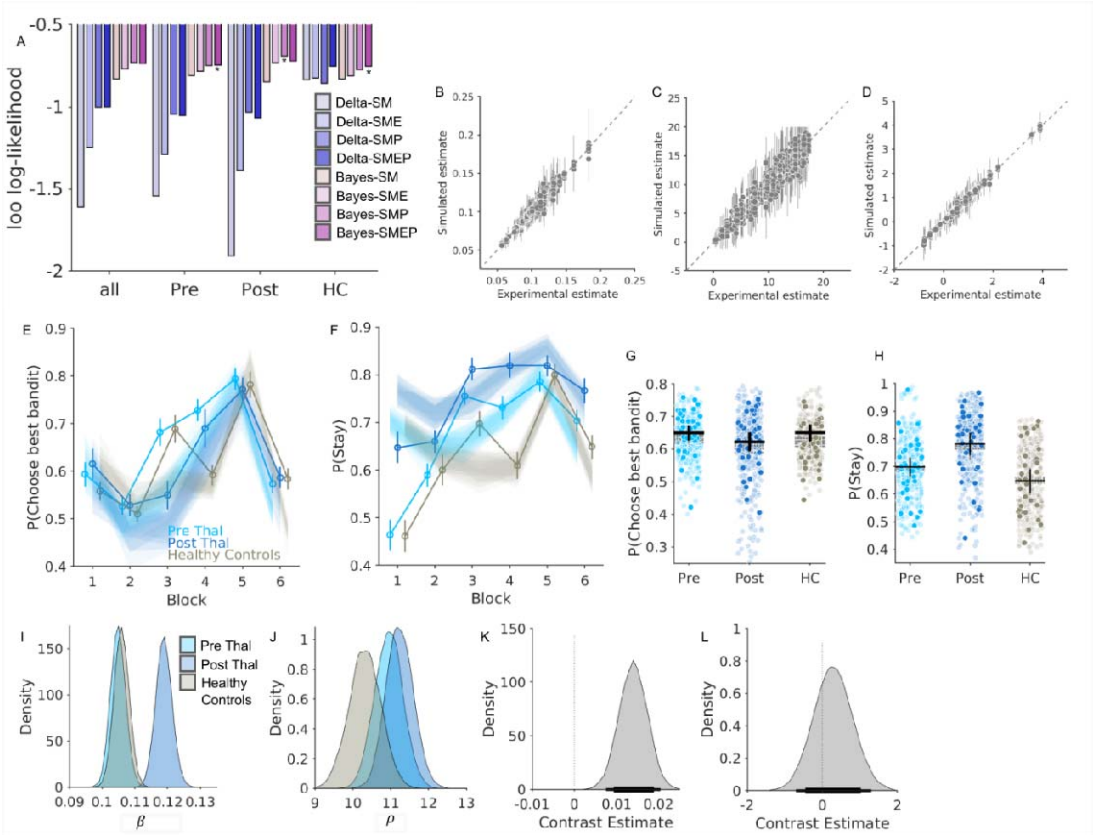

Fig 3. Reinforcement learning model fitting to thalamotomy choice reveals mechanism of cognitive inflexibility. (A) Model fits from 8-model variants fitted to pre-, post- thalamotomy and healthy control groups. Each group's winning model delineated by asterix (Pre-thalamotomy and Healthy controls; Bayes-SMEP, Post-thalamotomy Bayes-SMP). (**B-D**) Each grey dot represents average parameter estimate recovered from fitting 10 simulated choices using the individual's estimated parameter from each groups winning model fitted to the experimental choices. The vertical lines represent the 95% Higher density interval (HDI). Successful parameter recovery for the reward sensitivity parameter, (B), perseveration bonus, ,(C) and exploration bonus, , (D) was indexed by in strong linear correlations between the experimentally estimated parameters and those recovered from simulated choices (; R^2 = 0.97, ; R^2= 0.95, ; R^2 =0.99 , all p<0.001). In (E &F) the experimental best bandit choices P(best bandit) and stay choices, P(stay) are plotted, across six – 50 trial blocks, with the same analysis applied to the simulated model choices for each group. Shaded areas present 95% confidence limits across simulations. (G) Simulated individual P(best bandit) average are in light coloured dots, overlayed with experimental equivalent with dark coloured. Black vertical and horizontal lines are experimental mean and 95% confidence limits, grey dotted lines simulated. (H) Illustration of simulated versus experimental P(stay) values through the task demonstrating the simulated data could replicate the increase in these choices post-thalamotomy.(I) Posterior distributions for , (I) and , (J) illustrated for each group. Posterior differences in the mean for (K) and (L). Thin and thick horizonal black lines represent 95 and 85% HDI.

Using the winning models for each group to divide each choice into the trinary classification described above, patients made more exploitative choices through the task after thalamotomy (0.77±0.01) as before the procedure (0.74±0.01), Fig.4A&D (rmANOVA F(1,360) = 5.36, p = 0.02). Patients also made the less random exploratory choices post-operatively (0.14±0.01) than pre-operatively (0.17±0.01, rmANOVA F(1,360) = 7.54, p < 0.01), Fig. 4B&E. The proportion of directed exploration choices were reduced following the thalamotomy (0.07±0.01) compared to pre-operative (0.1±0.01),

rmAVOVA F(1,360) = 8.08, p < 0.01, Fig4 C&F. These results confirmed that the increase in stay choices post thalamotomy was best explained by several factors. Firstly, reflected by the increase in , choice repetition was driven by a greater extent to which decisions were biased by the neural estimate of all four bandits value. This was expressed into the task behaviour by a switch towards more exploitative choices, to the higher valued option, and reduced random section of lower valued alternatives. Secondly, the reluctance to switch choices can also be explained by an increase in confidence in the accuracy of these value estimates, as thalamotomy lead to significant suppression of directed exploration, a decision strategy expressed when decision confidence is low [65].

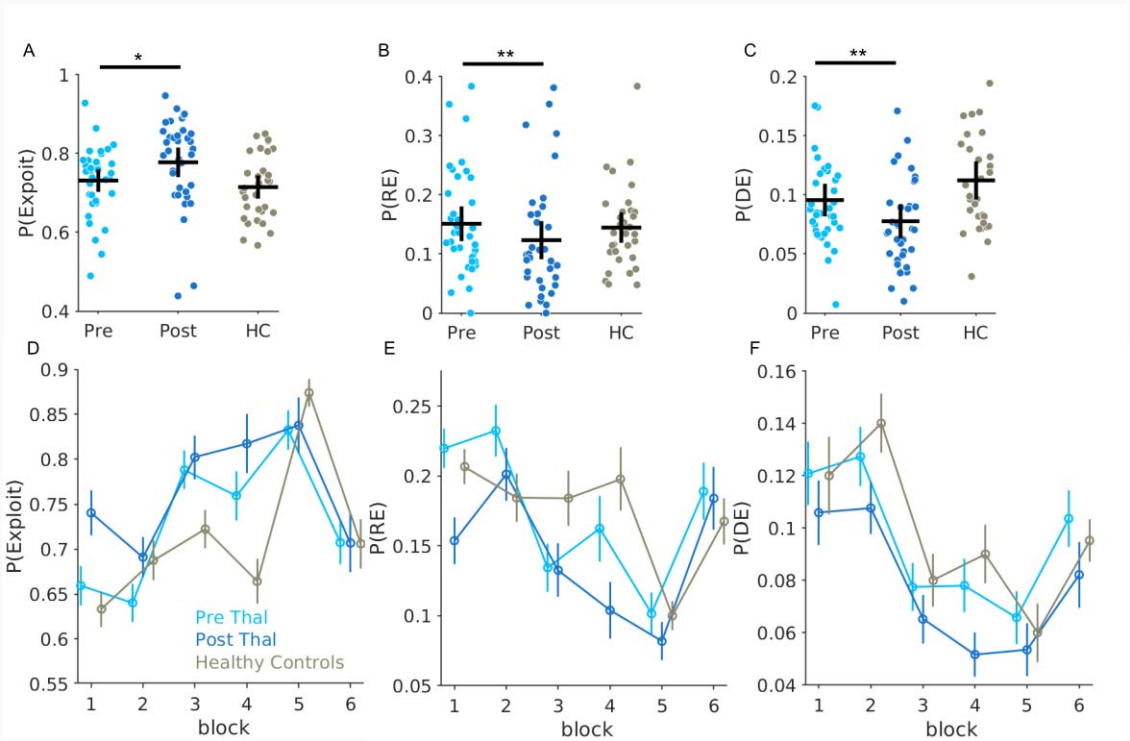

Fig 4. Effect of thalamotomy on value-based decision types. (A) Proportion of choices to the highest value option P(Exploit) across the task in the three groups. (**B**) Proportion of random exploratory choices, P(RE), and (**C**) directed exploratory choices p(DE). Light and dark blue dots represent individual patient before (pre) and after (post) thalamotomy. Healthy controls (HC) in grey. Black horizontal and vertical lines represent group mean and 95% confidence limits. *p<0.05, **p<0.01. The time course across six, 50 trial blocks of each decision type for all three groups is plotted in D to F.

Neuroimaging correlates of thalamotomy induced cognitive inflexibility

Next, we analysed the relationship between the extent of post-operative vasogenic oedema into individual thalamic nuclei and the main behavioural effect of thalamotomy, increased stay choices, on the task. There was a trend for greater oedema extension through the thalamus as a whole (mean oedematous voxels = 2102±1108) correlated, rho(23) = 0.51, Bonferroni adjusted p-value = 0.056) with the increase in P(Stay). Specifically, extension into the MD nucleus (mean voxels = 105±79), rho(23) = 0.65, Bonferroni adjusted p-value = 0.003) but not the VIM, CM or VLa nuclei predicted the individual changes with the increase in P(Stay) (Fig.5C & Supplementary Fig. 1 A-H & Supplementary Table 2). To minimise the likelihood of the behavioural being a consequence of multiple nuclei being

impacted by the thalamotomy we undertook a further multivariate analysis using a GLM including each individuals oedema extension into the four nuclei. Only the number of voxels in the MD nucleus correlated with the increase in P(Stay) $R^2 = 0.43$ p= 0.01 (MD, T = 2.8, p < 0.01; VIM T = 0.71, p = 0.48; CM, T = -0.01, p = 0.98; VLa, T = -0.27, p = 0.7). In contrast, a second GLM which included interactions between the voxel extension into the MD and the other three nuclei was unable to predict individual effect on P(Stay), ($R^2 = 0.28$, p=0.07). We calculated the voxel most likely to be associated with the behavioural effect. Weighting each individual patient MD-lesion mask by the change in perseveration (P(Stay) post – pre) confirmed an infero-lateral portion of the MD nucleus (Fig.5B) was most likely to increase P(stay) (Voxel-Wise Signed Rank Z = 3.8,p<0.05, FDR q = 0.05, MNI co-ordinate x = -7.9, y = -17,z = 3).

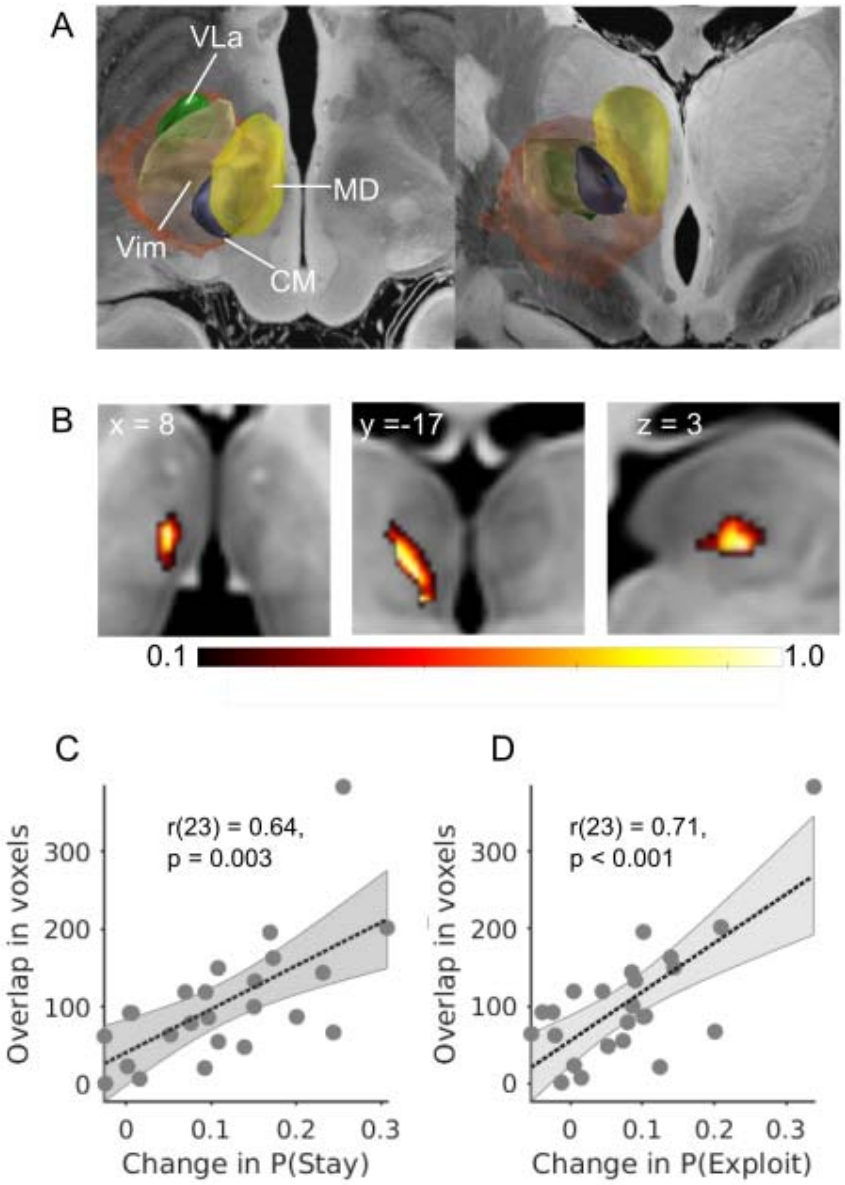

Fig5. Spatial distribution of thalamotomy oedema correlates with changes in behaviour. (**A**) 3D projections of thalamotomy oedema overlap (orange) across all patients in relation to the Ventral intermediate (VIM, brown) and surrounding centromedian (CM, blue), Ventro-lateral anterior (VLa, green) and Mediodorsal (MD,yellow) nuclei. Background from Edlow et al., (2019)[68]. **(B**) Axial, coronal and sagittal slices with the yellow-red heatmap illustrate the probability of lesion oedema extending into the MD nucleus across 24 patients with post-operative imaging at the time of testing. (**C**) Correlations between change in P(Stay) over the task, pre- minus post-thalamotomy, and the number of oedematous voxels from the thalamotomy that overlap into each segmented nucleus surrounding the surgically targeted VIM (segmentation based upon Su et. al., (2019)) [49]. (D) The same analysis applied to the change in P(Exploit) over the task (pre-, minus post-thalamotomy) demonstrating the increase in exploitation co-varied with the extent of oedema extension into the MD nucleus on an individual patient basis.

Using the overlap of thalamotomy oedema into each segmented nuclei (MD, VIM, CM and VLa) as seed regions to a group connectome derived from resting state functional connectivity of healthy volunteers[69], we derived whole-brain functional connectivity maps for each patient (Fig.6). The correlation between the functional connectivity and the thalamotomy induced change in P(Stay) was calculated on a voxel-by-voxel basis. The only seed region whose connectivity pattern could predict this effect was the MD nucleus. The R-Map in Fig.6A depicts thresholded correlations between functional connectivity from thalamotomy oedema volumes to all brain voxels and thalamotomy induced increase in stay choices. Spatial correlation of individual patients' thalamic seed connectivity profiles with this R-Map significantly explained variants in the between subject increase in P(Stay), (R = 0.61, p <0.001, Permutation test: R = 0.65, p = 0.004). The same analysis applied to the VIM, CM or VLa did not reveal statistically significant R-maps which exceeded the permutation testing threshold (VIM; R = 0.32, p = 0.11, Permutation test R = 0.33, p = 0.14, CM; R = 0.40 , p = 0.06 , Permutation test R = 0.38 , p = 0.09, VLa; R = 0.42, p = 0.04, permutation test R = 0.43, p = 0.1). Supplementary Table 3 includes the details of the peak R-map values and voxel co- ordinates of the cortical regions whose connectivity with the MD nucleus predicted increased stay choices.

To further validate the conclusion that the increased stay choices were mediated by lesion effects on the MD nucleus to heighten reward sensitivity, we repeated this analysis using the difference in exploitative choices after thalamotomy as the behavioural regressor. As for the analysis of stay choices, the number of oedematous voxels in the MD nucleus predicted the increased in proportion of exploitative choices, rho(23) = 0.71, Bonferroni adjusted p-value <0.001 (Fig.5D & Supplementary Fig. 1 A-H & Supplementary Table 2). Furthermore, increased exploitation caused by the thalamotomy, could predicted by the functional connectivity profile of the MD nucleus with the rest of the brain (R = 0.61, p < 0.001, Permutation test R = 0.62, p = 0.005, Supplementary Fig. 2 & Supplementary Table 4).

Finally, we performed structural connectivity analysis on individual patient DTI images acquired pre-operatively, to confirm which prefrontal cortical regions shared white matter tract with the region of the mediodorsal nucleus swollen by the thalamotomy. Probabilistic fibre tracking confirmed significant clusters of connectivity between voxels where oedema extended into the MD nucleus and prefrontal cortex which was maximal with the frontal pole (One sample T-test T = 7.74, p<0.001 TFCE-FWE-corrected) (voxels=37884; MNI *x, y, z*: -16,57,-10), Fig.6D, but in addition exhibited significant clusters of connectivity with the DLPFC (p<0.05 TFCE-FWE-corrected, Supplementary Table 5). In a further exploratory analysis, we correlated the change in task behaviour with the thalamo-cortical structural

connectivity profile. Significant clusters of connectivity which positively correlated with both the increase in P(Stay) choices and P(Exploit) choices following thalamotomy were identified DLPFC (BA 9/46), OFC (BA11) and Frontal Pole (all p<0.05 TFCE-uncorrected), Supplementary Tables 6 & 7.

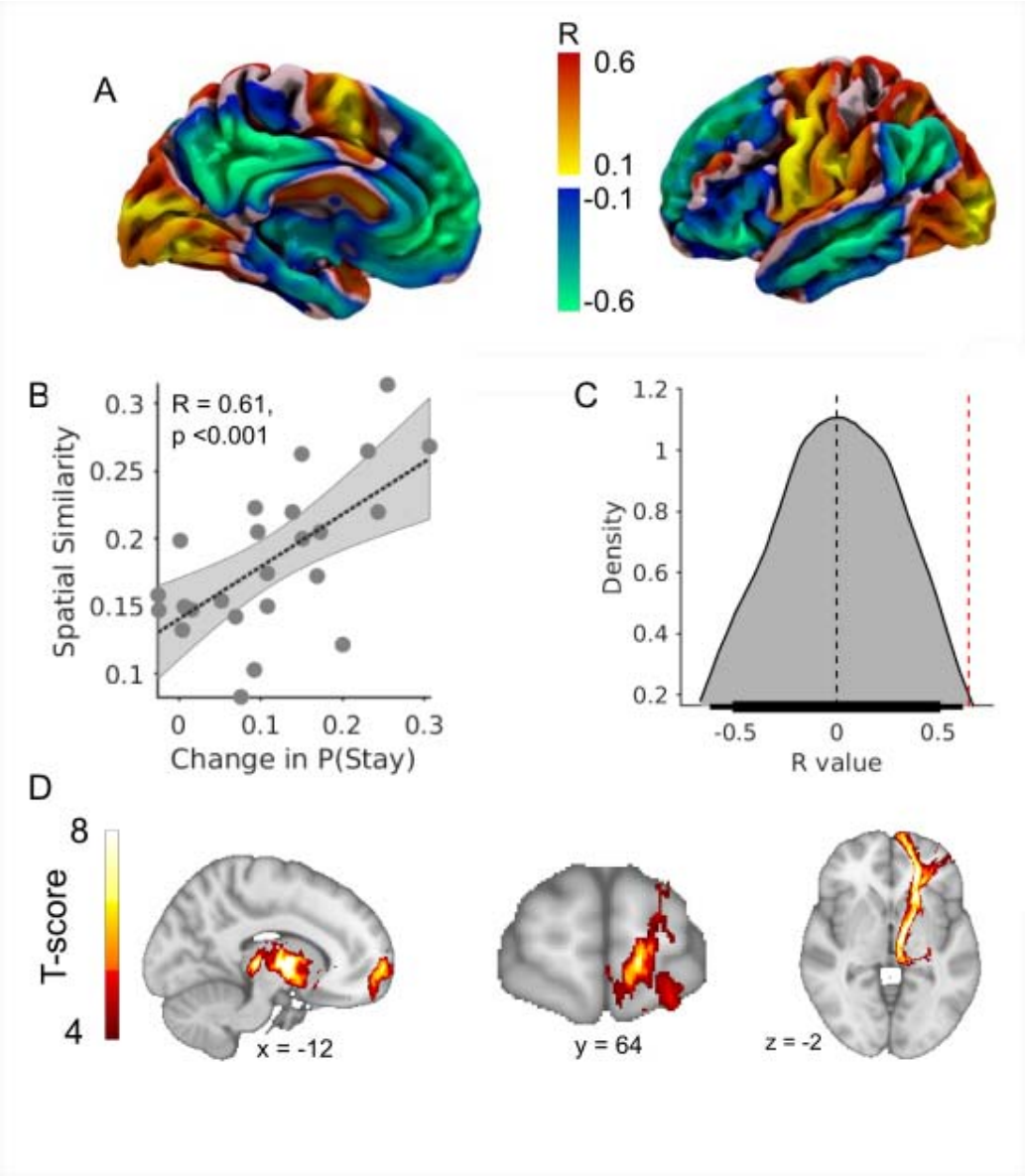

Fig6. Mediodorsal thalamus-prefrontal cortical connectivity predicts thalamotomy effect on choice persistence. Medial (**A**) and lateral cortical surface projections of the whole-brain voxel-wise R- Map demonstrates the optimal functional connectivity profile between thalamotomy oedema extension into the MD nucleus and increase in P(stay). Warm colours show cortical voxels where functional connectivity to the voxels of the MD nucleus associated with increased P(stay). Cool colours indicate voxels where stronger functional connectivity was associated with less marked effect of thalamotomy on the same behaviour. (**B**) The more the individual functional connectivity profile matched the 'optimal' R-Map, the greater was the increase in P(Stay) induced by thalamotomy related oedema into the MD nucleus (R = 0.61, p<0.001). In (**C**) we plot the R value distribution derived from 1000 re-permuted correlations between the individual change in P(Stay) caused by the thalamotomy and the optimal R-map. The probability of seeing the same correlation by chance was p<0.01 (represented by the red dashed vertical line). Horizontal thin and thick lines represent 95 and 85% HDI. (**D**) Group level T-scores of structural connectivity between thalamotomy overlap with MD nucleus and the prefrontal cortex.peak connectivity in broadman area 10, Frontal pole

MNI T = T = 7.14 [-16, 57, -10], p_{FWE} <0.001. Colour bar thresholded for all voxels with significant connectivity FWE corrected at p<0.05.

Discussion

In this study we examined the immediate behavioural consequences of MRgFUS thalamotomy on the 4-armed restless bandit task. A significant body of research has implicated the prefrontal cortex as the core neural processing hub where neural estimate of a decision's value is integrated into the eventual decision, guiding selection towards the optimal goal[12,41]. Via constant updates, with experiential learning, decisions are adjusted to rapidly changing environmental conditions[8]. Prior to this study, whether the thalamus might influence the way decision value estimates are integrated into human behaviour has been proposed but not firmly established[71]. Here we provide causal evidence that the immediate consequence of transiently deafferenting the cortico-thalamic circuit is a marked reduction in decision flexibility. This observation would support a role for the physiologically intact thalamo-cortical circuit in maintenance of choice flexibility, a result supported by clinical observations that medial thalamic lesions typically mimic the effects of those seen in lesions of the prefrontal cortex (such as frontal perseverative syndrome)[72]. A simple explanation such as cognitive inattention or laterality bias could not explain this effect, as our patients were no slower, less accurate, nor more inclined to choose bandits displayed to one hemifield over another, than before thalamotomy.

By fitting a computational model to the patient's behaviour, our results further elucidate the specific influence of the thalamus on decision flexibility. We included model variants whose free parameters, reward sensitivity (β), exploration, (φ) and preservation, (ρ), bonuses, allowed us to identify how the thalamotomy influenced different aspects of the decision value estimate. Cognitive inflexibility, in the form of response perseveration, as classically described in prefrontal lesions, is unlikely as ρ, which augments a *single* action's value estimate to promote choice repetition, was unaffected by thalamotomy. Instead, simulated choices using the winning model (Bayes-SMP, Fig.3F&H), could reproduce the effect of thalamotomy only with a significant increase in the reward sensitivity parameter β. Because this multiplies the value estimate for *all* actions, decreased decision flexibility is better explained by an overall increase in the participants choices being influenced by the learnt value estimates across all options rather than one option in isolation. By this mechanism, our modelling suggests that following thalamotomy, our patients make more stay choices because they have greater confidence that the neural value estimates are a reliable guide with which to maximise reward seeking. Accordingly, we found a shift towards greater exploitative choices and less exploratory ones, consistent with heighted decision confidence mediating an effect on explore-exploit trade-off[73]. Interestingly, this interpretation is consistent with a proposed role for thalamo-prefrontal coupling in promoting changes in decision strategy under conditions of low decision confidence and high perceptual uncertainty[74,75]. By decoupling this negative feedback loop, thalamotomy effectively raised the threshold for questioning the need to change decision, leading to a greater likelihood of repeating the same choice.

We also found a significant reduction in the proportion of directed exploratory (DE) choices. This was paralleled with a change in the winning model to one with no exploration bonus (Bayes-SMP) as this

parameter linearly scales the proportion of DE choices in the model uses[9]. Choices which explore lower valued bandits are classified as directed, rather than randomly explored, when these are made to the option which has the greatest choice uncertainty. By actively choosing the bandit with least recent sampling history, the information gained from a directed exploration is used to update and reduce uncertainty in the agent's model, in this case the value estimates, of the environment's structure. The effect of the thalamotomy to substantially reduce this type of choice, can be considered analogous to diminishing curiosity[76] or reflective, prospective decision-making[77] about the reward environment. In combination with an overinflated estimate of the bandit's value (from the increase in β), the behavioural indifference to change leads to maintenance of a decision *status-quo*. These results are consistent with the role of basal-ganglia and thalamus in regulating the extent to which metacognitive variables such as confidence are expressed in behaviour[12,23,78,79].

Our neuroimaging analysis of the effect of thalamotomy oedema on stay choices would support that the medio-dorsal nucleus is the most likely structure mediating this behavioural effect in our patients. We found that individual variation in the effect of thalamotomy on the proportion of stay choices correlated, with; i) the volume of oedema into this nucleus, ii) the functional connectivity of this nucleus with the rest of the brain, iii) white matter tract connectivity between the MD nucleus and PFC. This was not the case for any other thalamic nucleus, including the clinical target (VIM) for tremor control using MRgFUS. The MD nucleus is known to have the largest reciprocal connectivity profile with the prefrontal cortex[28,31,42]. It is also increasingly recognised as a regulator of information coding within the pre-frontal cortex relevant to decision making[31,40,42]. In rodent models, optogenetic activation of MD nucleus neurons promotes transition in PFC population coding which leads to switching and exploratory behaviour[80]. In humans, lesions of the MD nucleus lead to diminished behavioural flexibility and difficulty with tasks that require rapid cognitive switching which parallel those seen in our patients post thalamotomy[81]. Probabilistic analysis of the day one lesion oedema at the time of testing, demonstrated this most likely was mediated by oedema extending into an infero-lateral portion of the nucleus. The lateral, parvocellular (MDpc) portion projects to the lateral regions of the pre-frontal cortex including Broadman area 46/9 (DLPFC, premotor and SMA/dACC)[28,29,35,82]. These are areas generally considered part of a cognitive control network which encode in neural activity predicting decisions to stay or switch [83–85]. Cortical topography of our functional connectivity analysis showed strong positive correlations with this network whilst negative correlations were observed with the ventral, medial and orbital PFC. Interestingly, these regions project to and receive input from the medial or magnellocellar portion of the MD (MDmc)[35,36]. Lesions of the MDmc lead to premature switching and reward insensitivity [41], which is the opposite behavioural effect to what we observed in our patients. These contrasting behavioural effects may suggest a marked difference in the functional role of MDmc and MDpc on value-based decision making which is mediated by their distinct connectivity profiles to different regions of the PFC.

Clinical implications

MRgFUS lesion formation has been recognised for its unpredictability. One series reported 18% incidence of lateral tail extension of the lesion into the internal capsule[86]. These were associated with a high rate of motor complications which are more obvious clinically than cognitive changes that might be expected from medial extension into the MD nucleus. Previous reports may therefore have under reported the impact of MRgFUS lesions on cognition by relying upon cognitive batteries insensitive to set-shifting errors analogous to the transient effects of focused ultrasound heating on the MD nucleus seen here. This is likely to be particularly relevant to patients undergoing staged bilateral lesioning where the capacity for cognitive reserve may be less robust[87,88].

When performing the restless bandit task, Parkinson's disease patients who are apathetic are impaired at both choosing the highest valued bandit and persisting in this once it has been identified. This reward insensitivity is manifest behaviourally by an increased exploration and a lower proportion of stay choices[47] . This decision-making signature of apathy in PD is opposite to the effect of thalamotomy, and specifically oedema extension into the MD nucleus, which drove up stay choices, increased reward sensitivity, and shifted the explore-exploit trade-off towards exploitation. This raises the possibility that neuromodulation, targeting the inferolateral MD nucleus may be a viable future treatment to restore value-based learning and motivation in Parkinson's disease. In support of this, Gilmour et., al., (2024) also identified a correlate of preserved motivation in non-apathetic PD that was suppression of thalamic activity during exploit choices. Deafferentation of thalamic negative feedback to PFC, though compensatory mechanisms, or neuromodulation, may promote consolidation of value-estimates so that the value of learnt experience can be expressed into motivated goal-directed behaviour.

Study Limitations

Thalamotomy performed using MRgFUS deliberately targets the motor VIM nucleus. Our study relies upon the indirect extension of vasogenic oedema outside of this target and into surrounding structures, including MD nucleus which we propose mediates the behavioural effects on our task. Although unlikely, given in our voxel-based analysis we found no additional predictive value of including combinations of surrounding nuclei along with the MD, we cannot definitively exclude the possibility that oedema extending to multiple nuclei led to the observed change in task behaviour. Future studies should aim to address this by selective, anatomically circumscribed, focused neuromodulation of the MD nucleus to replicate the behaviour seen here.

Conclusions

By prospectively studying a group of patients undergoing ultrasound thalamotomy we have demonstrated a causal role for the thalamus in influencing the cognitive strategy necessary for flexible reward-based learning. The increase in reward sensitivity leading to reduction in choice flexibility after thalamotomy, may suggest that the thalamus, and specifically the MD nucleus, can actively undermine the extent to which models of value, generated in pre-frontal cortex, influence decision making. In turn, this may support the role of the intact thalamus in biasing choices away from an

overconfidence in the extent to which these models are incorporated into decision making behaviour. Under conditions of volatility and decision uncertainty, this function may contribute to optimally tune the level of decision flexibility necessary to maximise reward acquisition[89,90].

Data Availability

Behavioural data from these experiments and python scripts used for computational modelling are openly available at https://osf.io/w3x7d/.

Acknowledgements

The authors would like to thank both the DBS team at the Queen Elizabeth University Hospital, Glasgow (Dr's Vicky Marshall, Michael Canty & Edward Newman) and the MRgFUS team of the Walton Centre, Liverpool (Dr Jay Panicker, Dr Dinesh Damodaran, Dr Mark Radon & Dr Rajesha Srinivasaiah) who's collective clinical efforts in patient selection allowed this research to be undertaken. We would also like to thank the patients who participated in this study.

Funding

This work was supported by the University of Dundee Movement Disorders research endowment fund and a Tenovus Scotland research grant to TPG.

Competing Interests

The authors report no competing interests.

References

1. Passingham, R. *Understanding the Prefrontal Cortex: Selective Advantage, Connectivity, and Neural Operations*. (Oxford University Press, 2021).

2. Rudebeck, P. H., Bannerman, D. M. & Rushworth, M. F. S. The contribution of distinct subregions of the ventromedial frontal cortex to emotion, social behavior, and decision making. *Cogn. Affect. Behav. Neurosci.* **8**, 485–497 (2008).

3. Rudebeck, P. H. *et al.* Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J. Neurosci.* **28**, 13775–13785 (2008).

4. Rushworth, M. F., Noonan, M. P., Boorman, E. D., Walton, M. E. & Behrens, T. E. Frontal cortex and reward-guided learning and decision-making. *Neuron* **70**, 1054–1069 (2011).

5. Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* **9**, 545–556 (2008).

6. Gläscher, J. *et al.* Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. *Proc. Natl. Acad. Sci.* **109**, 14681–14686 (2012).

7. Rushworth, M. F. S., Noonan, M. P., Boorman, E. D., Walton, M. E. & Behrens, T. E. Frontal Cortex and Reward-Guided Learning and Decision-Making. *Neuron* **70**, 1054–1069 (2011).

8. Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879 (2006).

9. Chakroun, K., Mathar, D., Wiehler, A., Ganzer, F. & Peters, J. Dopaminergic modulation of the exploration/exploitation trade-off in human decision-making. *Elife* **9**, e51260 (2020).

10. Wilson, R. C., Bonawitz, E., Costa, V. D. & Ebitz, R. B. Balancing exploration and exploitation with information and randomization. *Curr. Opin. Behav. Sci.* **38**, 49–56 (2021).

11. Tomov, M. S., Truong, V. Q., Hundia, R. A. & Gershman, S. J. Dissociable neural correlates of uncertainty underlie different exploration strategies. *Nat. Commun.* **11**, 2371 (2020).

12. Gilbertson, T. & Steele, D. Tonic dopamine, uncertainty and basal ganglia action selection. *Neuroscience* **466**, 109–124 (2021).

13. Hunt, L. T. *et al.* Mechanisms underlying cortical activity during value-guided choice. *Nat. Neurosci.* **15**, 470–476 (2012).

14. Schwartenbeck, P. *et al.* Computational mechanisms of curiosity and goal-directed exploration. *eLife* **8**, e41703 (2019).

15. Gershman, S. J. Deconstructing the human algorithms for exploration. *Cognition* **173**, 34–42 (2018).

16. Exploration Disrupts Choice-Predictive Signals and Alters Dynamics in Prefrontal Cortex: Neuron. https://www.cell.com/neuron/fulltext/S0896-6273%2817%2931130-3?elsca1=etoc&amp%3Belsca2=email&amp%3Belsca3=0896-6273_20180117_97_2_&amp%3Belsca4=Cell+Press&code=cell-site.

17. Hogeveen, J. *et al.* The neurocomputational bases of explore-exploit decision-making. *Neuron* **110**, 1869-1879.e5 (2022).

18. Dixon, M. L. & Christoff, K. The lateral prefrontal cortex and complex value-based learning and decision making. *Neurosci. Biobehav. Rev.* **45**, 9–18 (2014).

19. Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. Conflict monitoring and cognitive control. *Psychol. Rev.* **108**, 624–652 (2001).

20. MacDonald, A. W., Cohen, J. D., Stenger, V. A. & Carter, C. S. Dissociating the Role of the Dorsolateral Prefrontal and Anterior Cingulate Cortex in Cognitive Control. *Science* **288**, 1835–1838 (2000).

21. Gläscher, J., Hampton, A. N. & O'Doherty, J. P. Determining a Role for Ventromedial Prefrontal Cortex in Encoding Action-Based Value Signals During Reward-Related Decision Making. *Cereb. Cortex* **19**, 483–495 (2009).

22. Mc, K.-F., A, B. & Mfs, R. Medial and orbital frontal cortex in decision-making and flexible behavior. *Neuron* (2022).

23. de A Marcelino, A. L. *et al.* Pallidal neuromodulation of the explore/exploit trade-off in decision-making. *Elife* **12**, e79642 (2023).

24. Yang, W., Tipparaju, S. L., Chen, G. & Li, N. Thalamus-driven functional populations in frontal cortex support decision-making. *Nat. Neurosci.* **25**, 1339–1352 (2022).

25. Shine, J. M., Lewis, L. D., Garrett, D. D. & Hwang, K. The impact of the human thalamus on brain-wide information processing. *Nat. Rev. Neurosci.* **24**, 416–430 (2023).

26. Sherman, S. M. & Guillery, R. W. *Exploring the Thalamus and Its Role in Cortical Function*. (MIT press, 2006).

27. de Bourbon-Teles, J. *et al.* Thalamic Control of Human Attention Driven by Memory and Learning. *Curr. Biol.* **24**, 993–999 (2014).

28. Perry, B. A., Lomi, E. & Mitchell, A. S. Thalamocortical interactions in cognition and disease: The mediodorsal and anterior thalamic nuclei. *Neurosci. Biobehav. Rev.* **130**, 162–177 (2021).

29. Mitchell, A. S. The mediodorsal thalamus as a higher order thalamic relay nucleus important for learning and decision-making. *Neurosci. Biobehav. Rev.* **54**, 76–88 (2015).

30. Pergola, G. *et al.* The regulatory role of the human mediodorsal thalamus. *Trends Cogn. Sci.* **22**, 1011–1025 (2018).

31. Wolff, M. & Halassa, M. M. The mediodorsal thalamus in executive control. *Neuron* (2024).

32. Goldman-Rakic, P. S. & Porrino, L. J. The primate mediodorsal (MD) nucleus and its projection to the frontal lobe. *J. Comp. Neurol.* **242**, 535–560 (1985).

33. Daum, I. & Ackermann, H. Frontal-Type Memory Impairment Associated with Thalamic Damage. *Int. J. Neurosci.* **77**, 187–198 (1994).

34. Van der Werf, Y. D., Witter, M. P., Uylings, H. B. & Jolles, J. Neuropsychology of infarctions in the thalamus: a review. *Neuropsychologia* **38**, 613–627 (2000).

35. Mengxing, L., Lerma-Usabiaga, G., Clascá, F. & Paz-Alonso, P. M. High-Resolution Tractography Protocol to Investigate the Pathways between Human Mediodorsal Thalamic Nucleus and Prefrontal Cortex. *J. Neurosci.* **43**, 7780–7798 (2023).

36. Klein, J. C. *et al.* Topography of connections between human prefrontal cortex and mediodorsal thalamus studied with diffusion tractography. *NeuroImage* **51**, 555–564 (2010).

37. Liebermann, D., Ploner, C. J., Kraft, A., Kopp, U. A. & Ostendorf, F. A dysexecutive syndrome of the medial thalamus. *Cortex* **49**, 40–49 (2013).

38. Mukherjee, A., Lam, N. H., Wimmer, R. D. & Halassa, M. M. Thalamic circuits for independent control of prefrontal signal and noise. *Nature* **600**, 100–104 (2021).

39. Schmitt, L. I. *et al.* Thalamic amplification of cortical connectivity sustains attentional control. *Nature* **545**, 219–223 (2017).

40. Halassa, M. M. & Kastner, S. Thalamic functions in distributed cognitive control. *Nat. Neurosci.* **20**, 1669–1679 (2017).

41. Chakraborty, S., Kolling, N., Walton, M. E. & Mitchell, A. S. Critical role for the mediodorsal thalamus in permitting rapid reward-guided updating in stochastic reward environments. *eLife* **5**, e13588 (2016).

42. Pergola, G. *et al.* The Regulatory Role of the Human Mediodorsal Thalamus. *Trends Cogn. Sci.* **22**, 1011–1025 (2018).

43. Collomb-Clerc, A. *et al.* Human thalamic low-frequency oscillations correlate with expected value and outcomes during reinforcement learning. *Nat. Commun.* **14**, 6534 (2023).

44. Elias, W. J. *et al.* A Randomized Trial of Focused Ultrasound Thalamotomy for Essential Tremor. *N. Engl. J. Med.* **375**, 730–739 (2016).

45. Morel, A., Magnin, M. & Jeanmonod, D. Multiarchitectonic and stereotactic atlas of the human thalamus. *J. Comp. Neurol.* **387**, 588–630 (1997).

46. Harary, M., Essayed, W. I., Valdes, P. A., McDannold, N. & Cosgrove, G. R. Volumetric analysis of magnetic resonance–guided focused ultrasound thalamotomy lesions. *Neurosurg. Focus* **44**, E6 (2018).

47. Gilmour, W. *et al.* Impaired value-based decision-making in Parkinson's disease apathy. *Brain* **147**, 1362–1376 (2024).

48. Wintermark, M. *et al.* Imaging Findings in MR Imaging–Guided Focused Ultrasound Treatment for Patients with Essential Tremor. *Am. J. Neuroradiol.* **35**, 891–896 (2014).

49. Su, J. H. *et al.* Thalamus Optimized Multi Atlas Segmentation (THOMAS): fast, fully automated segmentation of thalamic nuclei from structural MRI. *NeuroImage* **194**, 272–282 (2019).

50. Horn, A. *et al.* Lead-DBS v2: Towards a comprehensive pipeline for deep brain stimulation imaging. *NeuroImage* **184**, 293–316 (2019).

51. Golden, E. C., Graff-Radford, J., Jones, D. T. & Benarroch, E. E. Mediodorsal nucleus and its multiple cognitive functions. *Neurology* **87**, 2161–2168 (2016).

52. Ilyas, A., Pizarro, D., Romeo, A. K., Riley, K. O. & Pati, S. The centromedian nucleus: Anatomy, physiology, and clinical implications. *J. Clin. Neurosci.* **63**, 1–7 (2019).

53. de Bourbon-Teles, J. *et al*. Thalamic Control of Human Attention Driven by Memory and Learning. *Curr. Biol.* **24**, 993–999 (2014).

54. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **57**, 289–300 (1995).

55. Holmes, A. J. *et al.* Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. *Sci. Data* **2**, 150031 (2015).

56. Boes, A. D. *et al.* Network localization of neurological symptoms from focal brain lesions. *Brain* **138**, 3061–3075 (2015).

57. Subthalamic neuromodulation improves short-term motor learning in Parkinson's disease | Brain | Oxford Academic. https://academic.oup.com/brain/article/142/8/2198/5512025.

58. Horn, A. *et al.* Connectivity Predicts deep brain stimulation outcome in Parkinson disease. *Ann. Neurol.* **82**, 67–78 (2017).

59. Fischl, B. Automatically Parcellating the Human Cerebral Cortex. *Cereb. Cortex* **14**, 11–22 (2004).

60. Iglesias, J. E. *et al.* A probabilistic atlas of the human thalamic nuclei combining ex vivo MRI and histology. *NeuroImage* **183**, 314–326 (2018).

61. https://psychiatry.ucsd.edu/research/programs-centers/snl/_files/ants2.pdf. https://psychiatry.ucsd.edu/research/programs-centers/snl/_files/ants2.pdf.

62. Behrens, T. E., Berg, H. J., Jbabdi, S., Rushworth, M. F. & Woolrich, M. W. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *neuroimage* **34**, 144–155 (2007).

63. Desikan, R. S. *et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* **31**, 968–980 (2006).

64. Smith, S. M. & Nichols, T. E. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage* **44**, 83–98 (2009).

65. Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A. & Cohen, J. D. Humans use directed and random exploration to solve the explore–exploit dilemma. *J. Exp. Psychol. Gen.* **143**, 2074–2081 (2014).

66. Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S. & Wyart, V. Computational noise in reward-guided learning drives behavioral variability in volatile environments. *Nat. Neurosci.* **22**, 2066–2077 (2019).

67. Vehtari, A., Mononen, T., Tolvanen, V., Sivula, T. & Winther, O. Bayesian Leave-One-Out Cross-Validation Approximations for Gaussian Latent Variable Models. *J. Mach. Learn. Res.* **17**, 1–38 (2016).

68. Edlow, B. L. *et al.* 7 Tesla MRI of the ex vivo human brain at 100 micron resolution. *Sci. Data* **6**, 244 (2019).

69. Thomas Yeo, B. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).

70. Domenech, P., Rheims, S. & Koechlin, E. Neural mechanisms resolving exploitation-exploration dilemmas in the medial prefrontal cortex. *Science* **369**, eabb0184 (2020).

71. Hummos, A., Wang, B. A., Drammis, S., Halassa, M. M. & Pleger, B. Thalamic regulation of frontal interactions in human cognitive flexibility. *PLOS Comput. Biol.* **18**, e1010500 (2022).

72. Luria, A. R. TWO KINDS OF MOTOR PERSEVERATION IN MASSIVE INJURY OF THE FRONTAL LOBES. *Brain* **88**, 1–10 (1965).

73. Boldt, A., Blundell, C. & De Martino, B. Confidence modulates exploration and exploitation in value-based learning. *Neurosci. Conscious.* **2019**, niz004 (2019).

74. Kosciessa, J. Q., Lindenberger, U. & Garrett, D. D. Thalamocortical excitability modulation guides human perception under uncertainty. *Nat. Commun.* **12**, 2430 (2021).

75. Wang, B. A. & Pleger, B. Confidence in Decision-Making during Probabilistic Tactile Learning Related to Distinct Thalamo–Prefrontal Pathways. *Cereb. Cortex* **30**, 4677–4688 (2020).

76. Schwartenbeck, P. *et al.* Computational mechanisms of curiosity and goal-directed exploration. *eLife* **8**, e41703 (2019).

77. Smith, R. *et al.* Lower Levels of Directed Exploration and Reflective Thinking Are Associated With Greater Anxiety and Depression. *Front. Psychiatry* **12**, (2022).

78. Kosciessa, J. Q., Lindenberger, U. & Garrett, D. D. Thalamocortical excitability modulation guides human perception under uncertainty. *Nat. Commun.* **12**, 2430 (2021).

79. Balsdon, T. & Philiastides, M. G. Confidence control for efficient behaviour in dynamic environments. *Nat. Commun.* **15**, 9089 (2024).

80. Lam, N. H. *et al.* Prefrontal transthalamic uncertainty processing drives flexible switching. *Nature* **637**, 127–136 (2025).

81. Van der Werf, Y. D. *et al.* Deficits of memory, executive functioning and attention following infarction in the thalamus; a study of 22 cases with localised lesions. *Neuropsychologia* **41**, 1330–1344 (2003).

82. Aquino, T. G., Cockburn, J., Mamelak, A. N., Rutishauser, U. & O'Doherty, J. P. Neurons in human pre-supplementary motor area encode key computations for value-based choice. *Nat. Hum. Behav*. **7**, 970–985 (2023).

83. Kolling, N. *et al.* Value, search, persistence and model updating in anterior cingulate cortex. *Nat. Neurosci.* **19**, 1280–1285 (2016).

84. Shenhav, A., Cohen, J. D. & Botvinick, M. M. Dorsal anterior cingulate cortex and the value of control. *Nat. Neurosci.* **19**, 1286–1291 (2016).

85. Cole, M. W. & Schneider, W. The cognitive control network: Integrated cortical regions with dissociable functions. *NeuroImage* **37**, 343–360 (2007).

86. Segar, D. J. *et al.* Lesion location and lesion creation affect outcomes after focused ultrasound thalamotomy. *Brain* **144**, 3089–3100 (2021).

87. Aubignat, M. & Godefroy, O. Cognitive Effects of MR-Guided Focused Ultrasound Thalamotomy—Time to Evaluation. *JAMA Neurol.* **82**, 309 (2025).

88. Kaplitt, M. G. *et al.* Safety and Efficacy of Staged, Bilateral Focused Ultrasound Thalamotomy in Essential Tremor: An Open-Label Clinical Trial. *JAMA Neurol.* **81**, 939–946 (2024).

89. Phillips, J. M. *et al.* Primate thalamic nuclei select abstract rules and shape prefrontal dynamics. 2024.03.13.584871 Preprint at https://doi.org/10.1101/2024.03.13.584871 (2025).

90. Zhang, X., Mukherjee, A., Halassa, M. M. & Chen, Z. S. Mediodorsal thalamus regulates task uncertainty to enable cognitive flexibility. *Nat. Commun.* **16**, 2640 (2025).