Supplementary Information: Methods

The Structural Connectivity of Subthalamic Deep Brain Stimulation Correlates with Impulsivity in Parkinson's

Philip E. Mosley [1,2,3,4], Saee Paliwal [5], Katherine Robinson [1], Terry Coyne [3,6], Peter Silburn [2,3], Marc Tittgemeyer [7], Klaas E. Stephan [5,7,8], Alistair Perry [1,9,10] †, Michael Breakspear [1,11] †

[1] *Systems Neuroscience Group, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia*
[2] *Neurosciences Queensland, St Andrew's War Memorial Hospital, Spring Hill, Queensland, Australia*
[3] *Queensland Brain Institute, University of Queensland, St Lucia, Queensland, Australia*
[4] *Faculty of Medicine, University of Queensland, Herston, Queensland, Australia*
[5] *Translational Neuromodeling Unit (TNU), Institute for Biomedical Engineering, University of Zürich and Swiss Federal Institute of Technology (ETH Zürich), Zürich, Switzerland*
[6] *Brizbrain and Spine, the Wesley Hospital, Auchenflower, Queensland, Australia*
[7] *Max Planck Institute for Metabolism Research, Cologne, Germany.*
[8] *Wellcome Centre for Human Neuroimaging, University College London, London, UK*
[9] *Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Germany*
[10] *Centre for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany*
[11] *Brain and Mind Priority Research Centre, Hunter Medical Research Institute, University of Newcastle, NSW, Australia*

†Co-senior author

Correspondence to:
Dr Philip E Mosley, Neurosciences Queensland, Level 1, St Andrew's Place, 33 North Street, Spring Hill, Queensland, 4000, Australia

Correspondence to:
Dr Philip E Mosley, Neurosciences Queensland, Level 1, St Andrew's Place, 33 North Street, Spring Hill, Queensland, 4000, Australia

E-mail:
philip.mosley@qimrberghofer.edu.au

Telephone:
+61 (7) 3839 3688

1.1 Neuropsychiatric Assessment Instruments

1.1.1 Barratt Impulsiveness Scale

The BIS is one of the most widely used instruments for assessing trait impulsivity and is often considered the gold standard instrument against which other measures are compared. It is a 30-item self-report questionnaire assessing the prevalence of impulsive behaviours. Respondents must rate each item (e.g. '*I act on the spur of the moment*') from 1 to 4 according to the frequency of occurrence (i.e. rarely / never; occasionally; often; always / almost always). Higher scores indicate greater impulsivity. The mean BIS score is significantly greater in persons with PD with ICDs compared to non-impulsive persons with PD (Isaias *et al.,* 2008; Voon *et al.,* 2007), suggesting that this instrument also has construct validity in the assessment of impulsivity in PD.

1.1.2 Questionnaire for Impulsive-Compulsive disorders in PD

The QUIP-RS is a 28-item self-report questionnaire assessing the prevalence of ICDs including compulsive spending, hypersexuality, pathological gambling, binge eating, hobbyism, punding and dopamine dysregulation (Weintraub *et al.,* 2012). Respondents much rate each item (e.g. '*Do you have urges or desires for the following behaviours that you feel are excessive or cause you distress?*') from 0 (never) to 4 (very often). The total sum obtained for each compulsive behaviour indicates the current severity of that behaviour. The instrument was designed for use in the PD population and was previously validated against a semi-structured clinical interview.

1.1.3 Hayling Test

The Hayling Test is a sentence completion task, during which participants must insert a nonsense word at the end of a sentence, inhibiting the pre-potent stimulus to complete the sentence with a word that makes sense. The test assesses the construct of inhibition and is sensitive to frontal lobe dysfunction (Burgess *et al.,* 1997). For example, in the sentence: '*the whole town came to hear the Mayor*…' a correct response could be '*banana*'. Participants would be penalised for completing the sentence with the clearly related words '*speak*', or '*talk*' (referred to as category A errors), as well as with words that are only partially connected such as '*explode*' (referred to as category B errors). Persons with PD make more category A and B errors than controls on this task (O'Callaghan *et al.,* 2013a; Obeso *et al.,* 2011).

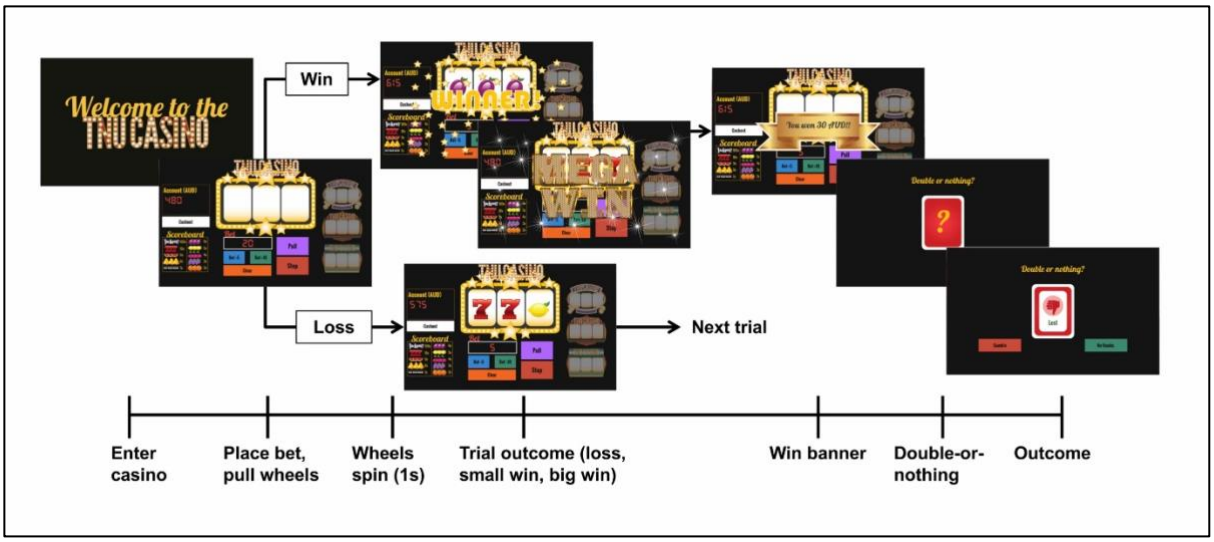
1.1.4 Excluded Letter Fluency Task

The ELF is an additional measure of inhibitory control (Shores *et al.,* 2006). Participants are given 3 trials of 90 seconds to produce as many words as possible that do not contain a specified vowel. Words

must be longer than 3 letters and cannot be proper nouns or derivations of the same word stem. Scoring includes an overall correct total, the number of rule violations and the number of word repetitions. In a sample of 50 persons with PD, the number of rule violations was previously shown to be significantly greater compared to age-matched controls and was highly correlated with anatomical changes in brain regions implicated in inhibition (O'Callaghan *et al.*, 2013b).

1.1.5 Delay Discounting Task

An assessment of delay aversion, the tendency to prefer sooner, smaller rewards over those that are larger but temporally more distant (Kirby *et al.*, 1999). The task was designed to assess impulsivity in individuals with substance use disorders; behaviours that share face validity with the impulse-control disorders (ICDs) observed in a subset of persons with PD. Participants are presented with a series of 27 choices between an amount of money distributed immediately and a larger sum after a specified delay. After the task is complete, participants have the opportunity to win the amount of money they have chosen in a choice selected at random, either immediately or after a delay, depending on the choice they have made. Subsequently, the pattern of choices is analysed to calculate a discount parameter, or the point of indifference between delayed and immediate rewards for a given sum. Individuals with greater delay aversion have a higher discount parameter. The extent of delay aversion was previously shown to be greater amongst persons with PD than healthy controls (Milenkova *et al.*, 2011), as well as being greater amongst persons with PD with ICDs than non-impulsive persons with PD (Housden *et al.*, 2010; Voon *et al.*, 2011).



Supplementary Figure 1 | Slot Machine Gambling Paradigm

The task consists of 100 trials. On every trial, players are able to place a bet of unlimited magnitude, switch slot machines or *'cash out*', exiting the casino and returning again on another virtual '*day*'. The overall win probability is 25 %, with wins split into big wins and small wins. The two possible types of losses are near-misses, in which the first two wheels are the same and the third is different (i.e. AAB) or a true loss, in which all the wheels are different (i.e. ABC). Game play proceeds as follows. Each trial begins with the slot machine main screen loading, displaying the player's account value. The player then places a continuous-valued bet amount, incremented in units of 5 or 10 AUD. After the player has placed a bet, he or she presses the '*Pull*' button and watches as the wheels begin to spin. At any point, the player has the ability to press the '*Stop*' button, ending the trial and subsequently revealing the outcome of the three wheels. Unbeknownst to the participant, pressing the stop button has no effect on the trial outcome. If the stop button is not pressed, the trial times out after 5 seconds, and the player sees the outcome of the first, second and third wheel sequentially. On trials in which the outcome is a win, there are ten possible reward grades (or multiples of the bet amount). After every win trial, players are offered a possible '*double-up*' option, during which players are given 3 seconds to decide whether or not to engage in a '*double-or-nothing*' option, thereby risking his or her entire win amount. If the player elects to engage in this gamble, a card flips over revealing the result, and subjects are taken to the next trial. If the player does nothing, or decides not to gamble, he or she is taken to the next trial. For each loss trial, players are taken directly to the beginning of the next trial. Again, the trajectory of win-loss outcomes is fixed, ensuring comparable inference upon perceptual and response parameters across participants.

1.2 The Virtual Casino

We employed a modified version of an established slot machine gambling paradigm validated in healthy controls (Paliwal *et al.*, 2014). The task was designed to have standard features normally attributed to slot machines (colours, sounds, banners), and its features were designed to mirror the specifications of Swiss and German slot machines (Figure 2). Due to its game-like feel, the task successfully elicits impulsive, risk-taking and exploratory behaviour in participants. Task behaviour has been previously shown to correlate with standard measures of impulsivity (i.e. the BIS). Players began the slot machine with 2000 AUD in their account, and played through 100 trials. The trajectory of win-loss outcomes was predetermined, ensuring that participants' experience of rewards and losses were comparable in order and quantity. The trajectory results in a positive outcome (net winnings) for most participants. At the end of the task, participants were awarded up to 30 AUD in real money based on the size of these virtual winnings.

The task began with five training trials, after which the participant played through the main task, consisting of 100 trials. Only data from the main task were used for further analysis. For the main task, the win probability was 25 %, with wins split into big wins (12 % of trials) and small wins (88 % of trials). Players won when all three wheels showed the same symbols (e.g. all three wheels display an apple image). There were two possible types of losses. The first was a near-miss, in which the first two wheels of the slot machine displayed the same symbol, and the third was different (e.g. cherry, cherry, apple). The second was a true loss, in which all the wheels displayed different images (e.g. cherry, apple, orange).

Game play proceeded as follows: at the onset of each trial, the main screen loaded, displaying the player's account value. Players were then able to execute one of the following actions: place a bet (of unlimited magnitude – by loading the machine in increments of 5-10 AUD), switch slot machines, or '*cash out'*, which involved '*exiting*' the casino and returning again on another virtual '*day*'. While these actions might at first glance appear to relate to different behaviours, they all share a common theme in that they enhance outcome variance and thus risk (compare the definition of risk in behavioural economics). For example, for a machine switch, regardless of whether the player is performing well or poorly on the current machine, the decision to switch machines incurs the risk that the new machine chosen may be punishing or rewarding, thereby making the player vulnerable to the variance of the task. Similarly, a bet

increase is a risk-inducing shift in the face of uncertainty, again making the player more susceptible to larger wins and losses. And in the same respect, casino switches and double-ups again expose players to the risk that their environment will change dramatically, and for the worst. Each of the above actions thus leads to greater outcome variance (risk), and risk-taking, in turn, is one critical component of impulsivity (Whiteside and Lynam, 2001).

If the player chose to bet, after loading the machine, they pressed the '*Pull*' button and watched as the wheels begin to spin. The player had the option of pressing the '*Stop*' button at any time during wheel spin, ending the trial and subsequently revealing the outcome of the three wheels, Pressing the stop button had no effect on the trial outcome; though this was not told to the participant. In the absence of the stop button being pressed, the trial timed out after five seconds, and revealed the outcome to the player, with the first, second and third wheel stopping sequentially. For winning trials, there were ten possible reward amounts. Each possible reward was called a reward grade, and indicated a different multiple of the bet size placed (e.g. reward grade 1 indicated a reward amount that was double the bet amount placed). After each win trial, the player is offered a '*double-up*' option, during which they were given three seconds to decide whether or not to engage in a '*double-or-nothing*' option. The double-or-nothing option had two possible outcomes: if the player won the double-or-nothing gamble, they doubled their win amount from that trial, if they lost the double-or-nothing option, they lost their entire win amount from the corresponding trial. If the player did nothing, or decided not to gamble, they are taken to the next trial. For losses, players were taken directly to the beginning of the next trial.

In the context of the analyses presented, this version of the slot machine deviated from the version presented in Paliwal *et al.* (2014) in several important ways: here, players were given the ability to place unlimited bet sizes with the ability to increase their bets in increments of 5 or 10 AUD; there were no '*fake win*' results, simply wins and losses; the task was considerably (fifty percent) shorter; and finally, the task was aesthetically remodelled in order to be more naturalistic.

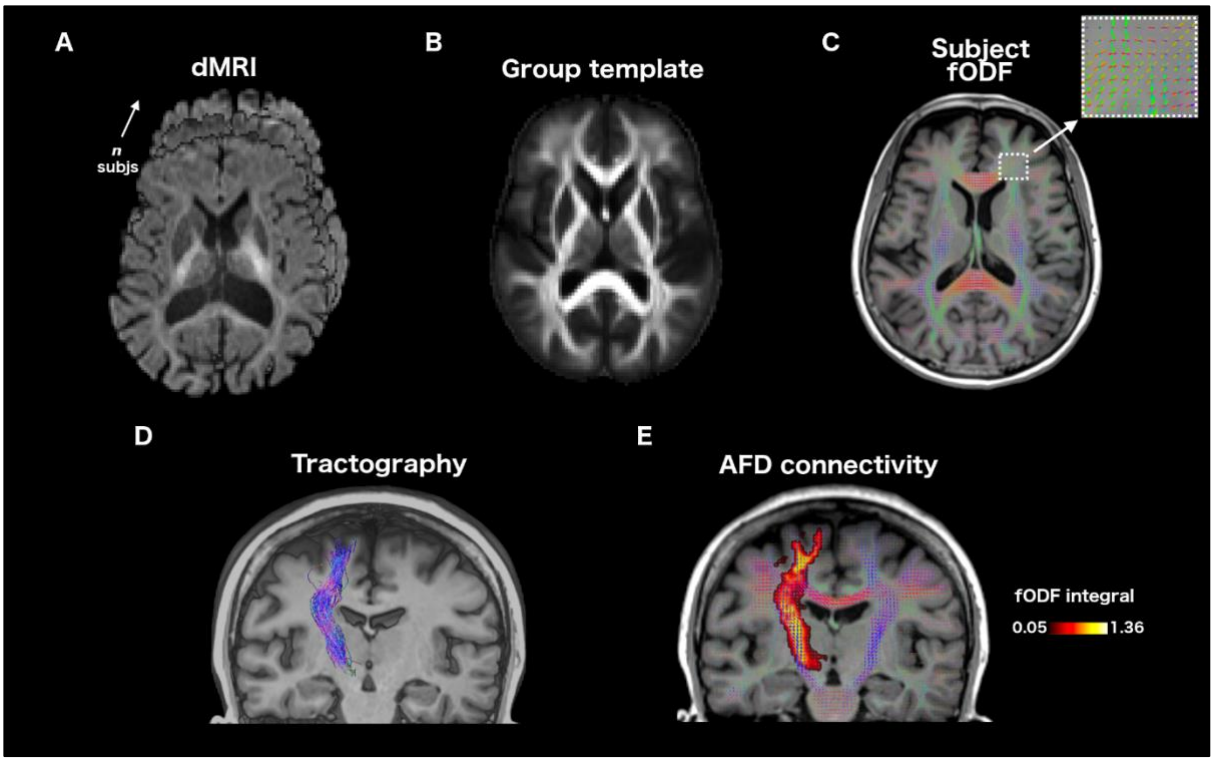
Each trial in the game followed a pre-programmed result sequence consisting of the following trial types:

i. big wins: top three out of 10 reward grades

ii. small wins: lower 7 out of 10 reward grades
iii. near misses: when the same symbol appeared in wheel 1 and two, and a different symbol appeared in wheel 3
iv. true losses (all three wheels showed different symbols).

After all win trials, a participant was allowed to engage in a secondary double-or-nothing gamble, and had three seconds to decide whether or not to do so. A more detailed trial breakdown across the various types of wins and losses is listed below:

i. 25 out of 100 trials were wins
ii. 3 out of 100 trials were big wins
iii. 22 out of 100 trials were small wins
iv. 22 out of 100 trials were near misses
v. 52 out of 100 trials were full losses.
vi. All win trials contained a double-up option



Supplementary Figure 2 | Diffusion Processing Pipeline

High angular-resolution diffusion-weighted imaging was acquired along 90 directions using a 3T scanner and a 64-channel array head coil, with a *b*-value of 3000s/mm^2 and voxel size of 1.7 mm^3 isotropic. After denoising, and correction for motion, susceptibility, bias and eddy-current induced distortions, fractional anisotropy (FA) maps were calculated for each participant. B: FA maps were non-linearly registered to a population-average FA template, in order to derive an average white matter mask, which was then warped back into individual space to permit intensity normalisation on the diffusion data. This ensured that the median *b*0 white matter value was uniform across the study population. C: From the intensity-normalised diffusion data, signal responses across different tissue types (grey-matter, white-matter, CSF) were estimated and averaged across all participants to obtain a group-wise response function. Constrained spherical deconvolution of the average white-matter signal furnished fibre orientation distribution functions (fODF) for each participant. These functions provide local estimates of the density of fibres according to their angular orientation and can resolve complex organisations of crossing fibres more effectively than single tensor models. Our acquisition protocol incorporating 90 directions was designed to optimise this process. D: Fibre bundles were reconstructed using a probabilistic streamline algorithm, through sampling a probability density of the fODF at each path point, tracking the most plausible fibre propagations between seed and target regions. E: Quantitative estimates of structural connectivity between seed and target regions were derived from the apparent fibre density (AFD), calculated by summing the fODF lobe integrals along the pathway of interest and dividing by mean streamline length, to estimate the mean cross-sectional area of the fibre bundle.

1.3 Diffusion Image Acquisition

Diffusion MRI data were acquired using a 3T Siemens PRISMA scanner, with a 64-channel array head coil. For each scan, the diffusion-weighting was distributed in an isotropic manner (1.7 mm_3 voxel resolution) along 90 directions with a *b*-value of 3000s/mm_2 (phase-encoding in anterior-posterior direction), using a twice-refocused spin-echo planar imaging (EPI) sequence. Twelve non-diffusion-weighted images (*b*0) were acquired and interleaved throughout this sequence. An additional sequence of 8 non-diffusion weighted images were acquired in the opposite phase-encoding direction (posterior-anterior).

1.4 Diffusion Image Pre-Processing

dMRI data were pre-processed using functions provided within MRtrix3 (www.mrtrix3.org) (https://github.com/MRtrix3/mrtrix3), called from a pre-processing pipeline developed in-house (https://github.com/breakspear/diffusion-pipeline). The dMRI data were first denoised (Veraart *et al.*, 2016) and then corrected for motion, susceptibility, and eddy-current induced distortions within FSL *eddy* (version 5.0.11) (Andersson and Sotiropoulos, 2016), leveraging the reverse phase-encoding acquisition to estimate the inhomogeneity fields. FSL eddy (using --repol) was also used to detect slice signal outliers due to bulk motion, corrected using a non-parametric replacement method (Andersson *et al.*, 2016). Finally, bias-intensity correction was performed (Zhang *et al.*, 2001).

1.5 Intensity Normalization and Fibre Reconstruction

To permit the comparison of fibre reconstruction and structural connectivity estimates across participants, group-average intensity normalization was performed. Briefly, fractional anisotropy (FA) maps were first calculated from the bias-corrected diffusion images of each participant (Fig 1A), and non-linearly registered to a generated population-average FA template (Fig 1B). The population template was subsequently used to derive an average white-matter mask (FA > 0.4), which was warped back into individual space, and then intensity normalization was performed on the dMRI data. This ensured that the median *b*0 white-matter value was uniform across study participants (Raffelt *et al.*, 2012).

From the intensity-normalized diffusion data, the signal responses across different tissue-types (grey-matter, white-matter, CSF) were estimated (Dhollander *et al.*, 2016; Dhollander *et al.*, 2018) and then averaged across all participants to obtain a group-wise response function. Through constrained spherical deconvolution (CSD) (*lmax* = 8, *msmt_csd*) (Jeurissen *et al.*, 2014) of the average white-matter signal contribution, fibre orientation distribution functions (fODF) were estimated for each participant (Fig 1C). The fODF provides local estimates of the apparent density of fibres as a function of angular orientation (Tournier *et al.*, 2004; Tournier *et al.*, 2007). CSD is able to resolve complex local fibre orientations and amplitudes at high-precision (Raffelt *et al.*, 2015) and the acquisition protocol in this investigation was designed to maximise this capability.

1.6 Tractography

The probabilistic streamline algorithm iFOD2 (Tournier *et al.*, 2010) was used to reconstruct fibre-bundles. Through sampling a probability density function of the fODF at each path point, the iFOD2 algorithm tracked the most plausible fibre propagations between seed and target regions, until 100 streamlines in total were reconstructed (Fig 1D). The default (for the acquired voxel resolution) tracking parameters were as follows: step size = 0.86 mm, minimum length = 8.56 mm, max length = 250 mm, seed/termination fODF threshold = 0.05, curvature constraint = 1 mm radius.

All cortical and basal ganglia regions used in seed-based tractography were in MNI ICBM nonlinear asymmetric space. Parcellations were first transformed into individual anatomical space through non-linear co-registration of the skull-stripped T1 image and the ICBM template (Avants *et al.*, 2008). Resultant parcellations (now in anatomical space) were then co-registered into diffusion space, applying a transformation matrix derived from boundary-based registration (Greve and Fischl, 2009) of the anatomical and mean *b*0 diffusion images.

1.7 Apparent Fibre Density

Estimates of structural connectivity between each seed and target region was derived from the *apparent fibre density* (AFD) (Raffelt *et al.*, 2012), calculated by summing the fODF lobe integrals along each pathway of interest, approximating the total fibre volume. To correct for differences in fibre length, the total fibre volume was divided by the mean streamline length to

estimate the cross-sectional area of the fibre bundle, providing a measure of fibre 'density' independent of fibre length.

AFD affords a more biologically-interpretable quantification of structural connectivity along a given bundle in comparison to traditional tensor-derived metrics such as fractional anisotropy (FA) and mean diffusivity (MD) (Calamante *et al.*, 2015; Raffelt *et al.*, 2012). Tensor-based metrics provide an average value across all voxels traversed by the pathway of interest and their biological interpretation in neurological disorders is contested, especially in tissue that contain crossing fibres (Raffelt *et al.*, 2015; Riffert *et al.*, 2014; Scheck *et al.*, 2015).

1.8 Networks

Cortical targets for the reward evaluation and response inhibition networks were selected from a gold-standard subdivision of the cortex based on multimodal MRI data (Glasser *et al.*, 2016), which were initially projected onto volumetric MNI ICBM nonlinear asymmetric 2009a space (Horn, 2016). These included areas 10r and 10v (vmPFC), OFC and pOFC (OFC), a24 and p24 (ACC), 45 and 47l (IFG), 6ma and 6mp (SMA). The basal ganglia parcellations (within 2009b space) that served as seeds within these tractography networks included the VS (Choi *et al.*, 2012), the VTA (Pauli *et al.*, 2018), and the STN (Ewert *et al.*, 2018). All cortical and basal ganglia parcels were non-linearly transformed into native diffusion space via the skull-stripped anatomical image.

1.9 PLS Path Modelling

Partial least squares path modelling (PLS-PM) was employed to represent the multivariate relationships between connectivity and behavioural measures (McIntosh and Lobaugh, 2004; Shaw *et al.*, 2016), controlling for relevant demographic and disease-related factors. PLS-PM is a form of structural equation modelling in which complex associations between multivariate data sets can be estimated. Each model specifies the linear weighting of one set of variables that best co-varies with a linear weighting of another. For example, in this investigation, connectivity variables were created from the reward evaluation and response inhibition networks as a weighted mixture of the connectivity of each tract within the network. Behavioural variables were formed from each neuropsychiatric instrument and each gambling output (although as these were assessed individually, the relationship between each behavioural variable and observed behaviour was monotonic). Each model then represents the path coefficients and yields the corresponding significance values for the relationship between these connectivity and behavioural variables, in addition to describing the weighted contribution that each tract makes to the connectivity variable. In each model, continuous measures including age, years since diagnosis of Parkinson's disease and LEDD were also entered as co-variates, with disease subtype and gender examined with a permutation test. Interaction (or moderating) effects of these co-variates on the effect of connectivity on behaviour were also modelled. Confidence intervals for estimates of the path coefficients were determined by bootstrapping, in which the data set was repeatedly sampled with replacement to create 10,000 independent bootstrapped data sets, with the sample size equal to the number of participants. Each PLS path model was developed using a bi-hemispheric structural network, but results for each hemisphere in isolation are also reported.

For each outcome of interest, a number of alternative PLS path models of varying complexity could be proposed. However, there is no consensus method for determining the optimal trade-off between model fit and model complexity (Henseler and Sarstedt, 2013). Therefore, model complexity was constrained *a priori* to be equivalent across all models; each PLS path model included only one structural network and all included age, years since diagnosis and LEDD as covariates. One interaction term with the structural network was included (e.g. the interaction of LEDD or age with the reward evaluation network). The winning model from all permutations was selected based on the maximum R_2 value prior to bootstrapping: in the setting

of equivalent complexity of all estimated models, we thus use model fit (R^2) as the single summary metric for comparing models.

Formal mathematical treatment of the PLS-PM model is given by Tenenhaus *et al.* (2005). In brief, each PLS path model is formed by two submodels: the structural (or inner model) and the measurement (or outer model). The structural model assesses the relationships between latent variables (in this case connectivity and behaviour), which are unobserved (latent) representations that cannot be directly measured but may underlie associations between measurable variables. The measurement model describes how each latent variable relates to a block of manifest (observed) variables (such as how the connectivity variable is comprised of a mixture of tracts from each frontostriatal network that can be quantified).

Linear relationships within the structural (inner) model are expressed as:

$$LV_j = \beta_0 + \sum_{i \to j} \beta_{ji} LV_i + error_j$$

where the *j*-th latent variable LV_j, formed by the block X_j of variables, is predicted by a weighted sum of the remaining latent variables LV_i, with weight coefficients β_{ji} (path coefficients) representing the strength and direction of the relations between LV_j and predictors LV_i.

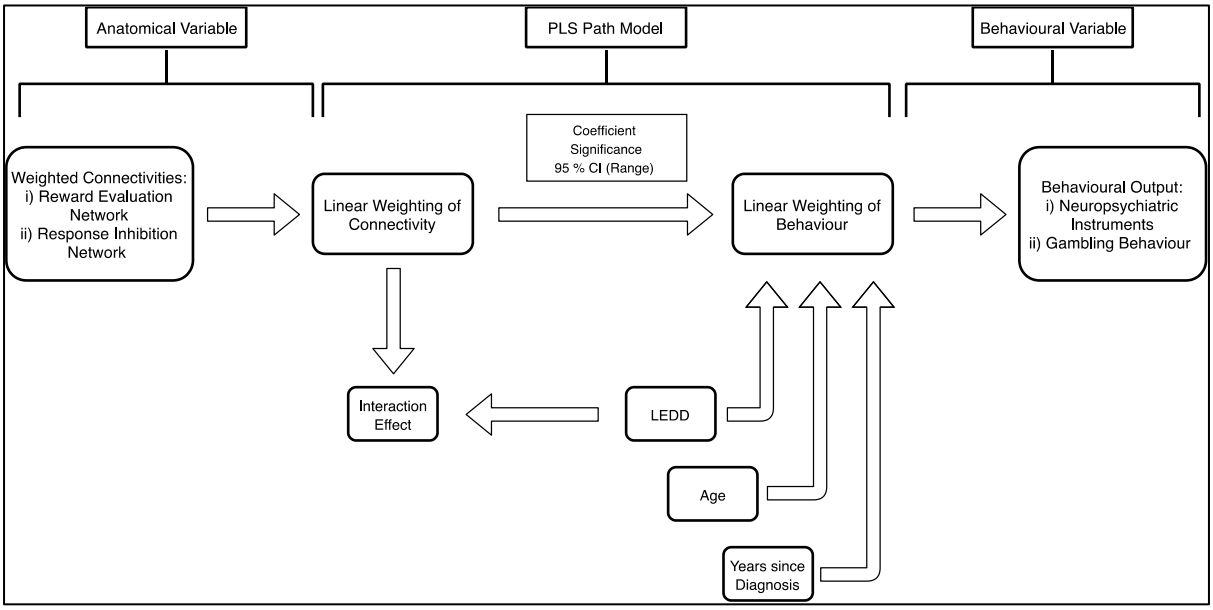
In the 'formative' measurement model employed here to define the connectivity latent variables, manifest (observed) variables (individual tracts within each network) are considered to be 'forming' the latent variable, which is calculated as a weighted sum of its manifest variables:

$$LV_j = \sum_k \omega_{jk} X_{jk}$$

with each variable k within each block X_j of variables denoted by X_{jk} and weights denoted by ω_{jk}. For alternative types of measurement models in PLS-PM, please see Tenenhaus *et al.* (2005).

The PLS algorithm is an iterative procedure in which abitary weights are first assigned to the outer model, in order to approximate the inner latent variables as linear combinations of their manifest variables. Then, relationships between latent variables in the inner model are considered to obtain approximations of the inner model. Subsequently, new outer weights are calculated and the inner model respecified until the model converges on the outer weights. At this point path coefficients are estimated using ordinary least squares regression between latent variables.

In summary, the goal of PLS-PM is to analyse a system of linear relationships between multiple blocks of variables. Latent (unobserved) variables are calculated as linear combinations of their associated formative (manifest) variables. Linear combinations are optimised so that the final obtained latent variables express the relationships of the structural and measurement models in a way that maximises the explained variance.



Supplementary Figure 3 | PLS Path Modelling.

A PLS path model represents the relationship between structural network connectivity and impulsivity. A connectivity variable is constructed from the apparent fibre density of each white matter tract in the structural network under investigation. The individual contribution of each tract to the connectivity variable is quantified by a 'weight' and the connectivity variable is formed as a linear mixture of the corresponding apparent fibre density values that best co-varies with the behavioural variable under investigation. The relationship between the connectivity and behavioural latent variables is quantified in the path model by a path coefficient (that can be tested for statistical significance). Relevant demographic and disease-related co-variates are also represented in the inner model and path coefficients can be determined for these relationships. An interaction (moderating) effect can be modelled; in this case, the interaction of LEDD with the connectivity variable. Bootstrapping of the model yields 95 % confidence intervals for the path coefficients of interest.

1.10 Latent Change Score Modelling

A latent change score model was applied to test for longitudinal cross-domain (connectivity-behaviour) coupling. In these models, the score of an individual i on a construct of interest $Behaviour$ at time t can be used to calculate a change score between $t1$ and $t2$:

$$\Delta Behaviour_i = Behaviour_{i,t2} - Behaviour_{i,t1}$$

The latent change score factor $\Delta Behaviour_i$ is modelled as a latent variable. It captures the change between timepoints and is proportional to the scores at $t1$ through an autoregressive parameter β:

$$\Delta Behaviour_i = \beta \cdot Behaviour_{i,t1}$$

In this manuscript, both $Connectivity$ and $Behaviour$ variables are including in a bivariate extension of the latent change score model. Through cross-domain coupling, the extent of change in one domain $\Delta Behaviour_i$ can be seen as a function of the baseline level in the other $Connectivity_{i,t1}$. This can be quantified as follows:

$$\Delta Behaviour_i = \beta \cdot Behaviour_{i,t1} + \gamma \cdot Connectivity_{i,t1}$$

In the context of this investigation, this relationship captures the extent to which changes in impulsivity are a function of the initial condition of brain measures (i.e. structural network connectivity at baseline). For further mathematical treatment of these relationships, please see Kievit *et al.* (2018).

1.11 Linear Mixed-Effects Modelling

Linear mixed-effects models complemented the latent change score models by providing an explicit change term in interpretable units, such as dollars wagered or percentage of 'double or nothing' gambles accepted. These took the form:

$$\text{Gambling variable}_{ij} \sim \mathbf{Connectivity}_{ij} + \text{LEDD}_i + \text{Age}_i + \text{Gender}_i + \text{Years Since Diagnosis}_i + \text{Timepoint}_{ij} + (1|\text{ID}) + (1|\text{Timepoint}_{ij})$$

with i denoting participant and j denoting timepoint.

Models were adjusted for levodopa equivalent dose, age, gender, years since diagnosis of Parkinsons' disease and timepoint (pre- or post-DBS), with random intercepts specified for each participant and each timepoint.

The term in bold (the effect of connectivity on gambling variable) is the coefficient of interest. Hypothesis testing on a null model (omitting connectivity) was performed with the *anova* function in the *lavaan* package.

REFERENCES

Andersson, J.L.R., Graham, M.S., Zsoldos, E., Sotiropoulos, S.N., 2016. Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. Neuroimage. 141**,** 556-572.

Andersson, J.L.R., Sotiropoulos, S.N., 2016. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. Neuroimage. 125**,** 1063-1078.

Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C., 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal. 12**,** 26-41.

Burgess, P.W., Shallice, T., Thames Valley Test Company., 1997. The Hayling and Brixton tests, Vol., Thames Valley Test Company, Bury St Edmunds.

Calamante, F., Smith, R.E., Tournier, J.D., Raffelt, D., Connelly, A., 2015. Quantification of voxel-wise total fibre density: Investigating the problems associated with track-count mapping. Neuroimage. 117**,** 284-293.

Choi, E.Y., Yeo, B.T., Buckner, R.L., 2012. The organization of the human striatum estimated by intrinsic functional connectivity. J Neurophysiol. 108**,** 2242-63.

Dhollander, T., Raffelt, D., Connelly, A., 2016. Unsupervised 3-tissue response function estimation from single-shell or multi-shell diffusion MR data without a co-registered T1 image. ISMRM Workshop on Breaking the Barriers of Diffusion MRI. Lisbon, Portugal.

Dhollander, T., Raffelt, D., Connelly, A., 2018. Accuracy of response function estimation algorithms for 3-tissue spherical deconvolution of diverse quality diffusion MRI data. 26th International Society of Magnetic Resonance in Medicine. Paris, France.

Ewert, S., Plettig, P., Li, N., Chakravarty, M.M., Collins, D.L., Herrington, T.M., *et al.*, 2018. Toward defining deep brain stimulation targets in MNI space: A subcortical atlas based on multimodal MRI, histology and structural connectivity. Neuroimage. 170**,** 271-282.

Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., *et al.*, 2016. A multi-modal parcellation of human cerebral cortex. Nature. 536**,** 171-178.

Greve, D.N., Fischl, B., 2009. Accurate and robust brain image alignment using boundary-based registration. Neuroimage. 48**,** 63-72.

Henseler, J., Sarstedt, M., 2013. Goodness-of-fit indices for partial least squares path modeling. Computational Statistics. 28**,** 565-580.

Horn, A., 2016, HCP-MMP1.0 projected on MNI2009a GM (volumetric) in NIfTI format. https://neurovault.org/images/24150/. 3rd August 2018

Housden, C.R., O'Sullivan, S.S., Joyce, E.M., Lees, A.J., Roiser, J.P., 2010. Intact reward learning but elevated delay discounting in Parkinson's disease patients with impulsive-compulsive spectrum behaviors. Neuropsychopharmacology. 35**,** 2155-64.

Isaias, I.U., Siri, C., Cilia, R., De Gaspari, D., Pezzoli, G., Antonini, A., 2008. The relationship between impulsivity and impulse control disorders in Parkinson's disease. Mov Disord. 23**,** 411-5.

Jeurissen, B., Tournier, J.D., Dhollander, T., Connelly, A., Sijbers, J., 2014. Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. Neuroimage. 103**,** 411-426.

Kievit, R.A., Brandmaier, A.M., Ziegler, G., van Harmelen, A.L., de Mooij, S.M.M., Moutoussis, M., *et al.*, 2018. Developmental cognitive neuroscience using latent change score models: A tutorial and applications. Dev Cogn Neurosci. 33**,** 99-117.

Kirby, K.N., Petry, N.M., Bickel, W.K., 1999. Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. J Exp Psychol Gen. 128**,** 78-87.

McIntosh, A.R., Lobaugh, N.J., 2004. Partial least squares analysis of neuroimaging data: applications and advances. Neuroimage. 23 Suppl 1**,** S250-63.

Milenkova, M., Mohammadi, B., Kollewe, K., Schrader, C., Fellbrich, A., Wittfoth, M., *et al.*, 2011. Intertemporal choice in Parkinson's disease. Mov Disord. 26**,** 2004-10.

O'Callaghan, C., Naismith, S.L., Hodges, J.R., Lewis, S.J., Hornberger, M., 2013a. Fronto-striatal atrophy correlates of inhibitory dysfunction in Parkinson's disease versus behavioural variant frontotemporal dementia. Cortex. 49**,** 1833-43.

O'Callaghan, C., Naismith, S.L., Shine, J.M., Bertoux, M., Lewis, S.J., Hornberger, M., 2013b. A novel bedside task to tap inhibitory dysfunction and fronto-striatal atrophy in Parkinson's disease. Parkinsonism Relat Disord. 19**,** 827-30.

Obeso, I., Wilkinson, L., Casabona, E., Bringas, M.L., Alvarez, M., Alvarez, L., *et al.*, 2011. Deficits in inhibitory control and conflict resolution on cognitive and motor tasks in Parkinson's disease. Exp Brain Res. 212**,** 371-84.

Paliwal, S., Petzschner, F.H., Schmitz, A.K., Tittgemeyer, M., Stephan, K.E., 2014. A model-based analysis of impulsivity using a slot-machine gambling paradigm. Front Hum Neurosci. 8**,** 428.

Pauli, W.M., Nili, A.N., Tyszka, J.M., 2018. A high-resolution probabilistic in vivo atlas of human subcortical brain nuclei. Sci Data. 5**,** 180063.

Raffelt, D., Tournier, J.D., Rose, S., Ridgway, G.R., Henderson, R., Crozier, S., *et al.*, 2012. Apparent Fibre Density: a novel measure for the analysis of diffusion-weighted magnetic resonance images. Neuroimage. 59**,** 3976-94.

Raffelt, D.A., Smith, R.E., Ridgway, G.R., Tournier, J.D., Vaughan, D.N., Rose, S., *et al.*, 2015. Connectivity-based fixel enhancement: Whole-brain statistical analysis of diffusion MRI measures in the presence of crossing fibres. Neuroimage. 117**,** 40-55.

Riffert, T.W., Schreiber, J., Anwander, A., Knosche, T.R., 2014. Beyond fractional anisotropy: extraction of bundle-specific structural metrics from crossing fiber models. Neuroimage. 100**,** 176-91.

Scheck, S.M., Pannek, K., Raffelt, D.A., Fiori, S., Boyd, R.N., Rose, S.E., 2015. Structural connectivity of the anterior cingulate in children with unilateral cerebral palsy due to white matter lesions. Neuroimage Clin. 9**,** 498-505.

Shaw, P., Weingart, D., Bonner, T., Watson, B., Park, M.T., Sharp, W., *et al.*, 2016. Defining the neuroanatomic basis of motor coordination in children and its relationship with symptoms of attention-deficit/hyperactivity disorder. Psychol Med. 46**,** 2363-73.

Shores, E.A., Carstairs, J.R., Crawford, J.R., 2006. Excluded Letter Fluency Test (ELF): Norms and Test–Retest Reliability Data for Healthy Young Adults. Brain Impairment. 7**,** 26-32.

Tenenhaus, M., Vinzi, V.E., Chatelin, Y.-M., Lauro, C., 2005. PLS path modeling. Computational Statistics & Data Analysis. 48**,** 159-205.

Tournier, J.D., Calamante, F., Gadian, D.G., Connelly, A., 2004. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. Neuroimage. 23**,** 1176-85.

Tournier, J.D., Calamante, F., Connelly, A., 2007. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. Neuroimage. 35**,** 1459-72.

Tournier, J.D., Calamante, F., Connelly, A., 2010. Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions. Proceedings of the International Society for Magnetic Resonance in Medicine. Stockholm.

Veraart, J., Novikov, D.S., Christiaens, D., Ades-Aron, B., Sijbers, J., Fieremans, E., 2016. Denoising of diffusion MRI using random matrix theory. Neuroimage. 142**,** 394-406.

Voon, V., Thomsen, T., Miyasaki, J.M., de Souza, M., Shafro, A., Fox, S.H., *et al.*, 2007. Factors associated with dopaminergic drug-related pathological gambling in Parkinson disease. Arch Neurol. 64**,** 212-6.

Voon, V., Sohr, M., Lang, A.E., Potenza, M.N., Siderowf, A.D., Whetteckey, J., *et al.*, 2011. Impulse control disorders in Parkinson disease: a multicenter case--control study. Ann Neurol. 69**,** 986-96.

Weintraub, D., Mamikonyan, E., Papay, K., Shea, J.A., Xie, S.X., Siderowf, A., 2012. Questionnaire for impulsive-compulsive disorders in Parkinson's Disease–Rating Scale. Movement Disorders. 27**,** 242-247.

Whiteside, S.P., Lynam, D.R., 2001. The Five Factor Model and impulsivity: using a structural model of personality to understand impulsivity. Personality and Individual Differences. 30**,** 669-689.

Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans Med Imaging. 20**,** 45-57.