

# Distinct computational mechanisms underlying cognitive flexibility deficits in impulsivity and compulsivity

Vasilisa Skvortsova (✉ [vasilisaskv@gmail.com](mailto:vasilisaskv@gmail.com))

University College London <https://orcid.org/0000-0002-5582-5670>

Tobias Hauser

Max Planck UCL Centre for Computational Psychiatry and Ageing Research <https://orcid.org/0000-0002-7997-8137>

---

## Article

**Keywords:** impulsivity, compulsivity, learning noise, learning imprecision, choice stochasticity, adaptive learning, crowd-sourced smartphone experiments

**Posted Date:** January 27th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1280535/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

## **Distinct computational mechanisms underlying cognitive flexibility deficits in impulsivity and compulsivity**

Vasilisa Skvortsova<sup>1,2\*</sup>, Tobias U Hauser<sup>1,2</sup>

<sup>1</sup> Max Planck UCL Centre for Computational Psychiatry and Ageing Research, London  
WC1B 5EH, United Kingdom.

<sup>2</sup> Wellcome Centre for Human Neuroimaging, University College London, London WC1N 3BG,  
United Kingdom.

\* Corresponding author: Vasilisa Skvortsova (vasilisaskv@gmail.com).

## **Abstract**

Cognitive flexibility, the ability to quickly adapt to changing environmental demands, is a hallmark of human behaviour, and is impaired across multiple psychiatric disorders. Especially compulsivity and impulsivity disorders have been linked to impaired adaptive learning and flexibility. Initial computational investigations suggested these distinct psychiatric dimensions suffer from the same underlying neurocognitive impairments, related to stochasticity during choice. However, a recent advance in computational neuroscience has demonstrated that imprecision in the learning process itself can account for a large portion of behavioural variability traditionally attributed to choice-stochasticity. Here, in a series of large-scale experiments using both lab-designed and gamified citizen-science tasks, we show that distinct computational markers are affected in compulsivity and impulsivity. Whilst impulsivity is tied to an imprecision in learning across valence domains, (hygiene-related) compulsivity is linked to choice-stochasticity. This double-dissociation demonstrates that distinct neurocomputational mechanisms can drive seemingly similar behavioural deficits, only dissociable using targeted computational approaches.

**Key words:** impulsivity, compulsivity, learning noise, learning imprecision, choice stochasticity, adaptive learning, crowd-sourced smartphone experiments

## Introduction

The ability to learn and adjust to changes in action-outcome contingencies lies at the core of flexible behaviour and is critical for survival in dynamic environments. Imbalances in cognitive flexibility are a hallmark of “impulsive-compulsive” disorders such as obsessive-compulsive disorder (OCD), attention-deficit/hyperactivity disorder (ADHD), and substance abuse<sup>1-5</sup>. Phenomenologically, compulsive behaviours are believed to arise from excessive rigidity and cognitive inflexibility, which leads to maladaptive compulsive perseveration and reduced goal-directed behaviour<sup>6-9</sup>. In contrast, premature and exaggerated switching is archetypical for impulsivity<sup>1,10,11</sup>.

Despite the apparent phenomenological opponency between compulsivity and impulsivity, behavioural and computational analyses of cognitive flexibility thus far have found rather similar impairments showing deficits in the choice rather than the learning phase of adaptive (reversal) learning paradigms. Compulsivity was often linked to increased choice switching<sup>12-18</sup>, similarly to impulsivity, which has been linked to increased switching often seen as exaggerated exploration<sup>19-22</sup>.

However, recent advances in the neurocomputational understanding of perceptual and value-guided decision-making have put previous modelling attempts into question<sup>23,24</sup> by identifying a new source of behavioural variability that has previously not been accounted for, and which stems from the imprecisions in the inference process itself<sup>23,24</sup>. In reward-guided learning<sup>24</sup>, this inference imprecision translates into imprecisions in the learning process itself, i.e., random variability in feedback processing. These learning-driven imprecisions account for over 2/3 of behavioural variability that has previously been attributed solely to choice variability<sup>24</sup>. They also

demand to re-consider several behavioural phenomena, such as choice hysteresis (repetition of previous choices; as proposed in compulsivity) and the presence of what was typically assumed to be exploratory decisions (as proposed in impulsivity)<sup>24,25</sup>.

In this series of large-scale computational studies, we thus examined whether and how distinct computational mechanisms can account for the cognitive flexibility impairments seen in impulsivity and compulsivity. In the first dimensional online study, we used this recently developed task and model<sup>24</sup> to demonstrate that impulsivity is linked to an exaggerated learning imprecision but not choice deficits, whilst compulsivity (primarily washing compulsions) is linked to imbalances in choice stochasticity but not learning. In a second experiment, we demonstrate that this double-dissociation holds true for novel forms of data collection using crowd-sourced, citizen-science smartphone apps. Our findings thus not only provide a re-interpretation of the cognitive flexibility deficit mechanisms, but also reveal that seemingly similar impairments can have distinct computational origins, which provide us with critical new information about their underlying neural mechanisms<sup>15,20,26–29</sup>.

## Results

### *Learning imprecisions account for choice deficits in adaptive learning*

To assess cognitive flexibility in compulsivity and impulsivity, we collected a large non-clinical sample (N=392; age  $32.6 \pm 12.4$  (mean $\pm$ s.d.), range 18-72, 204F) via the Prolific online workers platform. Subjects played a restless 2-arm bandit reward learning task online, similar to the one we previously used to examine learning noise imprecisions<sup>24</sup> (Fig.1A-B, Methods, STable1).

Overall, subjects accumulated more rewards than chance, demonstrating they adapted their behaviour based on the received information ( $t(391)=36.29$ ,  $p<0.0001$ ), and their performance was comparable to the performance obtained in the previous laboratory study (N=30)<sup>24</sup> (Fig.1D) confirming that careful online experimentation can elicit similar data as lab-based studies<sup>30</sup>.

Recent work has highlighted the importance of accounting for learning errors when modelling behavioural variability in reward-guided decisions under uncertainty<sup>24,25</sup>. To probe whether this also held true for our online sample, we ran a model comparison and found that indeed a 'noisy' model with learning imprecision better explained subjects' behaviour than a traditional 'exact' reinforcement learning (RL) model without any learning noise (Fig.1E; posterior model exceedance probability  $P_{exc}>0.99$ ).

Importantly, this winning model allowed for two sources of variability: at the moment of choice (through a softmax inverse temperature parameter) and at the moment of action value update through learning imprecision (a learning noise parameter that scaled with the absolute magnitude of prediction error on each trial). This 'noisy' model also fit subjects' choices better than a model with just learning imprecisions (i.e., noisy 'argmax' model) (Fig.1E), suggesting that both sources of variability are needed to best explain behaviour.

Finally, we examined to which extent the learning noise explained variability previously attributed to choice stochasticity<sup>31</sup>. We found that learning noise explained more than  $73.9 \pm 1.5\%$  (mean  $\pm$  s.e.m) of total behavioural variability in our data (Fig.1E), in line with previous findings<sup>24</sup>. This suggests that previous findings linked to choice stochasticity as a primary source of inter-individual differences may be inadequate and could be explained through other mechanisms.

### ***Impulsivity and Compulsivity are linked to choice stochasticity in the exact model***

Following previous studies<sup>32–36</sup>, we characterized compulsivity and impulsivity dimensionally using questionnaires (OCI-R<sup>37–40</sup> to measure compulsivity; BIS-11<sup>41,42</sup> to measure impulsivity; SFig.3).

First, we wanted to assess whether both dimensions were linked to an imbalanced choice process in the standard, exact (i.e., noise-free) learning.<sup>18,43–45</sup> In line with previous findings, we found that both compulsivity  $R_{OCI-R \text{ total}} = 0.16, p = 0.002$  and impulsivity  $R_{BIS \text{ total}} = 0.13, p = 0.01$  were positively associated with the choice stochasticity (inverse softmax temperature parameter) parameter<sup>17–19,36</sup>, suggesting that these personality traits might increase choice-driven variability.

To make sure that these associations were not driven by other covariates, we additionally controlled for age, gender and other computational model parameters. We replicated the above effects (Fig.2a;  $\beta_{BIS \text{ total}} = 0.11 (SE = 0.04), p = 0.02$ ;  $\beta_{OCI-R \text{ total}} = 0.13 (SE = 0.04), p = 0.002$ ), which means that these covariates did not explain the observed association, and we kept them in all subsequent analyses.

We did not observe any relationship between the learning rate for the chosen option and either of the two psychiatric traits (both  $p$ -values  $> 0.17$ ), suggesting that inter-individual differences in compulsivity and impulsivity did not depend on the exact learning process.

### ***Distinct associations of impulsivity and compulsivity when accounting for learning imprecisions***

We next investigated the model parameters of the better fitting ‘noisy’ model, which additionally captures learning imprecisions. Interestingly, when learning noise was taken into account, the choice stochasticity was no longer linked to impulsivity (Fig.2B;  $\beta_{BIS\ total} = 0.02$  ( $SE = 0.04$ ),  $p = 0.51$ ), but it remained associated with compulsivity ( $\beta_{OCI-R\ total} = 0.11$  ( $SE = 0.04$ ),  $p = 0.002$ ). Instead, we found a significant association of learning noise parameter with impulsivity (Fig.2C;  $\beta_{BIS\ total} = 0.08$  ( $SE = 0.04$ ),  $p = 0.02$ ), but not compulsivity ( $\beta_{OCI-R\ total} = 0.03$  ( $SE = 0.04$ ),  $p = 0.49$ ).

We, again, observed no relationship between the learning rates and either psychiatric trait ( $\beta_{BIS\ total} = 0.03$  ( $SE = 0.05$ ),  $p = 0.61$ ,  $\beta_{OCIR\ total} = -0.05$  ( $SE = 0.05$ ),  $p = 0.35$ ) suggesting impulsivity is not associated with poorer or exaggerated learning per se, but rather, it is linked to a more inconsistent learning over time<sup>22,46</sup>. We found a similar double dissociation in another task setting, in which participants received feedback for both (chosen and unchosen) stimuli (Supplementary Information).

Finally, in line with previous study<sup>24</sup>, accounting for learning imprecisions resulted in re-categorizing of more than 70% ( $74.6 \pm 1.5\%$ ) (mean  $\pm$  s.e.m) of the decisions labelled as ‘non-greedy’ (i.e., choices where an option with lower expected value has been chosen) under the exact model, as ‘greedy’ under the noisy model. This re-evaluation resulted in a dissociation



between compulsive and impulsive traits previously linked to the non-greedy decisions<sup>7,33,36</sup> with only compulsivity positively contributing to the exploratory decisions (cf Supplementary Information).

### ***Learning imprecisions are sufficient to account for choice hysteresis***

A tendency to repeat previous choices irrespective of their value has been traditionally captured with an additional parameter during choice<sup>47–49</sup>. This “choice hysteresis” (tendency to repeat same action) parameter has previously been linked to compulsivity<sup>7,14,15,17</sup>. However, we found that choice hysteresis was no longer necessary when accounting for learning imprecisions, aligned with previous simulations<sup>24</sup> (Supplementary Information). While in the exact setting, adding choice hysteresis parameter improved model fitting ( $P_{\text{exc with bias}} > 0.97$ ), the same parameter in the noisy model was not necessary:  $P_{\text{exc no bias}} > 0.99$  (SFig.1C, Supplementary Information).

In the exact model, the choice hysteresis parameter was significantly different from 0:  $1.06 \pm 0.05$ ,  $t(391) = 21.8$ ,  $p < 0.0001$  and was negatively associated with OCI-R total score in line with previous studies<sup>7,14,15,17</sup>:  $\beta_{\text{OCIR total}} = -0.17$  ( $SE = 0.05$ ),  $p = 0.0007$  (controlling for the covariates:  $\beta_{\text{OCIR total}} = -0.13$  ( $SE = 0.05$ ),  $p = 0.003$ ). We examined whether a similar association was present in the a (worse fitting) noisy model with choice hysteresis and did no longer observe any association between compulsivity and choice hysteresis ( $\beta_{\text{OCIR total}} = -0.07$  ( $SE = 0.04$ ),  $p = 0.12$ ). This suggests that choice hysteresis is superfluous in this task when accounting for learning noise.

### ***Associations with specific compulsivity and impulsivity components***

Impulsivity and compulsivity as captured here are multifactorial constructs that aggregate across what is believed to be separable components of compulsivity and impulsivity<sup>1,10,50,51</sup>. Previous factor analyses of BIS and OCI-R revealed multiple subscales, which characterize different facets of these endophenotypes<sup>37-39</sup> and contribute to different behavioural deficits<sup>1,8,52,53</sup>.

We thus assessed which components of compulsivity and impulsivity were most strongly linked to the model parameters observed above. As these factors are correlated with each other (R-values > 0.3, SFig.6C-D), we conducted separate regression analyses and then corrected for multiple comparisons.

When investigating choice stochasticity, we found that this parameter was primarily linked to the compulsivity subscales of “washing” (Fig.3B,  $\beta_{OCI-R \text{ washing}} = 0.12$  ( $SE = 0.04$ ),  $p = 0.001$ ,  $\alpha_{corrected}=0.004$ ) and “hoarding” ( $\beta_{OCI-R \text{ hoarding}} = 0.10$  ( $SE = 0.03$ ),  $p = 0.003$ ,  $\alpha_{corrected}=0.004$ ) (Fig.3A, STable2). We did not find any relationship between choice stochasticity and any of the impulsivity subscales (all p-values > 0.05, Fig.3A, STable2) suggesting that choice imbalances are primarily associated with behavioural compulsions related to washing and hoarding.

Next, we investigated learning noise as the parameter primarily linked to impulsivity. We found that learning noise was specifically associated with the *motor impulsiveness* subscale (Fig.3B;  $\beta_{BIS \text{ motor impulsiveness}} = 0.17$  ( $SE = 0.03$ ),  $p < 0.0001$ ,  $\alpha_{corrected}=0.004$ ), capturing the propensity to act prematurely without foresight<sup>37,39</sup> (all other subscales p-values > 0.05, STable2).

### ***Cognitive flexibility in crowd-sourced smartphone app data***

Online experiments, as used above, have been highly successful for studying inter-individual differences in psychiatric traits and transdiagnostic approaches<sup>34,54,55</sup>. Nevertheless, using online worker platforms has several limitations<sup>30,56</sup>. First, the participant pool is limited to registered workers whose primary interest is most often financial. Moreover, the platforms are restricted in age (adults only), language (mostly English speakers) and most often confined to a specific geographic location (here: United Kingdom). More importantly, using laboratory tasks online does not answer the question whether obtained results could generalize across different contexts and have sufficient ecological validity for a potential translation into clinical applications<sup>57,58</sup>. Thus, several research groups turned towards studying cognition through gamified tasks and more ecologically valid settings using crowd-sourced game platforms, such as smartphone app games<sup>30,59–62</sup>.

The goal of this second experiment was thus not only to replicate our findings in a more ecologically realistic sample, but also examine whether similar results could be obtained using a substantially shorter task in a less well controlled environment which is likely to be more reflective of any potential future clinical translation<sup>57,58</sup>.

We designed a short, gamified version of the 2-arm bandit task from experiment 1 (Fig.1C) and used it as part of the smartphone-based Brain Explorer research app ([www.brainexplorer.net](http://www.brainexplorer.net)). We collected an independent large convenience sample of app users (N=2610, age  $41.8y \pm 15.9$  (s.d.), range (18–85y), 1499F; STable1) who played the game voluntarily and without reimbursement (comparison of samples cf SFig.3-5).

Subject played a “Milky Way” game with the goal to cumulate as much space milk as possible from two space cows (i.e., bandits with drifting reward magnitudes, Fig.1C). The performance

between the two experiments was comparable, but slightly lower in the smartphone sample (Fig.1D; exp.1:  $8.12 \pm 0.22$ , exp.2:  $7.07 \pm 0.09$  (mean $\pm$ s.e.m),  $t(535.51)=4.31$ ,  $p=0.0002$ ). This difference is expected as the latter task is substantially shorter, has a much briefer training and no pre-task quiz (as used in experiment 1; also see<sup>30</sup>). However, the second sample still performed significantly above chance ( $t(2609)=76.05$ ,  $p<0.0001$ ), showing that they understood the task and adapted based on the feedback.

As in experiment 1, the model that included both choice stochasticity and the learning noise term outperformed the models with either only the choice stochasticity (exact learning) or only the learning noise:  $P_{exc}>0.999$  (Fig.1F) with the best fitted parameter distributions laying in similar ranges but having slightly higher values as compared to experiment 1 (Fig.S2, Supplementary Information). The fraction of behavioural variability that was accounted for by the learning noise was even larger in the smartphone sample, accounting for  $78.7 \pm 0.06\%$  (mean $\pm$ s.e.m) (exp.1 vs exp.2:  $t(499.47)=2.60$ ,  $p=0.01$ ) (Fig.1F). This highlights that learning noise is of particular importance when assessing learning using smartphone apps like the one used here.

### ***Associations with impulsivity and compulsivity***

All users also completed the BIS and OCI-R questionnaires (STable1), with scores comparable between the two experiments (SFig.3-5; BIS total exp.1:  $60.7 \pm 11.1$ , exp.2:  $62.3 \pm 10.0$ ; OCI-R total exp.1:  $18.7 \pm 13.8$ , exp.2:  $15.2 \pm 10.1$  (mean $\pm$ s.d.)).

First, we replicated that choice stochasticity was significantly associated with impulsivity (Fig.2D;  $\beta_{BIS\ total} = 0.07$  ( $SE = 0.01$ ),  $p < 0.0001$ ) and compulsivity ( $\beta_{OCI-R\ total} = 0.04$  ( $SE = 0.02$ ),  $p = 0.008$ ) in the worse-fitting, exact, model that does not account for learning noise. This

confirms our hypothesis that imbalanced cognitive flexibility in compulsivity and impulsivity would be expressed as a bias in choice selection when we do not account for imprecisions during learning. When investigating the best-fitting model, which accounted for both sources of variability, we again found that choice stochasticity was no longer associated with impulsivity (Fig.2E;  $\beta_{BIS\ total} = 0.006$  ( $SE = 0.01$ ),  $p = 0.63$ ) but impulsivity was now significantly associated with learning noise instead (Fig.2F;  $\beta_{BIS\ total} = 0.10$  ( $SE = 0.01$ ),  $p < 0.00001$ ). These findings confirm our previous observation that general impulsivity was exclusively linked to learning imprecisions when accounting for this mechanism.

Interestingly, unlike in the first experiment, general compulsivity was no longer linked to choice stochasticity (Fig.2E;  $\beta_{OCI-R\ total} = -0.009$  ( $SE = 0.01$ ),  $p = 0.55$ ). In line with experiment 1, we also did not find any association between compulsivity and the learning noise parameter (Fig.2F;  $\beta_{OCIR\ total} = 0.01$  ( $SE = 0.008$ ),  $p = 0.22$ ).

### ***Compulsivity and impulsivity subscales reliably linked to behaviour***

In the first experiment, we identified that specific subscales were more closely associated with biases in both choice stochasticity and learning imprecisions. As these showed stronger associations with the parameters, we specifically tested whether we could find the same associations in the smartphone sample. We thus investigated whether motor impulsiveness was linked to learning noise, and whether hoarding and/or washing compulsions were linked to choice stochasticity.

We indeed confirmed the specific association between motor impulsiveness and learning noise (Fig.4B; STable3;  $\beta_{BIS\ motor\ impulsiveness} = 0.10$  ( $SE = 0.01$ ),  $p < 0.0001$ ) with no association between any impulsivity sub-scores and choice stochasticity, (Fig.4; STable3).

Next, even though we did not replicate the association between choice stochasticity and overall compulsivity, we were able to replicate the association between choice stochasticity and washing compulsivity (Fig. 4A;  $\beta_{OCI-R\ washing} = 0.03$  ( $SE = 0.01$ ),  $p = 0.02$ ) but not hoarding (Fig.4; STable3). Our findings thus suggest that the specific link between choice stochasticity and compulsive washing was strong enough to remain significant in this somewhat noisier smartphone-based data set.

### ***(Motor) impulsiveness drives learning imprecision across valence domains***

Lastly, we tested whether the learning imprecision effects observed in impulsivity would generalize across valence domains, i.e., whether they would also be present in learning to avoid punishment. This is particularly relevant for impulsivity as previous studies linked impulsivity to altered cognitive flexibility and learning from punishment<sup>63-66</sup>.

We thus implemented a version of the same smartphone game in the loss domain. Instead of collecting points users were instructed to learn how to avoid losing the points already won (Fig.5A). Overall, 670 ( $42.8y \pm 17.3$ , range 18-83, 415F) (STable1) users completed both reward and punishment (reward game always played first) games, together with the questionnaires and showed comparable performances across both games (Fig.5B, Supplementary Information).

We verified that the learning noise model also better accounted for this task ( $P_{exc}=0.99$ ) (Supplementary Information) with comparable parameters across both games (SFig.7A-D).

As in the reward domain, we replicated the positive association between general impulsivity (Fig.5C-D, Stable4, Supplementary Information). We further found that motor impulsivity was specifically linked to learning noise, suggesting that these associations are present across both valence domains.

## Discussion

Reduced cognitive flexibility and adaptive learning are hallmarks of many impulsive and compulsive disorders<sup>4,5,8,67,68</sup>. Yet, their computational mechanisms remain poorly understood and early modelling attempts suggested similar impairments despite apparent phenomenological differences. Here, we show that by considering new sources of learning variability<sup>23–25</sup>, we could deconvolve the impulsivity- and compulsivity-related reductions in cognitive flexibility and show that they are driven by different computational mechanisms. We demonstrate across multiple large-scale studies that while compulsivity is associated with increased choice stochasticity, impulsivity is related to the imprecisions in learning process itself. Whilst previous computational modelling approaches have linked both impulsivity and compulsivity to deficits during the choice process<sup>15,17,18,45,69</sup>, we highlight that this is not the case when considering multiple sources of variability. We also replicated previous findings<sup>24</sup> that showed that accounting for learning imprecisions not only substantially improved model performance and but their lack resulted in inflated estimates of choice-driven variability. In fact, we show that contrary to previous findings, impulsivity is not linked to imprecisions during choice (e.g., exploration<sup>19,22,36,45</sup>), but rather to increased variability during learning with likely different underlying neurobiological mechanisms.

Noradrenaline is believed to play a key role in the learning variability<sup>70–72</sup>. Previous work found that learning imprecisions were linked to indirect markers<sup>24</sup> of noradrenaline functioning<sup>73–75</sup> such as activity in the dorsal anterior cingulate cortex and pupil size fluctuations. This suggestion also aligns well with the assumption that many impulsivity disorders, such as ADHD, are linked to impaired noradrenaline functioning<sup>19,26,76,77</sup>. However, our study cannot rule out that an increased exploration during choice – as proposed in previous studies<sup>19,22,36,45</sup> – is still present in impulsivity. Studies that investigate exploration in the absence of any learning<sup>20,36,71</sup> suggest that increased exploration might still be linked to impulsivity and reflect noradrenaline functioning<sup>71</sup>. Interestingly, preliminary pharmacological studies suggest that noradrenaline is linked to increased learning imprecisions rather than choice stochasticity.<sup>25,78</sup>

Impulsivity represents a heterogenous symptom dimension which shows substantial variability<sup>10,50,68,79</sup>. Here, we observed a particularly strong association between motor impulsiveness (i.e., acting without thinking and the inability to withhold a response) and learning noise. Previous work showed that motor impulsiveness contributed to disinhibition possibly via noradrenaline system<sup>52,80</sup>, suggesting a potentially more specific link.

In contrast to impulsivity, our computational analysis of compulsivity – more specifically washing compulsivity – confirmed a reduced precision during choice, even when accounting for the learning noise. Importantly, the introduction of learning imprecisions also allowed us to reconsider other computational phenomena that have previously been attributed to compulsivity such as choice hysteresis (or ‘stickiness’)<sup>47,81,82</sup>. Whilst choice hysteresis improved model fitting in the noise-free model and correlated positively with compulsivity, in the better-

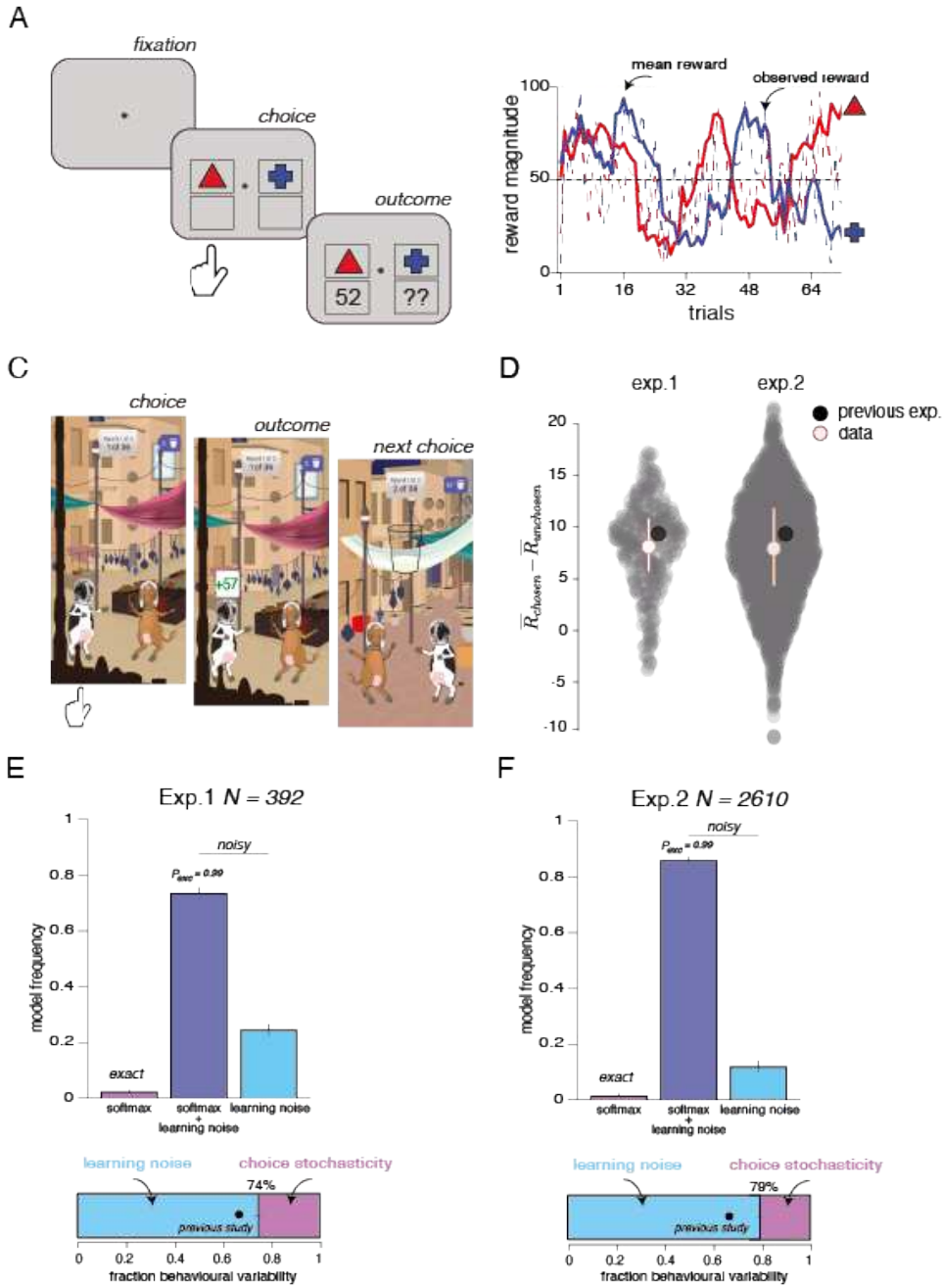


fitting noisy learning model, this parameter no longer improved model fits thus inviting to reconsider the previous findings<sup>14-17</sup>.

Our study was also a testbed for new technologies and the potential of translation and generalisation of such computational tasks. In the first study, we used a well-established paid worker platform, whilst in the second we analysed an unpaid, heterogenous sample of smartphone users which played a short, gamified version of the same experiment. Importantly, we replicated most of the key associations in both samples, which is critical for the findings' robustness and the potential for using such tasks in clinics<sup>57,58</sup>. However, it should also be noted that the association between choice stochasticity and compulsivity was more variable and only robust for washing compulsions. We believe that this is due to a generally increased noisiness in the shorter smartphone-based task, which may render the choice stochasticity parameter less sensitive. Such effects on signal quality are important to consider when building cognitive probes and could arise from differences in the samples (semi-professional vs lay participants), in the study length, instruction, or design. Crowd-source smartphone-based experiments provide an unprecedented access to the population that are difficult to reach in the laboratory settings (e.g., children and elderly) but one should be mindful of the larger variability and elevated measurement noise leading to lower effect sizes<sup>30,58</sup>.

In this study, we demonstrate that by taking into account recent developments in understanding multiple sources of behavioural variability, we can successfully dissociate differences in cognitive flexibility linked to compulsivity and impulsivity. By testing and replicating these effects across different sample and contexts, we demonstrate how advanced computational modelling can help pinpoint distinct neurocognitive mechanisms underlying seemingly similar deficits.

Figure 1



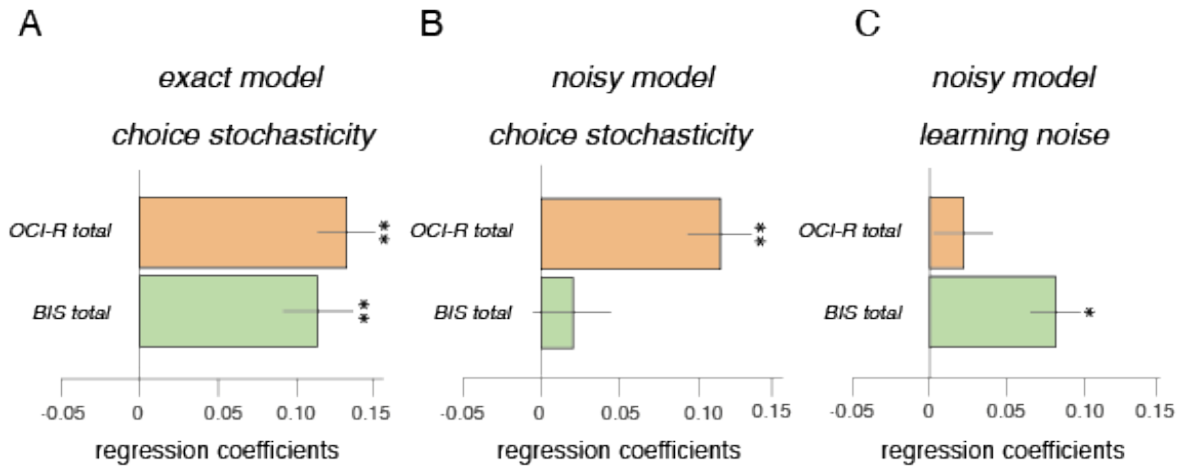
**Figure 1. Task structure and contributions of learning noise to behavioural variability**

**A.** Trial structure and design in two-armed bandit task in experiment 1. On every trial, participants were making a choice between two bandits represented by coloured shapes on the screen and observed the outcome (between 1 or 99 points) for the chosen option. **B.** Example of reward magnitudes for two bandits for one session in experiment 1. Reward magnitudes changed through trials and were sampled from two probability distributions with

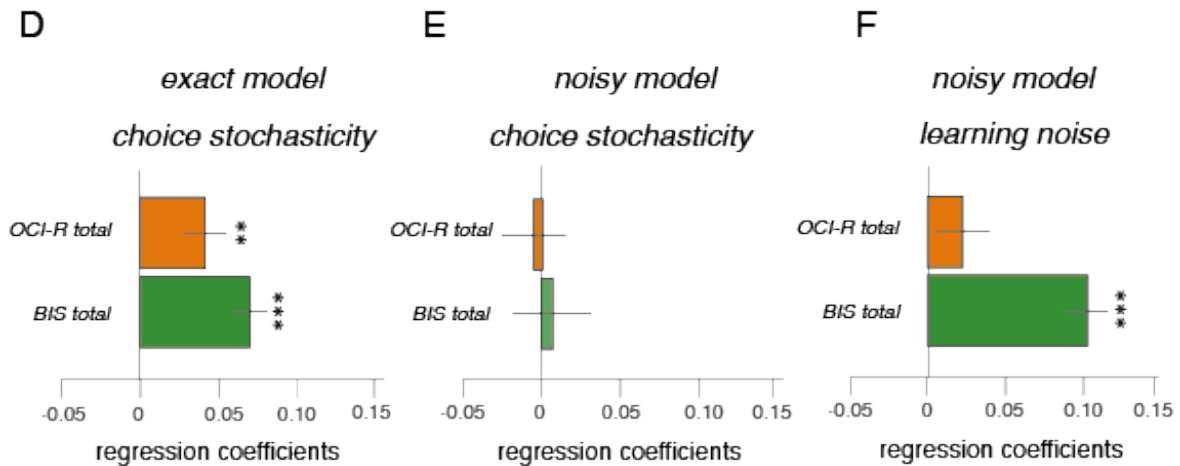
independently drifting means. Thick lines represent the drifting means of two distributions; dotted lines represent the reward outcomes that could have been observed by the subject conditioned on the choice. This observable reward magnitudes are sampled from the distribution mean with added Gaussian sampling noise to ensure continuous learning in the task. **C.** Task structure for reward learning two-armed bandit task in experiment 2 (Milky Way game in the Brain Explorer app). On every trial, participants were making a choice between two space cows and then observed how much space milk they have collected in the form of points (from 1 to 99). The underlying distributions of drifting mean rewards were the same as in experiment 1 but the task was substantially shorter. **D.** To analyse the subjects' overall performance in both experiments, we computed the difference between the total average reward won by each subject and the foregone reward (to account for the individually generated reward trajectories). A positive difference means that subjects were performing better than at chance level. In both experiment 1 and 2, subjects performed better than chance ( $p < 0.001$ ) White dots are medians, error bars are 25<sup>th</sup> and 75<sup>th</sup> percentile of the performance distributions. Grey dots represent individual subjects. Black dots are the median performance from the previous laboratory study<sup>24</sup>. **E. Top panels.** Bayesian model selection (BMS) results for experiment 1 (left panel) and experiment 2 (right panel). Bars represent estimated model frequencies for exact model (just softmax decision rule, left bar), noisy learning model with softmax decision rule (middle bar), and noisy learning model with argmax decision rule (right bar). Error bars are s.d. for model frequencies based on Dirichlet distribution. Model with two sources of variability (learning noise and softmax decision rule) better accounted for data in both experiments similarly to the previously observed results<sup>24</sup>. **Bottom panels.** Fraction of behavioural variability attributed to learning noise based on the best winning model in experiment 1 (left panel) and experiment 2 (right panel). In both experiments more than 2/3 of total variability in the decision was due to learning noise rather than choice stochasticity. Black dots represent the results obtained in the previous laboratory study<sup>24</sup>.

Figure 2

Exp.1 N = 392



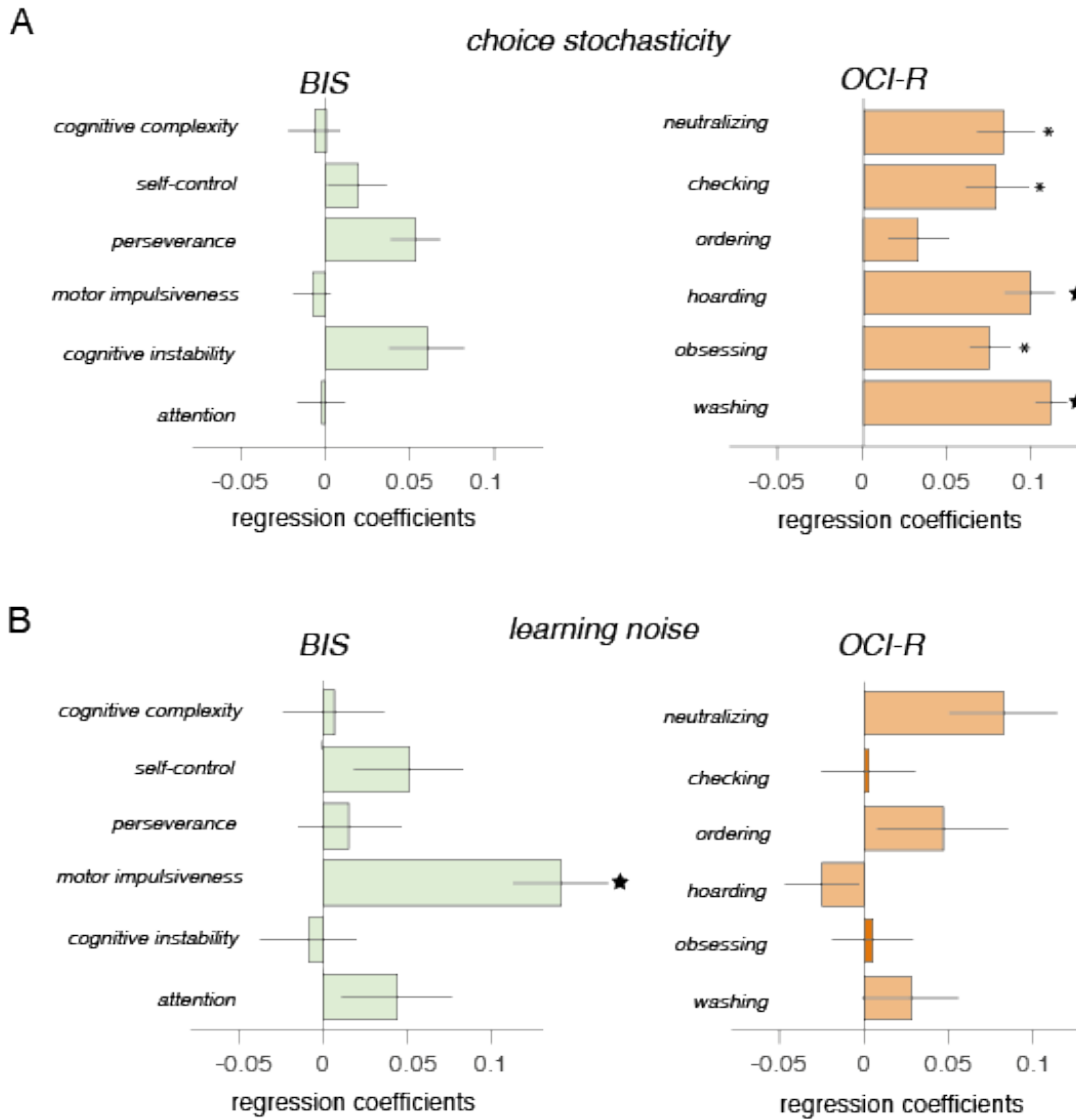
Exp.2 N = 2610



**Figure 2. Contributions of impulsivity and obsessive-compulsive traits to choice stochasticity and learning noise.**

**A.** In the exact model (no learning noise added), both impulsivity traits measured with BIS total score and OC traits measured with OCI-R total score showed a significant association with choice stochasticity (softmax parameter). **B.** Impulsivity traits were no longer associated with choice stochasticity when learning noise is added to the model. **C.** Impulsivity traits contributed to learning noise rather than choice stochasticity in the learning noise model. **D-F.** Same analysis as in A-C for the Brain Explorer (experiment 2) data. All regressions included age, gender, IQ (experiment 1), mental health status (experiment 2), and other model parameters as covariates. Dark bar colours indicate the replication tests based on the results from experiment 1. Error bars are standard errors, \* -  $p < 0.05$ , \*\* -  $p < 0.01$ , \*\*\* -  $p < 0.001$ .

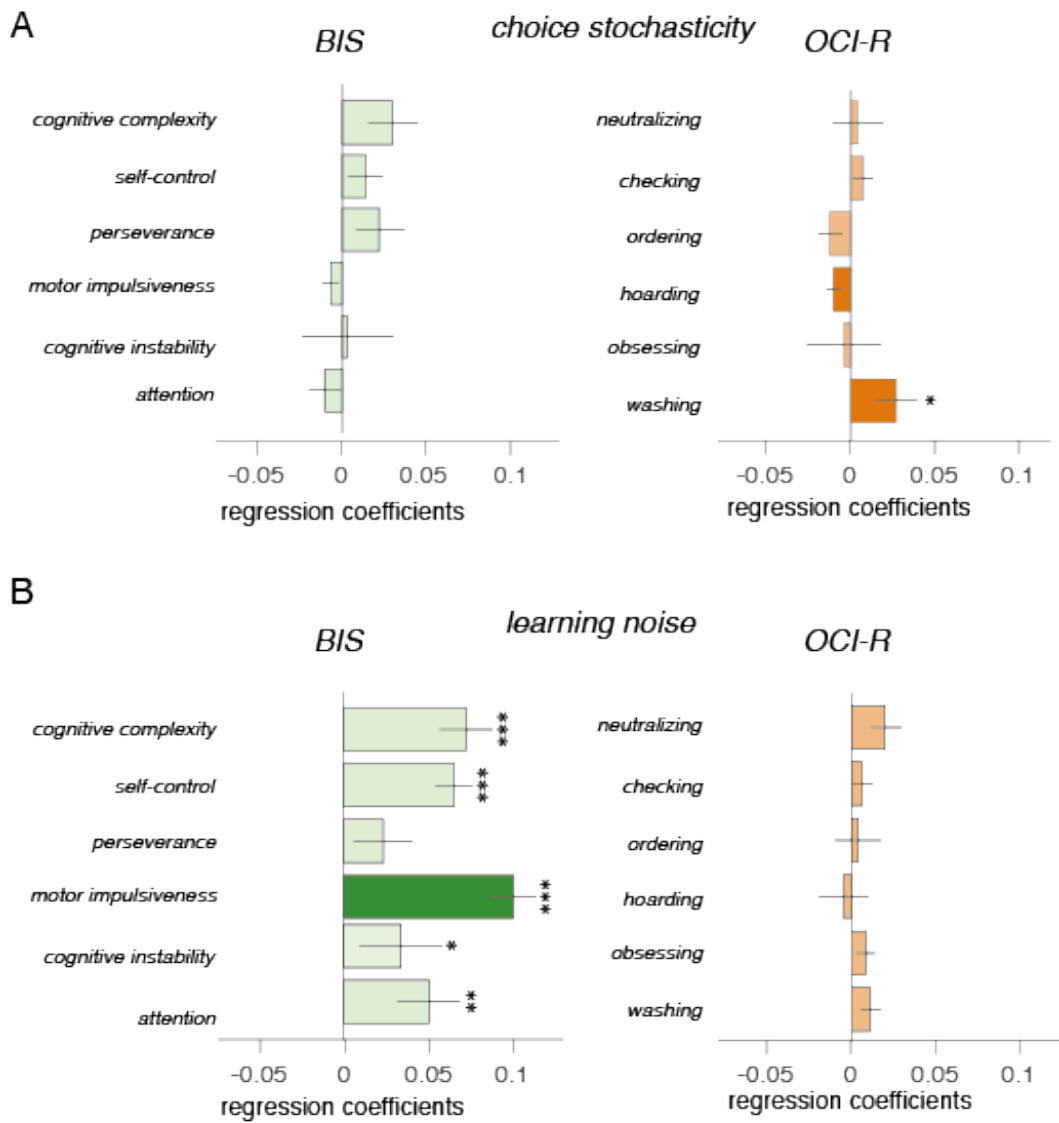
Figure 3 Experiment 1 N = 392



**Figure 3. Contributions of different subscales of impulsivity and obsessive-compulsive traits to choice stochasticity and learning noise in experiment 1 (N = 392).**

**A.** Analysis of impulsivity and compulsivity sub-scales show that hoarding and washing compulsions (right panel) were most closely associated with choice stochasticity in the noisy model, whilst there was no association with any of the impulsivity subscores (left panel). **B.** In contrast, learning noise was exclusively associated with motor impulsiveness, but no other impulsivity or compulsivity subscore. Error bars are standard errors, \* -  $p < 0.05$ , star represents test results Bonferroni corrected for multiple comparisons across regression models.

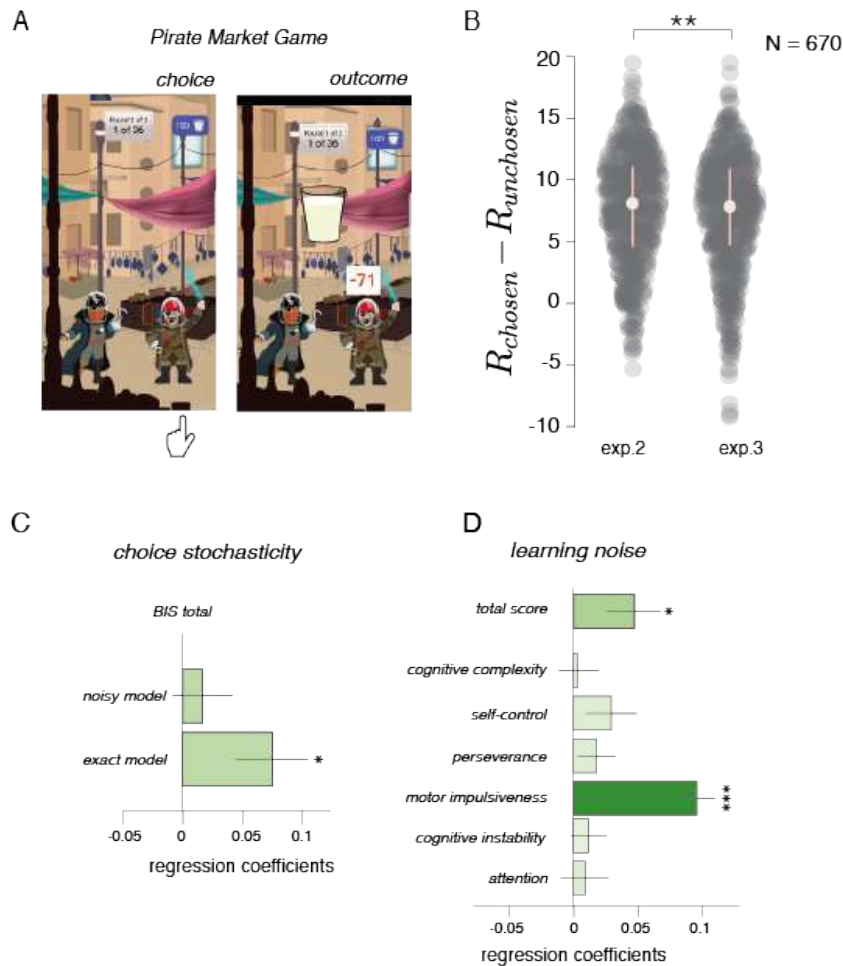
Figure 4 Experiment 2 N = 2610



**Figure 4. Contributions of different subscales of impulsivity and obsessive-compulsive traits to choice stochasticity and learning noise in experiment 2 (N = 2610).**

**A.** Choice stochasticity from the noisy model was again associated with washing compulsions (replicating the first experiment), but not with hoarding or any other subscore. **B.** Learning noise was again primarily associated with motor impulsiveness, as in experiment 1. No compulsivity subscale was associated with learning noise. Error bars are standard errors, \* -  $p < 0.05$ , \*\* -  $p < 0.01$ , \*\*\* -  $p < 0.001$ . Dark bar colours indicate the replication tests based on the results from experiment 1.

Figure 5



**Figure 5. Learning noise also linked to (motor) impulsivity in punishment learning (N = 670).**

**A.** Trial structure for the Pirate Market game in Brain Explorer app – a version of the “Milky Way” game but in the loss domain. This was framed as choosing the one of two pirates, who takes away less milk (hence, we called it ‘Pirate Market’). On every trial, users are endowed with 100 points (a full bucket of milk) and must choose a pirate that will steal less milk from them. The amount of stolen milk is presented in points (range from -99 to -1). The game setup and the outcome sequences were analogue to the Milky Way game but in the loss domain. **B.** Relative average rewards (chosen – unchosen) cumulated by the same sample of users (N = 670) in reward learning game (Milky Way, left) and in the punishment learning game (Pirate Market, right). White dots are medians, error bars are 25<sup>th</sup> and 75<sup>th</sup> percentile of the performance distributions. As in the reward learning context participants performed significantly better than chance ( $t(669)=38.8$ ,  $p<0.0001$ ), demonstrating they understood the task and learned to choose the less punishing bandit. Performance between the two domains was positively associated ( $r=0.33$ ,  $p<0.0001$ ), meaning that those who performed better in the reward version also performed better in the punishment version, even though they won slightly less in the punishment version ( $t(669)=3.00$ ,  $p=0.003$ , two-tailed). **C.** BIS total score was associated with choice stochasticity in the exact, but not in the noisy learning model. **D.** In the better fitting noisy model, BIS total score was associated with learning noise. Amongst the subscales, motor impulsiveness again showed the strongest association. \* -  $p < 0.05$ , \*\* -  $p < 0.01$ , \*\*\* -  $p < 0.001$ .

## Methods

### *Participants*

#### *Experiment 1*

420 participants living in the UK were recruited via the Prolific online testing platform (<https://www.prolific.co>). All participants were over 18 y.o. and gave their informed consent.

There were no other recruitment restrictions to gain a representative sample of the general public. The study was approved by University College London's Research Ethics Committee.

We excluded 28 participants because of incomplete dataset or failure to at least one of the attention checks (instructed questionnaire answers). This resulted in a final sample of 392 participants (age  $32.6 \pm 12.4$  (mean  $\pm$  s.d.), range (18-72), 204F). Participants received monetary compensation for their participation (8.25£) in the study, and an additional performance-based bonus of up to 2£ at the end of the experiment.

#### *Experiment 2 and experiment 3*

Between December 2020 and June 2021, 2790 users played at least one full session of the reward learning game ("Milky Way") using a smartphone-based app as part of the Brain Explorer project (<https://brainexplorer.net>). We excluded users with multiple but incomplete sessions, as well as with incomplete demographic or questionnaire data. In total, 2610 (age  $41.8 \pm 15.9$  (mean  $\pm$  s.d.), range (18-85), 1499F) out of 2790 users completed both OCI-R and BIS questionnaires, passed the data quality checks and were included in the final sample. Among these 2610 users, 670 (age  $42.8 \pm 17.3$  (mean  $\pm$  s.d.), range (18-83), 415F) also played at least one session of the punishment



learning game (“Pirate Market”). All participants first played the Milky Way reward learning game and then the Pirate market game.

## **Experimental protocol**

### ***Experiment 1.***

Participants played a restless, two-armed bandit game where their goal was to maximize the number of points won that were translated into a monetary bonus at the end of the game. On each trial, participants chose one of the two coloured shapes presented to the left and to the right of the fixation point on the screen and observed an outcome (Fig.1A). Choices were made in a self-paced manner with no time restrictions, but participants were instructed to complete the task within one hour. The task was divided into two sessions of 72 trials each. Each session involved a separate pair of coloured shapes.

All participants completed a training session and had to pass the quiz prior to starting the game. The task included continuous payoff ranging from 1 to 99 points which were sampled independently for each bandit from probability distribution whose properties were validated in the previous experiment<sup>24</sup> (Fig.1B). The mean payoffs on trial  $t$   $\hat{r}_t$  followed a random walk process and was sampled from a beta distribution with shape parameters  $\alpha = 1 + \hat{r}_{t-1} \exp(\tau)$  and  $\beta = 1 + (1 - \hat{r}_{t-1}) \exp(\tau)$ . This parameterization corresponds to a mode equal to  $\hat{r}_{t-1}$  and a spread growing monotonically with  $\tau$ , fixed to 3.0. Participants did not observe these mean payoffs directly but instead were shown the rounded to the nearest integer payoffs which were sampled from another beta distribution with shape parameters  $\alpha = 1 + \hat{r}_t \exp(\omega)$  and  $\beta = 1 + (1 -$

$\hat{r}_t) \exp(\omega)$ . The mode of this distribution on trial  $t$   $r_t$  corresponded to  $\hat{r}_t$  and a spread growing monotonically with  $\omega$ , fixed to 1.5. While parameter  $\tau$  controlled the overall volatility of the environment (with on average 1 reversal in every 16 trials), parameter  $\omega$  controlled additional instantaneous fluctuations around the mean payoff that were introduced to insure a sufficient level of uncertainty and continuous learning throughout the session.

To analyse the subjects' overall performance in the task, we computed the difference between the total average reward won by each subject and the foregone reward (to account for the individually generated reward trajectories). A positive difference means that subjects were performing better than choosing at chance (Fig.1D).

## ***Experiment 2***

*Milky Way game.* The game represented a gamified version of the laboratory task used in experiment 1. As before, participants played a restless, two-armed bandit game but instead of the coloured shapes, the two bandits were depicted as brown and black-and-white space cows (Fig.1C). The goal of the game consisted of accumulating as many points as possible that were translated into a "space milk". On each trial, participants made a self-paced choice between the two cows and observed the number of points won. Participants only observed the number of points from the chosen cow. Participants did not receive any monetary compensation for playing the game. Each game session consisted of 72 trials split into two rounds of 36 trials each. On average the users played  $1.3 \pm 0.58$  (mean  $\pm$  s.d.) sessions of the Milky Way game resulting in

93.0±42.1 (mean±s.d.) trials per user. Reward magnitudes were generated from the distributions with the same volatility and sampling parameters as in the experiment 1.

### ***Experiment 3***

*Pirate Market game.* The game was similar to the Milky Way game but instead of space cows that distributed “space milk”, participants encountered space pirates that were stealing the milk from them. The goal of the game was to preserve as much milk as possible but avoiding the most ravenous pirate (Fig.5A). On every trial, participants were endowed with 100 points and made a choice between two pirates, they next observed how many points (between 1 and 99) the chosen pirate had stolen, and the remaining points were added to their cumulative score. Each game session consisted of 72 trials split into two rounds. On average the users played 1.25±0.48 (mean±s.d.) sessions of the Milky Way game resulting in 90.2±34.8 (mean±s.d.) trials per user.

For both games, the payoffs were sampled from the beta distributions with the same shape parameters as in experiment 1 (Fig.1B). We verified that participants also in this version performed significantly better than chance ( $t(669)=38.8$ ,  $p<0.0001$ ), demonstrating they understood the task and learned to choose the less punishing bandit. Interestingly, performance between the two domains was positively associated ( $r=0.33$ ,  $p<0.0001$ ), meaning that those who performed better in the reward version also performed better in the punishment version, even though they won slightly less in the punishment version (Fig.5B;  $t(669)=3.00$ ,  $p=0.003$ ).

### ***Computational model***

To model participants' choice behaviour, we used the same models that were previously developed and validated<sup>24</sup>. Choice behaviour was modelled using two versions of the reinforcement learning algorithm. In the first model, we deployed a traditional *exact* Rescorla-Wagner<sup>83</sup> learning rule that was used to update the state-action Q-values on each trial  $t$  following action at trial  $t-1$  and obtained reward  $r$ :

$$Q_t = Q_{t-1} + \alpha (r_{t-1} - Q_{t-1})$$

where  $\alpha$  is the learning rate that scales the prediction error (PE) between obtained reward  $r_{t-1}$  and expected reward  $Q_{t-1}$ . In this formulation, the update of the values is deterministic and only depends on the learning rate.

In our previous work<sup>24</sup> we introduced a “noisy” formulation of this model which assumes stochasticity in the update rule. On each trial, the updated value is corrupted by additive random noise  $\varepsilon_t$ :

$$Q_t = Q_{t-1} + \alpha (r_{t-1} - Q_{t-1}) + \varepsilon_t$$

where  $\varepsilon_t$  is drawn from a Normal distribution with zero mean and standard deviation  $\sigma_t$  equal to a constant fraction  $\zeta$  of the magnitude of the PE:  $\sigma_t = \zeta |r_{t-1} - Q_{t-1}|$ . In this formulation of the noisy learning model, the noise added at each update scales with the prediction error similarly to Weber's law. Previous studies have demonstrated its better performance and advantage over models with just random noise (e.g., where the standard deviation  $\sigma_t$  does not scale with the prediction error)<sup>24,84</sup>.

The learning noise was applied to the update of both chosen and unchosen option values. While in the complete feedback (see Supplementary Information), the update of both options was based on the observed outcomes, in the partial feedback condition (as described in the main manuscript), we assumed a “decaying” learning rule for the unchosen option where the reward was replaced by the average payoff equal to 50 in our task similarly to the previous experiment<sup>24</sup>. In both models the choice process was modelled as a stochastic ‘softmax’ action selection policy, controlled by an ‘temperature’  $\beta$  and an optional ‘choice repetition bias’  $\xi$ :

$$P(A) a_t = \left( \frac{1}{1 + \exp(-\beta (Q_{t,A} - Q_{t,B}) - \xi \text{sign}(a_{t-1}))} \right)$$

where  $Q_{t,A}$  and  $Q_{t,B}$  are the values associated with options A and B at time point  $t$ . This stochastic action selection policy reduces to a purely greedy (value-maximizing) ‘argmax’ policy when  $\beta \rightarrow \infty$ .

### ***Model fitting procedure***

Model fitting was conducted using Monte Carlo methods<sup>85</sup> following the implementation developed in the previous work<sup>24</sup>. The model with the exact learning rule was fitted using an Iterated Batch Importance Sampler (IBIS)<sup>86</sup> to obtain the posterior distributions for the parameters  $\theta$  when the likelihoods  $p(a_t | a_{1:t-1}, r_{1:t-1}, \theta)$  are tractable<sup>24</sup>. This class of Monte Carlo methods could not be used for the noisy RL models because the corresponding likelihoods are intractable in this latter case. To obtain posterior parameter distributions for the noisy learning model we used the ‘SMC<sup>2</sup>’ algorithm<sup>87</sup>. To obtain action values trajectories in the noisy learning model, we applied smoothing procedures to estimate the smoothing distributions of the

trajectories of action values over the course of each block  $p(Q_{1:n}|a_{1:n}, r_{1:n}, \theta^{MAP})$ , where  $n$  corresponds to the number of trials in each block, and  $\theta^{MAP}$  to maxima a posteriori for model parameters  $\{\alpha, \zeta, \beta, \xi\}$ . To obtain samples approximately distributed under the smoothing distributions, we used the Forward Filter Backward Simulator (FFBSi)<sup>88,89</sup> to obtain  $K$  samples  $\tilde{Q}_{1:n,k}$ . For details of the computational model and model inversion procedures, see <sup>24</sup>.

### ***Model comparison***

We performed a series of Bayesian Model Selection (BMS) analyses to test whether learning imprecisions help account for choice data across the experiments. In all model comparisons we used model evidence conditioned on human decisions obtained from the particle MCMC fitting procedures (Methods,<sup>24</sup>). We used random-effect model selection procedure where models are treated as random effects that could differ between subjects and their prior frequencies are drawn from Dirichlet distribution. Model posterior frequencies and exceedance probabilities (e.g., how likely it is that a given model is more frequent in the population than other models in the set) were obtained through the BMS algorithm implemented in SPM12<sup>90,91</sup>.

To identify whether learning imprecision is an important source of behavioural variability, we compared three models across experiment 1 and 2: 1) *exact* model with just choice stochasticity in the softmax decision rule 2) model with just learning noise and no choice stochasticity (argmax choice policy) and 3) model with both sources of variability: learning noise and softmax. In both experiments model with both sources of variability better explained the data (Fig.1E-F).

To verify whether subjects inferred the value of the unchosen option we compared the model with no update of the counterfactual option with the model that assumed a progressive decay of

the counterfactual option to the mean of 50 in both exact and noisy learning model formulations (SFig.1B)<sup>24</sup>. We performed this analysis for both experiment 1 and experiment 2. In both, exact and noisy model formulations, the model which assumed the update of the unchosen option better accounted for the data (all exceedance probabilities > 0.99) (SFig.1B) suggesting that in both experiments users inferred the values of the unchosen option when making their choices. In both experiments, learning rates for the forgone option were lower than learning rates for the chosen option: experiment 1 chosen  $0.59 \pm 0.01$  (mean  $\pm$  s.e.m.), unchosen  $0.24 \pm 0.008$  (mean  $\pm$  s.e.m.), difference  $t(391)=25.81$ ,  $p < 0.0001$ ; experiment 2 chosen  $0.69 \pm 0.004$  (mean  $\pm$  s.e.m.), unchosen  $0.38 \pm 0.003$  (mean  $\pm$  s.e.m.), difference  $t(2609)=59.21$ ,  $p < 0.0001$ , suggesting participants were putting more weight on the observed rather than inferred option values.

Additionally, we tested an alternative model formulation, in order to investigate whether additional choice hysteresis parameter is necessary to improve model behaviour fitting. We compared model with and without repetition bias separately for exact and noisy learning formulations in the experiment 1 (SFig.1C).

### ***Analysis of inter-individual differences in impulsivity and compulsivity***

#### ***Linear regressions***

#### ***Relationships with choice stochasticity and learning noise parameter***

To examine the relationship between model parameters and impulsive and compulsive psychiatric traits we conducted a series of linear regressions. For both Experiment 1, 2 and 3, dependent variables were choice stochasticity  $1/\beta$ , and the learning noise parameter  $\zeta$  that

scaled the amount of stochasticity at each update step proportionally to the prediction error magnitude.

Larger values of  $\zeta$  resulted in larger deviations of the noisy action values from the exact action values. Additionally, in experiment 1 we also ran linear regressions with learning rate  $\alpha$  for the chosen option as dependent variable but did not proceed further since no significant results have been observed. For all dependent variables, other model parameters have been included as covariates to control for the shared variance due to the correlation between the model parameters (SFig.6A-B). In all models, age and gender have been included as covariates. In addition, in experiment 1 IQ scores was included as additional covariate, and in experiment 2 and 3, mental health status (a binary variable encoded as 1 if the user indicated to have a current or past diagnosed mental health condition and 0 otherwise). All continuous regressors were z-scored to ensure comparability of regression coefficients. We performed separate regressions for each total score and subscore. We computed robust multiple regression models using the *rlm* function of the MASS package<sup>92</sup> in R with a Hampel and bisquare methods to down-weight outliers and assessed their p values using the *rob.pvals* function of the repmud R package. For regression models in the experiment 1 we applied Bonferroni corrections for multiple comparisons over the number of tests performed for each scale (6 tests for BIS scale and 6 tests for OCI-R scale, total 12 tests, adjusted thresholds for Type 1 error at 0.05 results in a corrected  $\alpha=0.004$ ). For experiment 2, we did not correct our results for multiple comparisons as the performed tests were hypothesis-driven based on the results obtained in the experiment 1.



Additionally, we computed pair-wise correlations between model parameters, scales, and demographic variables: age, gender, IQ (experiment 1) and mental health status (experiment 2) (SFig.6).

### ***Analysis of contribution of learning noise to behavioural variability***

To estimate the fraction of behavioural variability explained by the learning noise illustrated on Figures 1E-F, we estimated separate contributions of learning noise and choice stochasticity to total behavioural variability. To estimate the behavioural variability generated by learning noise, we computed the standard deviation of the difference between the action values based on the noisy smoothing distributions  $\tilde{Q}_t$  (i.e., with added Weber noise at the value update at time step  $t-1$ ) and action values  $\hat{Q}_t$  obtained through the exact application of Rescorla-Wagner learning rule to noisy action values  $\tilde{Q}_{t-1}$  at time step  $t-1$ . To estimate the behavioural variability generated by choice stochasticity, we approximated the softmax (logistic) distribution as the cumulative probability function for the Gaussian normal distribution (see Supplementary Modeling notes of Findling et al., 2019<sup>24</sup> for details). From this approximation, we computed the standard deviation of the logistic distribution for the inverse temperature  $\beta$  as  $\frac{\pi}{\beta\sqrt{3}}$ .

To characterize the contribution of learning noise to exploration, we estimated the fraction of non-greedy decisions that could be caused by learning noise in experiment 1 and experiment 2. First, we labelled decisions as non-greedy based on the fits from the exact model. Non-greedy exact decisions were defined as choices where participant did not maximize their value  $\{Q_{t,A}, Q_{t,B}\}$  as predicted by the exact model<sup>31</sup>. Next, we generated the smoothing distributions of action values based on the fits from the noisy learning model and averaged across smoothing

trajectories to obtain noisy action values  $\{\tilde{Q}_{t,A}, \tilde{Q}_{t,B}\}$ . We then labelled as noise-driven all the exact non-greedy trials for which the noisy action values predicted an opposite ranking of action values relative the exact model:

$$\text{sign}(Q_{t,A} - Q_{t,B}) \neq \text{sign}(\tilde{Q}_{t,A} - \tilde{Q}_{t,B})$$

### **Code availability**

The games from experiment 2 and 3 are freely playable on the Brain Explorer app for Android and Apple handheld devices (access via [www.brainexplorer.net](http://www.brainexplorer.net) or the respective app stores).

The code and data to reproduce the main analyses are freely available in an Open Science Framework (OSF) repository at

[https://osf.io/h5ptx/?view\\_only=5c7ff5a76fd74cb0a4a9086a5f6a34b3](https://osf.io/h5ptx/?view_only=5c7ff5a76fd74cb0a4a9086a5f6a34b3)

### **Author Contributions**

V.S and T.U.H conceived of and designed the study and analysis plan. VS coded the experiment, collected and analysed the data. TUH and VS wrote the paper.

### **Acknowledgements**

TUH is supported by a Sir Henry Dale Fellowship (211155/Z/18/Z) from Wellcome & Royal Society, a grant from the Jacobs Foundation (2017-1261-04), the Medical Research Foundation, a 2018 NARSAD Young Investigator grant (27023) from the Brain & Behavior Research Foundation, and

a Philip Leverhulme Prize from the Leverhulme Trust (PLP-2021-040). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 946055). The Max Planck UCL Centre is a joint initiative supported by UCL and the Max Planck Society. The Wellcome Centre for Human Neuroimaging is supported by core funding from the Wellcome Trust (203147/Z/16/Z).

This research was funded in whole, or in part, by the Wellcome Trust (211155/Z/18/Z). For the purpose of Open Access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

### **Competing interests**

The authors declare no competing interests.

## References

1. Robbins, T.W., Gillan, C.M., Smith, D.G., de Wit, S., and Ersche, K.D. (2012). Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends in Cognitive Sciences* 16, 81–91.
2. Ziegler, G., Hauser, T.U., Moutoussis, M., Bullmore, E.T., Goodyer, I.M., Fonagy, P., Jones, P.B., Lindenberger, U., and Dolan, R.J. (2019). Compulsivity and impulsivity traits linked to attenuated developmental frontostriatal myelination trajectories. *Nature Neuroscience* 22, 992–999.
3. Chamberlain, S.R., Naomi Fineberg, M.A., Blackwell, A.D., Robbins, T.W., and Sahakian, B.J. (2006). BRIEF REPORTS [ajp.psychiatryonline.org](http://ajp.psychiatryonline.org) 11.
4. Lawrence, A.J., Luty, J., Bogdan, N.A., Sahakian, B.J., and Clark, L. (2009). Problem gamblers share deficits in impulsive decision-making with alcohol-dependent individuals. *Addiction* 104, 1006–1015.
5. Garavan, H., and Stout, J.C. (2005). Neurocognitive insights into substance abuse. *Trends in Cognitive Sciences* 9, 195–201.
6. Chamberlain, S.R., Fineberg, N.A., Menzies, L.A., Blackwell, A.D., Bullmore, E.T., Robbins, T.W., and Sahakian, B.J. (2007). Impaired cognitive flexibility and motor inhibition in unaffected first-degree relatives of patients with obsessive-compulsive disorder. *American Journal of Psychiatry* 164, 335–338.
7. Kanen, J.W., Luo, Q., Yellowlees, R., Arntz, F.E., Christmas, D.M., Sahakian, B.J., Apergis-Schoute, A.M., Cardinal, R.N., and Robbins, T.W. Perseverative behavior under uncertainty is enhanced by tryptophan depletion but diminished with subclinical obsessive-compulsive symptoms.
8. Gillan, C.M., Kalanthroff, E., Evans, M., Weingarden, H.M., Jacoby, R.J., Gershkovich, M., Snorrason, I., Campeas, R., Cervoni, C., Crimarco, N.C., et al. (2020). Comparison of the Association Between Goal-Directed Planning and Self-reported Compulsivity vs Obsessive-Compulsive Disorder Diagnosis. *JAMA Psychiatry* 77, 77–85.
9. Raio, C.M., Konova, A.B., and Otto, A.R. (2020). Trait impulsivity and acute stress interact to influence choice and decision speed during multi-stage decision-making. *Scientific Reports* 10, 7754.
10. Dalley, J.W., and Robbins, T.W. (2017). Fractionating impulsivity: neuropsychiatric implications. Nature Publishing Group.
11. Moeller, F.G., Barratt, E.S., Dougherty, D.M., Schmitz, J.M., and Swann, A.C. (2001). Psychiatric Aspects of Impulsivity. *American Journal of Psychiatry* 158, 1783–1793.
12. Izquierdo, A., and Jentsch, J.D. (2012). Reversal learning as a measure of impulsive and compulsive behavior in addictions. *Psychopharmacology* 219, 607–620.
13. Benzina, N., Pelissolo, A., Mallet, L., Burguière, E., and Mondor-Albert Chenevier, H. A cross-species assessment of behavioral flexibility in compulsive disorders.
14. Marzuki, A.A., Tomić, I., Ip, S.H.Y., Gottwald, J., Kanen, J.W., Kaser, M., Sule, A., Conway-Morris, A., Sahakian, B.J., and Robbins, T.W. (2021). Association of Environmental Uncertainty With Altered Decision-making and Learning Mechanisms in Youths With Obsessive-Compulsive Disorder. *JAMA Network Open* 4, e2136195.

15. Kanen, J.W., Ersche, K.D., Fineberg, N.A., Robbins, T.W., and Cardinal, R.N. (2019). Computational modelling reveals contrasting effects on reinforcement learning and cognitive flexibility in stimulant use disorder and obsessive-compulsive disorder: remediating effects of dopaminergic D2/3 receptor agents. *Psychopharmacology* 236, 2337–2358.
16. Fradkin, I., Ludwig, C., Eldar, E., and Huppert, J.D. (2020). Fradkin\_PlosCB.pdf. *PLoS Comput Biol* 16, e1007634.
17. Hauser, T.U., Iannaccone, R., Dolan, R.J., Ball, J., Hattenschwiler, J., Drechsler, R., Rufer, M., Brandeis, D., Walitza, S., and Brem, S. (2017). Increased fronto-striatal reward prediction errors moderate decision making in obsessive-compulsive disorder. *Psychological Medicine* 47, 1246–1258.
18. Pushkarskaya, H., Tolin, D., Ruderman, L., Henick, D., Kelly, J.M.L., Pittenger, C., and Levy, I. (2017). Value-based decision making under uncertainty in hoarding and obsessive-compulsive disorders. *Psychiatry Research* 258, 305–315.
19. Hauser, T.U., Fiore, V.G., Moutoussis, M., and Dolan, R.J. (2016). Computational Psychiatry of ADHD: Neural Gain Impairments across Marrian Levels of Analysis. *Trends in Neurosciences* 39, 63–73.
20. Dubois, M., Habicht, J., Michely, J., Moran, R., Dolan, R., and Hauser, T. (2020). Noradrenaline modulates tabula-rasa exploration. *bioRxiv*.
21. Dubois, M., Bowler, A., Moses-Payne, M., Habicht, J., Steinbeis, N., and Hauser, T.U. (2020). Tabula-rasa exploration decreases during youth and is linked to ADHD symptoms. *bioRxiv*.
22. Williams, J., and Dayan, P. (2005). Dopamine, learning, and impulsivity: A biological account of attention-deficit/hyperactivity disorder. *Journal of Child and Adolescent Psychopharmacology* 15, 160–179.
23. Drugowitsch, J., Wyart, V., Devauchelle, A.-D., and Koehlin, E. (2016). Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality. *Neuron*, 1–14.
24. Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S., and Wyart, V. (2019). Computational noise in reward-guided learning drives behavioral variability in volatile environments. *Nature Neuroscience* 22, 2066–2077.
25. Findling, C., and Wyart, V. (2021). Computation noise in human learning and decision-making: origin, impact, function. *Current Opinion in Behavioral Sciences* 38, 124–132.
26. del Campo, N., Chamberlain, S.R., Sahakian, B.J., and Robbins, T.W. (2011). The roles of dopamine and noradrenaline in the pathophysiology and treatment of attention-deficit/hyperactivity disorder. *Biological Psychiatry*.
27. Costa, A., la Fougère, C., Pogarell, O., Möller, H.J., Riedel, M., and Ettinger, U. (2013). Impulsivity is related to striatal dopamine transporter availability in healthy males. *Psychiatry Research - Neuroimaging* 211, 251–256.
28. Robinson, E.S., Eagle, D.M., Mar, A.C., Bari, A., Banerjee, G., Jiang, X., Dalley, J.W., and Robbins, T.W. (2008). Similar Effects of the Selective Noradrenaline Reuptake Inhibitor Atomoxetine on Three Distinct Forms of Impulsivity in the Rat. *Neuropsychopharmacology* 33, 1028–1037.

29. den Ouden, H.E.M., Daw, N.D., Fernandez, G., Elshout, J.A., Rijpkema, M., Hoogman, M., Franke, B., and Cools, R. (2013). Dissociable Effects of Dopamine and Serotonin on Reversal Learning. *Neuron* *80*, 1090–1100.
30. Gillan, C.M., and Rutledge, R.B. (2021). Annual Review of Neuroscience Smartphones and the Neuroscience of Mental Health.
31. Daw, N.D., O’Doherty, J.P., Dayan, P., Seymour, B., and Dolan, R.J. (2006). Cortical substrates for exploratory decisions in humans. *Nature* *441*, 876–9.
32. Ziegler, G., and Hauser, T.U. Compulsivity and impulsivity traits linked to attenuated developmental frontostriatal myelination trajectories. *Nature Neuroscience*.
33. Hauser, T.U., Moutoussis, M., Dayan, P., Dolan, R.J., and Consortium, N. (2017). Increased decision thresholds trigger extended information gathering across the compulsivity spectrum. *Translational Psychiatry* *7*, 1296.
34. Gillan, C.M., Kosinski, M., Whelan, R., Phelps, E.A., and Daw, N.D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goaldirected control. *eLife* *5*, 1–24.
35. Seow, T.X.F., and Gillan, C.M. (2020). Transdiagnostic Phenotyping Reveals a Host of Metacognitive Deficits Implicated in Compulsivity. *Scientific Reports* *10*, 1–11.
36. Dubois, M., and Hauser, T.U. (2022). Exploring too much? The role of exploration in impulsivity . *Nature Communications* *XX*, XX–XX.
37. Patton, J.H., Stanford, M.S., and Barratt, E.S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of clinical psychology* *51*, 768–774.
38. Foa, E.B., Huppert, J.D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., and Salkovskis, P.M. (2002). The obsessive-compulsive inventory: Development and validation of a short version. *Psychological Assessment* *14*, 485–496.
39. Mathias, C., Stanford, M.S., Mathias, C.W., Dougherty, D.M., Lake, S.L., Anderson, N.E., and Patton, J.H. (2009). Fifty years of the Barratt Impulsiveness Scale : An update and review Fifty years of the Barratt Impulsiveness Scale : An update and review. *Personality and Individual Differences* *47*, 385–395.
40. Wootton, B.M., Diefenbach, G.J., Bragdon, L.B., Steketee, G., Frost, R.O., and Tolin, D.F. (2015). A contemporary psychometric evaluation of the Obsessive Compulsive Inventory-Revised (OCI-R). *Psychological assessment* *27*, 874–882.
41. Patton, J.H., Stanford, M.S., and Barratt, E.S. (1995). Factor structure of the barratt impulsiveness scale. *Journal of Clinical Psychology* *51*, 768–774.
42. Stanford, M.S., Mathias, C.W., Dougherty, D.M., Lake, S.L., Anderson, N.E., and Patton, J.H. (2009). Fifty years of the Barratt Impulsiveness Scale: An update and review. *Personality and Individual Differences* *47*, 385–395.
43. Hauser, T.U., Fiore, V.G., Moutoussis, M., and Dolan, R.J. (2016). Computational Psychiatry of ADHD: Neural Gain Impairments across Marrian Levels of Analysis. *Trends in Neurosciences* *39*, 63–73.
44. Williams, J., and Taylor, E. (2006). The evolution of hyperactivity, impulsivity and cognitive diversity. *Journal of the Royal Society Interface* *3*, 399–413.
45. Hauser, T.U., Iannaccone, R., Ball, J., Mathys, C., Brandeis, D., Walitza, S., and Brem, S. (2014). Role of the medial prefrontal cortex in impaired decision making in juvenile attention-deficit/hyperactivity disorder. *JAMA Psychiatry* *71*, 1165–1173.

46. Lim, M.S.M., Jocham, G., Hunt, L.T., Behrens, T.E.J., and Rogers, R.D. Impulsivity and predictive control are associated with suboptimal action-selection and action-value learning in regular gamblers.
47. Rutledge, R.B., Lazzaro, S.C., Lau, B., Myers, C.E., Gluck, M.A., and Glimcher, P.W. (2009). Dopaminergic Drugs Modulate Learning Rates and Perseveration in Parkinson's Patients in a Dynamic Foraging Task. *Journal of Neuroscience* 29, 15104–15114.
48. Lau, B., and Glimcher, P.W. (2008). Value representations in the primate striatum during matching behavior. *Neuron* 58, 451–63.
49. Boorman, E.D., Behrens, T.E.J., Woolrich, M.W., and Rushworth, M.F.S. (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* 62, 733–43.
50. Evenden, J.L. (1999). Varieties of impulsivity. *Psychopharmacology* 146, 348–361.
51. Reynolds, B., Ortengren, A., Richards, J.B., and de Wit, H. (2006). Dimensions of impulsive behavior: Personality and behavioral measures. *Personality and Individual Differences* 40, 305–315.
52. Whitney, P., Jameson, T., and Hinson, J.M. (2004). Impulsiveness and executive control of working memory. *Personality and Individual Differences* 37, 417–428.
53. Cheung, A.M., Mitsis, E.M., and Halperin, J.M. (2004). The Relationship of Behavioral Inhibition to Executive Functions in Young Adults. *Journal of Clinical and Experimental Neuropsychology* 26, 393–404.
54. Loosen, A.M., Skvortsova, V., and Hauser, T.U. Obsessive-compulsive symptoms and information seeking during the Covid-19 pandemic.
55. Rouault, M., Seow, T., Gillan, C.M., and Fleming, S.M. (2018). Archival Report With Dissociable Shifts in Metacognition but Not Task Performance. *Biological Psychiatry* 84, 443–451.
56. Eyal, P., David, R., Andrew, G., Zak, E., and Ekaterina, D. (2021). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*.
57. Koutsouleris, N., Hauser, T.U., Skvortsova, V., and de Choudhury, M. (2022). From Promise to Practice: Crossing the Bars Towards AI-informed Mental Healthcare. *Lancet Psychiatry* XX, XX–XX.
58. Hauser, T.U., Skvortsova, V., de Choudhury, M., and Koutsouleris, N. (2022). The promise of a model-based psychiatry: building computational models of mental ill-health. *Lancet Digital Health* XX, XX–XX.
59. van Opheusden, B., Galbiati, G., Kuperwajs, I., Bnaya, Z., Li, Y., and Ma, W.J. (2021). Revealing the impact of expertise on human planning with a two-player board game. *PsyArXiv*.
60. Kosoy, E., Collins, J., Chan, D.M., Huang, S., Pathak, D., Agrawal, P., Canny, J., Gopnik, A., and Hamrick, J.B. (2020). Exploring Exploration: Comparing Children with RL Agents in Unified Environments.
61. Coughlan, G., Coutrot, A., Khondoker, M., Minihane, A.-M., Spiers, H., and Hornberger, M. (2019). Toward personalized cognitive diagnostics of at-genetic-risk Alzheimer's disease. *Proceedings of the National Academy of Sciences* 116, 9285.

62. Rutledge, R.B., Smittenaar, P., Zeidman, P., Brown, H.R., Adams, R.A., Lindenberger, U., Dayan, P., and Dolan, R.J. (2016). Risk Taking for Potential Reward Decreases across the Lifespan. *Current Biology* 26, 1634–1639.
63. Potts, G.F., George, M.R.M., Martin, L.E., and Barratt, E.S. (2006). Reduced punishment sensitivity in neural systems of behavior monitoring in impulsive individuals. *Neuroscience letters* 397, 130–134.
64. Wypych, M., Michałowski, J.M., Drożdziel, D., Borczykowska, M., Szczepanik, M., and Marchewka, A. (2019). Attenuated brain activity during error processing and punishment anticipation in procrastination – a monetary Go/No-go fMRI study. *Scientific Reports* 9, 11492.
65. Wypych, M., and Potenza, M.N. (2021). Impaired Learning From Errors and Punishments and Maladaptive Avoidance—General Mechanisms Underlying Self-Regulation Disorders? *Frontiers in Psychiatry* 11, 1667.
66. Economidou, D., Pelloux, Y., Robbins, T.W., Dalley, J.W., and Everitt, B.J. (2009). High Impulsivity Predicts Relapse to Cocaine-Seeking After Punishment-Induced Abstinence. *Biological Psychiatry* 65, 851–856.
67. Levy, F. (2008). Pharmacological and therapeutic directions in ADHD: Specificity in the PFC. *Behavioral and Brain Functions* 4, 12.
68. Robbins, T.W., Gillan, C.M., Smith, D.G., de Wit, S., and Ersche, K.D. (2012). Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends in Cognitive Sciences* 16, 81–91.
69. Dalley, J.W., Everitt, B.J., and Robbins, T.W. (2011). Impulsivity, Compulsivity, and Top-Down Cognitive Control. *Neuron* 69, 680–694.
70. Rostami Kandroodi, M., Cook, J.L., Swart, J.C., Froböse, M.I., Geurts, D.E.M., Vahabie, A.-H., Nili Ahmadabadi, M., Cools, R., and den Ouden, H.E.M. (2021). Effects of methylphenidate on reinforcement learning depend on working memory capacity. *Psychopharmacology* 238, 3569–3584.
71. Dubois, M., Habicht, J., Michely, J., Moran, R., Dolan, R.J., and Hauser, T.U. (2021). Human complex exploration strategies are enriched by noradrenaline-modulated heuristics. *eLife* 10, e59907.
72. Frank, M.J., Santamaria, A., O’Reilly, R.C., and Willcutt, E. (2007). Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. *Neuropsychopharmacology*.
73. Joshi, S., Li, Y., Kalwani, R.M.M., and Gold, J.I.I. (2015). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron* 89, 221–234.
74. Joshi, S., and Gold, J.I. Context-Dependent Relationships between Locus Coeruleus Firing Patterns and Coordinated Neural Activity in the Anterior Cingulate Cortex.
75. Aston-Jones, G., and Cohen, J.D. (2005). An integrative theory of locus coeruleus - norepinephrine function: Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience* 28, 403–450.
76. Arnsten, A.F.T., and Pliszka, S.R. (2011). Catecholamine influences on prefrontal cortical function: Relevance to treatment of attention deficit/hyperactivity disorder and related disorders. *Pharmacology Biochemistry and Behavior*.



77. Berridge, C.W., and Devilbiss, D.M. (2011). Psychostimulants as cognitive enhancers: the prefrontal cortex, catecholamines, and attention-deficit/hyperactivity disorder. *Biological psychiatry* 69, e101–e111.
78. Jepma, M., te Beek, E.T., Wagenmakers, E.-J., van Gerven, J.M.A., and Nieuwenhuis, S. (2010). The role of the noradrenergic system in the exploration-exploitation trade-off: a psychopharmacological study. *Frontiers in human neuroscience* 4, 170.
79. Caswell, A.J., Bond, R., Duka, T., and Morgan, M.J. (2015). Further evidence of the heterogeneous nature of impulsivity. *Personality and Individual Differences* 76, 68–74.
80. Lyke, J.A., and Spinella, M. (2004). Associations among aspects of impulsivity and eating factors in a nonclinical sample. *International Journal of Eating Disorders* 36, 229–233.
81. Lau, B., and Glimcher, P.W. (2005). Dynamic Response-by-Response Models of Matching Behavior in Rhesus Monkeys. *Journal of the Experimental Analysis of Behavior* 84, 555–579.
82. Gershman, S.J., Pesaran, B., and Daw, N.D. (2009). Human reinforcement learning subdivides structured action spaces by learning effector-specific values. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29, 13524–31.
83. Sutton, R.S., and Barto, A.G. (1998). *Introduction to Reinforcement Learning*. MIT Press Cambridge.
84. Findling, C., Chopin, N., and Koechlin, E. (2021). Imprecise neural computations as a source of adaptive behaviour in volatile environments. *Nature Human Behaviour* 5, 99–112.
85. Robert, C., and Casella, G. (2004). *Monte Carlo Statistical Methods* Springer.
86. Chopin, N. (2002). A sequential particle filter method for static models. *Biometrika* 89, 539–552.
87. Chopin, N., Jacob, P.E., and Papaspiliopoulos, O. (2013). SMC2: an efficient algorithm for sequential analysis of state space models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 75, 397–426.
88. Lindsten, F., and Schön, T.B. (2013). *Backward Simulation Methods for Monte Carlo Statistical Inference*. *Foundations and Trends® in Machine Learning* 6, 1–143.
89. Doucet, A., Godsill, S., and Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing* 10, 197–208.
90. Stephan, K.E., Penny, W.D., Daunizeau, J., Moran, R.J., and Friston, K.J. (2009). Bayesian model selection for group studies. *NeuroImage* 46, 1004–17.
91. Rigoux, L., Stephan, K.E., Friston, K.J., and Daunizeau, J. (2014). Bayesian model selection for group studies - Revisited. *NeuroImage* 84, 971–85.
92. Venables WN, and Ripley BD. (2002). *Modern Applied Statistics with S* Fourth Edition. (Springer).

# Figures

Figure 1

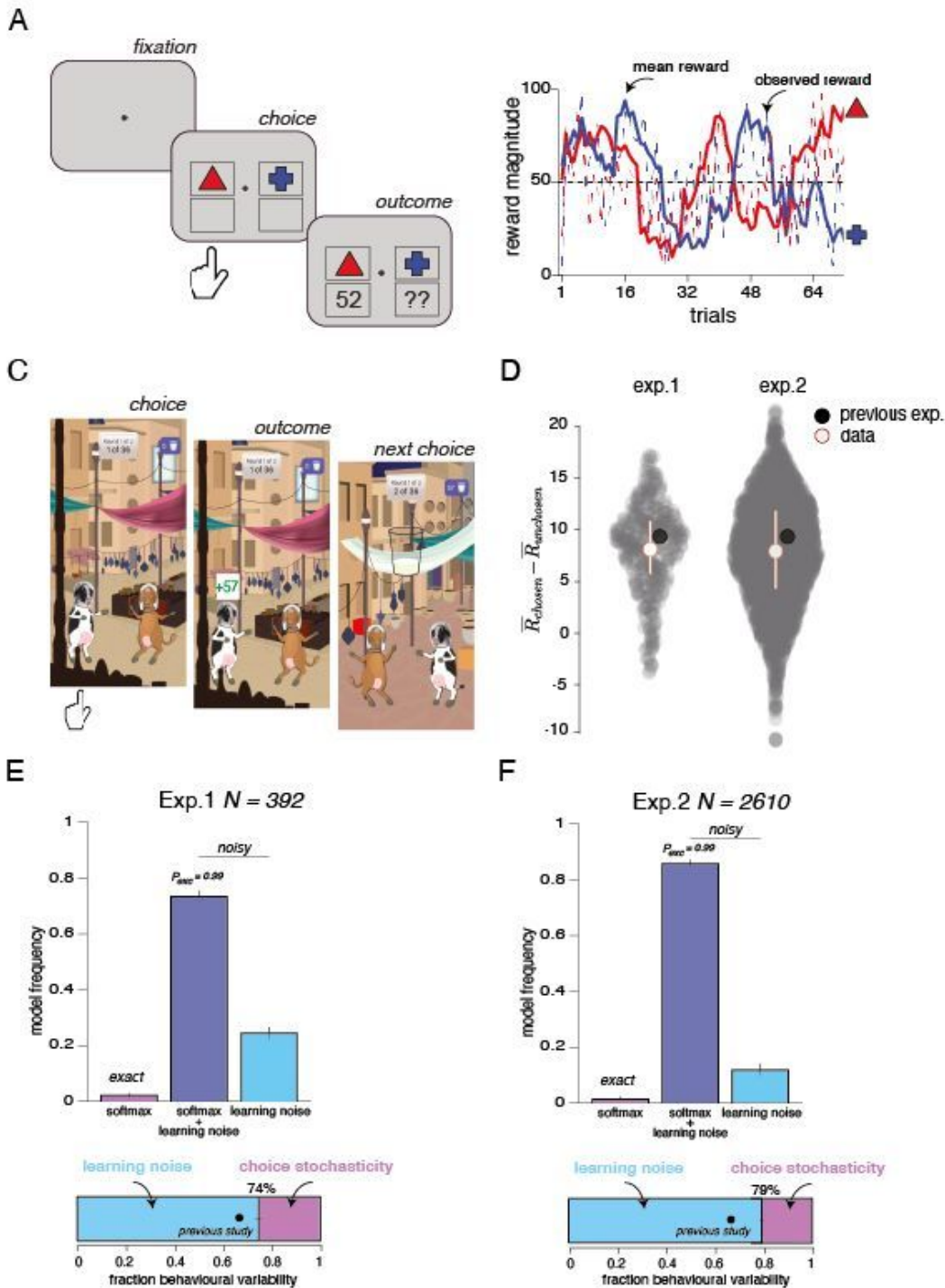


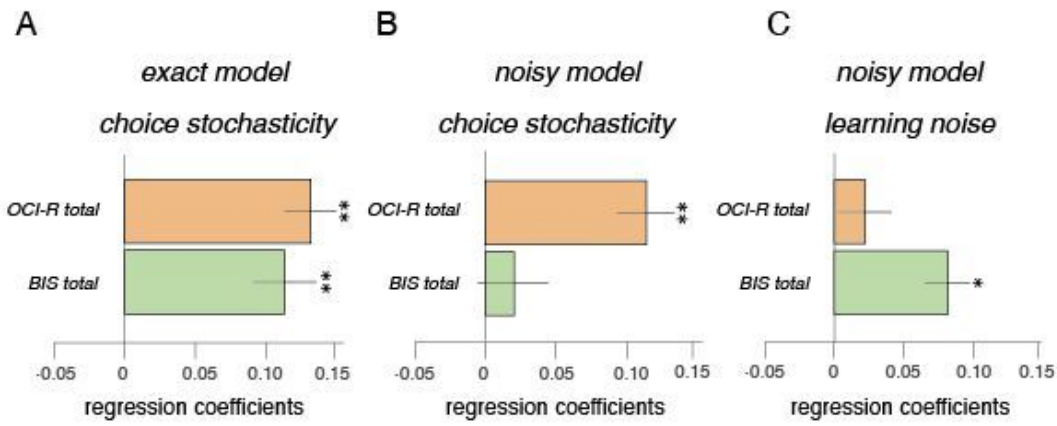
Figure 1

*Task structure and contributions of learning noise to behavioural variability*

**A.** Trial structure and design in two-armed bandit task in experiment 1. On every trial, participants were making a choice between two bandits represented by coloured shapes on the screen and observed the outcome (between 1 or 99 points) for the chosen option. **B.** Example of reward magnitudes for two bandits for one session in experiment 1. Reward magnitudes changed through trials and were sampled from two probability distributions with independently drifting means. Thick lines represent the drifting means of two distributions; dotted lines represent the reward outcomes that could have been observed by the subject conditioned on the choice. This observable reward magnitudes are sampled from the distribution mean with added Gaussian sampling noise to ensure continuous learning in the task. **C.** Task structure for reward learning two-armed bandit task in experiment 2 (Milky Way game in the Brain Explorer app). On every trial, participants were making a choice between two space cows and then observed how much space milk they have collected in the form of points (from 1 to 99). The underlying distributions of drifting mean rewards were the same as in experiment 1 but the task was substantially shorter. **D.** To analyse the subjects' overall performance in both experiments, we computed the difference between the total average reward won by each subject and the foregone reward (to account for the individually generated reward trajectories). A positive difference means that subjects were performing better than at chance level. In both experiment 1 and 2, subjects performed better than chance ( $p < 0.001$ ) White dots are medians, error bars are 25<sup>th</sup> and 75<sup>th</sup> percentile of the performance distributions. Grey dots represent individual subjects. Black dots are the median performance from the previous laboratory study<sup>24</sup>. **E. Top panels.** Bayesian model selection (BMS) results for experiment 1 (left panel) and experiment 2 (right panel). Bars represent estimated model frequencies for exact model (just softmax decision rule, left bar), noisy learning model with softmax decision rule (middle bar), and noisy learning model with argmax decision rule (right bar). Error bars are s.d. for model frequencies based on Dirichlet distribution. Model with two sources of variability (learning noise and softmax decision rule) better accounted for data in both experiments similarly to the previously observed results<sup>24</sup>. **Bottom panels.** Fraction of behavioural variability attributed to learning noise based on the best winning model in experiment 1 (left panel) and experiment 2 (right panel). In both experiments more than 2/3 of total variability in the decision was due to learning noise rather than choice stochasticity. Black dots represent the results obtained in the previous laboratory study<sup>24</sup>.

Figure 2

Exp.1 N = 392



Exp.2 N = 2610

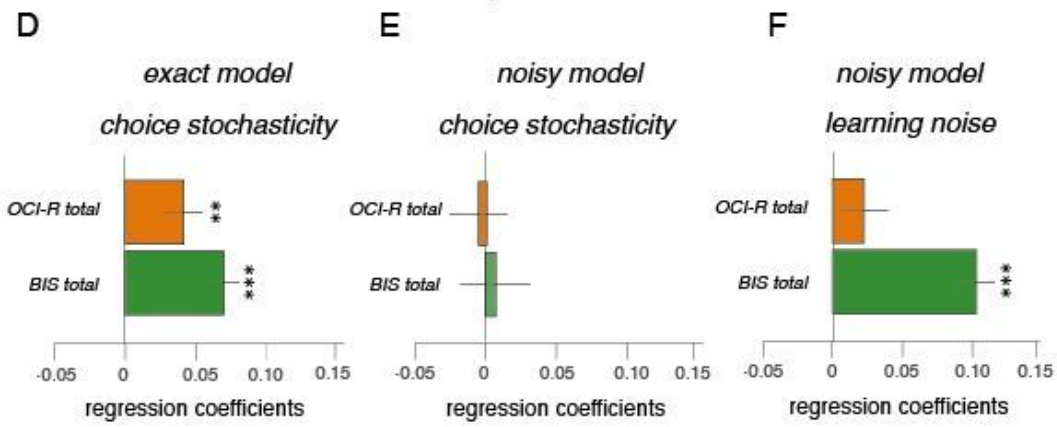


Figure 2

*Contributions of impulsivity and obsessive-compulsive traits to choice stochasticity and learning noise.*

**A.** In the exact model (no learning noise added), both impulsivity traits measured with BIS total score and OC traits measured with OCI-R total score showed a significant association with choice stochasticity (softmax parameter). **B.** Impulsivity traits were no longer associated with choice stochasticity when learning noise is added to the model. **C.** Impulsivity traits contributed to learning noise rather than choice stochasticity in the learning noise model. **D-F.** Same analysis as in A-C for the Brain Explorer (experiment 2) data. All regressions included age, gender, IQ (experiment 1), mental health status (experiment 2), and other model parameters as covariates. Dark bar colours indicate the replication tests based on the results from experiment 1. Error bars are standard errors, \* -  $p < 0.05$ , \*\* -  $p < 0.01$ , \*\*\* -  $p < 0.001$ .

Figure 3 Experiment 1 N = 392

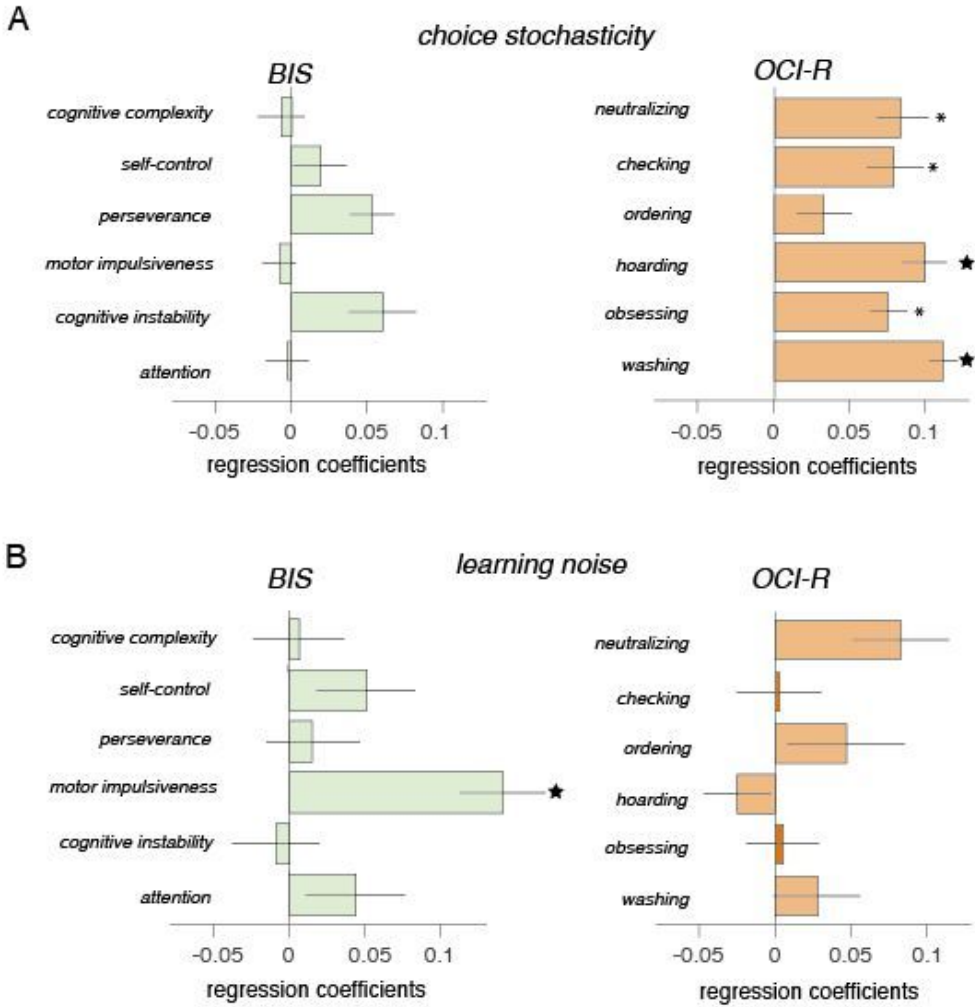


Figure 3

*Contributions of different subscales of impulsivity and obsessive-compulsive traits to choice stochasticity and learning noise in experiment 1 (N = 392).*

**A.** Analysis of impulsivity and compulsivity sub-scales show that hoarding and washing compulsions (right panel) were most closely associated with choice stochasticity in the noisy model, whilst there was no association with any of the impulsivity subscores (left panel). **B.** In contrast, learning noise was exclusively associated with motor impulsiveness, but no other impulsivity or compulsivity subscore. Error bars are standard errors, \* -  $p < 0.05$ , star represents test results Bonferroni corrected for multiple comparisons across regression models.

Figure 4 Experiment 2 N = 2610

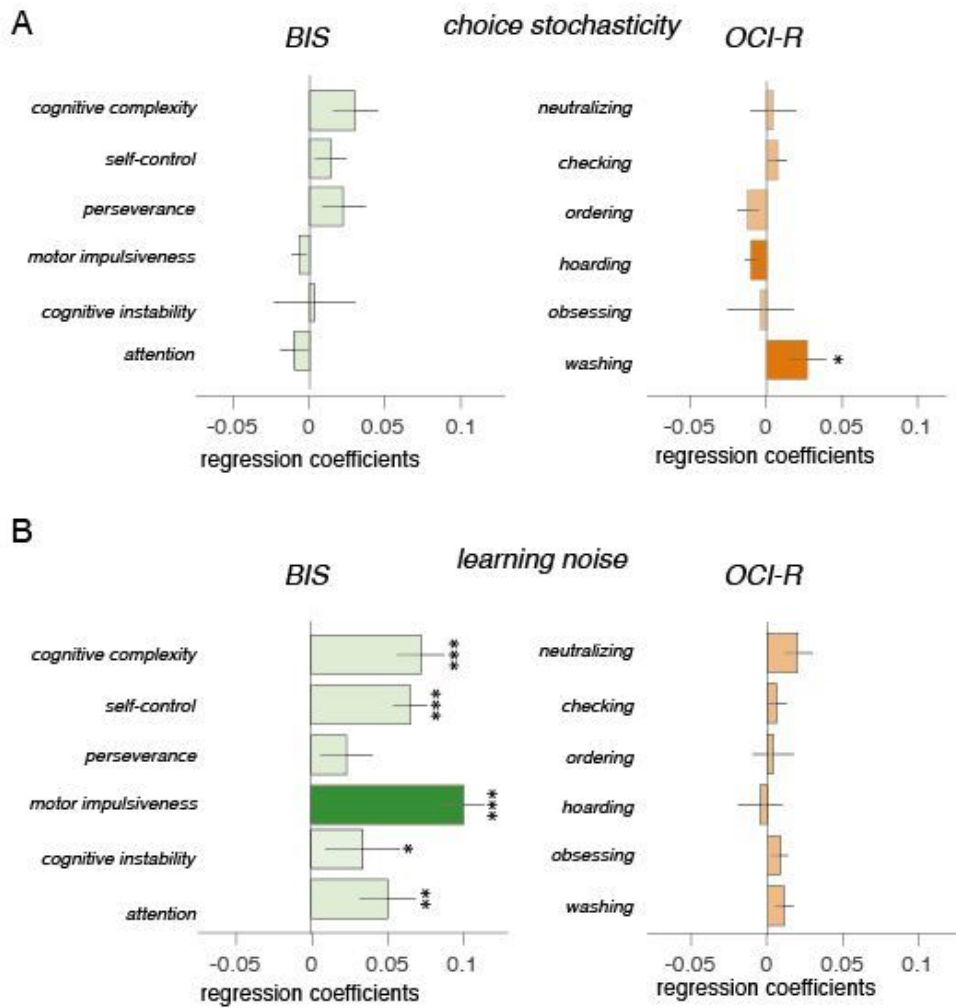


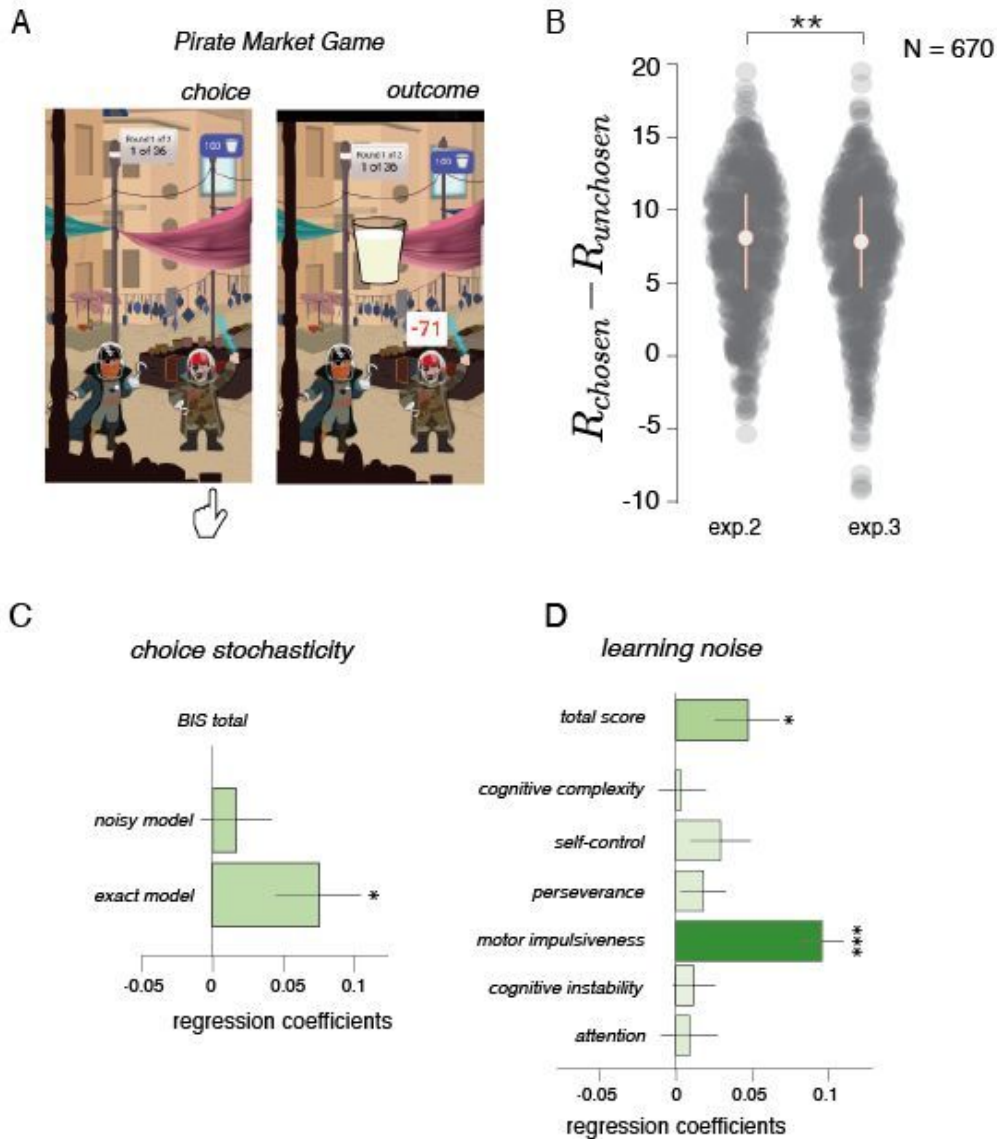
Figure 4

*Contributions of different subscales of impulsivity and obsessive-compulsive traits to choice stochasticity and learning noise in experiment 2 (N = 2610).*



**A.** Choice stochasticity from the noisy model was again associated with washing compulsions (replicating the first experiment), but not with hoarding or any other subscore. **B.** Learning noise was again primarily associated with motor impulsiveness, as in experiment 1. No compulsivity subscale was associated with learning noise. Error bars are standard errors, \* -  $p < 0.05$ , \*\* -  $p < 0.01$ , \*\*\* -  $p < 0.001$ . Dark bar colours indicate the replication tests based on the results from experiment 1.

Figure 5



## Figure 5

*Learning noise also linked to (motor) impulsivity in punishment learning (N = 670).*

**A.** Trial structure for the Pirate Market game in Brain Explorer app – a version of the “Milky Way” game but in the loss domain. This was framed as choosing the one of two pirates, who takes away less milk (hence, we called it ‘Pirate Market’). On every trial, users are endowed with 100 points (a full bucket of milk) and must choose a pirate that will steal less milk from them. The amount of stolen milk is presented in points (range from -99 to -1). The game setup and the outcome sequences were analogue to the Milky Way game but in the loss domain. **B.** Relative average rewards (chosen – unchosen) cumulated by the same sample of users (N = 670) in reward learning game (Milky Way, left) and in the punishment learning game (Pirate Market, right). White dots are medians, error bars are 25<sup>th</sup> and 75<sup>th</sup> percentile of the performance distributions. As in the reward learning context participants performed significantly better than chance ( $t(669)=38.8, p<0.0001$ ), demonstrating they understood the task and learned to choose the less punishing bandit. Performance between the two domains was positively associated ( $r=0.33, p<0.0001$ ), meaning that those who performed better in the reward version also performed better in the punishment version, even though they won slightly less in the punishment version ( $t(669)=3.00, p=0.003$ , two-tailed), **C.** BIS total score was associated with choice stochasticity in the exact, but not in the noisy learning model. **D.** In the better fitting noisy model, BIS total score was associated with learning noise. Amongst the subscales, motor impulsiveness again showed the strongest association. \* -  $p < 0.05$ , \*\* -  $p < 0.01$ , \*\*\* -  $p < 0.001$ .

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SkvortsovaHauserSupplInformation.docx](#)