

Value-free random exploration is linked to impulsivity

Received: 20 August 2020

Accepted: 1 July 2022

Published online: 04 August 2022

 Check for updatesMagda Dubois ^{1,2} ✉ & Tobias U. Hauser ^{1,2}

Deciding whether to forgo a good choice in favour of exploring a potentially more rewarding alternative is one of the most challenging arbitrations both in human reasoning and in artificial intelligence. Humans show substantial variability in their exploration, and theoretical (but only limited empirical) work has suggested that excessive exploration is a critical mechanism underlying the psychiatric dimension of impulsivity. In this registered report, we put these theories to test using large online samples, dimensional analyses, and computational modelling. Capitalising on recent advances in disentangling distinct human exploration strategies, we not only demonstrate that impulsivity is associated with a specific form of exploration—value-free random exploration—but also explore links between exploration and other psychiatric dimensions.

Protocol registration

The Stage 1 protocol for this Registered Report was accepted in principle on 19/03/2021. The protocol, as accepted by the journal, can be found at <https://doi.org/10.6084/m9.figshare.14346506.v1>.

That human and non-human animals differ in their impulsivity is one of the earliest and most influential observations of inter-individual differences¹. Impulsivity is often described as ‘acting without thinking’ and is traditionally assessed using self-report questionnaires^{1,2}. It is a broad and heterogeneous construct^{1,3–5} whose relevance not only comes from the observation of a substantial variation among a ‘healthy’ population, but also its importance in psychiatry. More recently, highly influential reinterpretations of psychiatric disorders have proposed impulsivity as an overarching symptom conglomerate encompassing multiple psychiatric disorders, such as addictions, manias, and—almost archetypically—attention-deficit/hyperactivity disorder⁶ (ADHD).

Despite the relevance of impulsivity, relatively little is known about the neurocomputational mechanisms that underlie this trait. Impulsivity has been linked to imbalances in catecholamine functioning^{7–13} but how these imbalances affect behaviour remains unknown. One suggestion about the function and role of impulsivity is as elegant as it remains speculative. Using simulations, Williams and Taylor¹⁴ suggested that impulsivity is characterised by heightened exploration behaviour, meaning that impulsive participants are more

likely to forego certain high valued outcomes for the benefit of exploring lesser known choice options that may hide even higher valued outcomes. Even though such a behaviour could be detrimental for an impulsive individual, the authors demonstrated that such behaviour could be of great benefit on a societal level¹⁴. Several theoretical accounts have embraced this concept and demonstrated how increased exploration can arise due to catecholamine imbalance^{6,12,15–17}, how this may be implemented neurally¹², and how an excessive exploration can explain other behaviours observed with impulsivity^{12,15}, such as delay discounting^{18–20}.

Outside of theoretical work, however, empirical evidence so far is still sparse. Only the first few studies using ADHD patients²¹ or looking at ADHD symptoms in youths²² have found empirical evidence for heightened exploration in impulsivity using computational methods. Their insights are particularly limited as recent work on exploration-exploitation trade-offs has shown that exploration itself is not a homogenous concept. Empirical work in healthy participants has clearly demonstrated that humans deploy multiple different exploration strategies and that these strategies differ in sophistication and

¹Max Planck UCL Centre for Computational Psychiatry and Ageing Research, London, UK. ²Wellcome Centre for Human Neuroimaging, University College London, London, UK. ✉e-mail: magda.dubois.18@ucl.ac.uk

computational demand^{23–25}. In particular, one can distinguish between sophisticated and complex exploration strategies, such as upper confidence bound²⁶ (UCB), which take the expectation as well as the uncertainty of all possible choice options into account^{25,27–30}, versus heuristic strategies, which require relatively less computation. Amongst the latter, we and others have found evidence for novelty exploration, a strategy that focuses only on completely unknown choice options^{23,29,31}. In addition, there is evidence for exploration strategies that deliberately omit all existing knowledge to choose all options equally likely, termed value-free random exploration²³. Such ‘value-free’ random exploration ignores all available information (i.e., expectation and uncertainty of choices) and thus forgoes any costly computation. This is in contrast to more refined ‘value-based’ random exploration that adds stochasticity during choice value computation or directed exploration, which biases choice towards information gain^{24,25}. Even though such mechanisms are suboptimal, their low computational demand has made them popular in artificial intelligence³² (i.e., ϵ -greedy) and we have found clear signatures in humans²³.

Whether exploration mechanisms (and if so, which ones) are altered in impulsivity remains unknown. However, this is of critical importance as recent animal³³ and human²³ work have demonstrated a specific role of catecholamine functioning in different forms of exploration. In particular, we have shown that only value-free random exploration is sensitive to noradrenaline (but not dopamine) functioning²³, which is a neurotransmitter that has repeatedly been suggested to be critical in impulsivity disorders^{6,7,12,17,34–38}.

In this study, we put the large body of theoretical work to test and exploited the recent advances in the exploration literature to empirically investigate the link between impulsivity and exploration. Here, we investigated impulsivity as a broad spectrum across the general population and also with respect to a more specific ADHD-related impulsivity. We used a preregistered, dimensional approach via large sample online testing to provide a clear answer. We advanced on a method that has recently proven the most promising for detecting meaningful mechanisms underlying psychiatric symptoms^{39–42}. We ran a big data dimensional study using our recently developed exploration task²³, which was designed to disentangle the exploration strategies that have been put forward in the literature, and which allowed us to provide an answer to whether exploration behaviour is linked to impulsivity. To determine not only whether, but also which, exploration mechanism predicts impulsivity, we made use of computational modelling. Supported by previous findings that impulsivity is associated to increased avoidance of mental effort⁴³ and that ADHD is associated to increased value-free random exploration²², we tested our hypothesis that it is specifically value-free random exploration (captured by our model parameter ϵ) which correlates with impulsivity measures (cf. Table 1), therefore determining which of these mechanisms is impaired in impulsive participants. In addition, our data allowed us to explore how exploration impairments may be linked to other psychiatric domains (e.g., to OCD and other avoidance of uncertainty disorders^{44,45}; to depression, anxiety and anhedonia^{46–48}) using data-driven methods.

Results

To capture different forms of exploration, we used our previously lab-validated Maggie’s Farm task²³ (cf. Fig. 1), which is essentially a 3-armed variant of the Horizon task²⁴. In this task, participants had to choose which bandit (depicted as trees) to draw a sample from (i.e., pick an apple) in order to maximise a sum of reward (represented by the apples’ size; Fig. 1a). To help them with their decision, at the beginning of each trial, participants had some information about how good each bandit was in the form of ‘initial samples’ (i.e., apples that have been picked before). Bandits carried either a lot, some, or no prior information (i.e., 3, 1 or 0 initial samples) and had either a

standard or a low reward mean. In effect, there were 4 different types of bandits: the certain-standard bandit (standard mean, 3 initial samples), the standard bandit (standard mean, 1 initial sample), the novel bandit (standard mean, 0 initial samples) and the low-value bandit (low mean). A real-life example would be having to choose between four different ice-cream flavours in an Italian city: chocolate, which you have enjoyed 3 times in the past, Toblerone, which you have enjoyed once in the past, hibiscus, which you have never tried, and spinach, which you have disliked once in the past. The decision horizon (cf. below), represents how often you will come back to this exact same ice-cream shop (e.g., the number of vacation days left, assuming you have once ice-cream per day). On each trial, 3 out of those 4 bandit types were used. In the analysis, the bandit with the highest mean reward of prior samples (either 1 or 3) is referred to as the ‘high-value bandit’.

This task allowed to distinguish between complex exploration strategies and exploration heuristics, namely, value-free random exploration and novelty exploration (cf. Methods for detail). We manipulated the number of prior samples and the rewards of the bandits. This allowed us to capture complex exploration strategies, because they take expected values and the uncertainty of the expected values into account. Value-free random exploration is a computationally very light heuristic that does not take any prior knowledge into account, de facto choosing randomly between options, even those known to be bad (e.g., associated to a low reward prior sample). The low-value bandit is thus a signature of such a heuristic and therefore allows quantification of its contribution. Similarly, the novel bandit allows us to capture novelty exploration, a heuristic which targets entirely novel options.

To promote and assess exploration, we manipulated the number of choices per trial (i.e., decision horizon; Fig. 1b), similarly to the Horizon Task²⁴. Participants could perform either one draw, encouraging exploitation (short horizon condition), or six draws, encouraging more substantial explorative behaviour (long horizon condition) as in the latter condition, the newly gained information could be subsequently exploited. Going back to the ice-cream example, knowing that you will come back to the same place many times will encourage you to explore different flavours (i.e., other than chocolate), as it can help guide your future choices. In the analysis, if not stated otherwise, we compared the short horizon’s single draw to the long horizon’s first draw in alignment with previous studies using the same manipulation^{23,24}. All tests were two-tailed. Detailed statistics for all measures can be found in Supplementary Table 2.

Step 1.1. Are exploitation and exploration horizon-modulated? Yes

To assess whether the horizon manipulation promoted exploration, we analysed which bandit participants chose in the long (versus short) horizon condition. We found, as hypothesised, that participants chose bandits with a lower expected value (computed as the mean of the bandits’ initial samples) in the long horizon compared to the short horizon, a sign of increased exploration in the condition where they could benefit from it (expected value of chosen bandit: Wilcoxon signed-rank two-tailed test: $V = 110057, p < 0.001$, Wilcoxon effect size: $r = 0.265$; Supplementary Fig. 3a). Further analysis revealed that this was driven by multiple behavioural shifts. We found a reduced frequency of picking the high-value bandit in the long horizon ($V = 157079.5, p < 0.001, r = 0.797$; Fig. 2a; Hypothesis 1.1.a. in Table 1), showing that participants forego the option with the best expected outcome. We found that this exploration was goal-directed, with participants choosing bandits they knew less about (lower number of initial samples, i.e., more informative) in the long horizon (number of initial samples of chosen bandit: $V = 160109.5, p < 0.001, r = 0.796$; Supplementary Fig. 3b). Concretely, they increasingly chose both the low-value bandit ($V = 34420, p < 0.001$,

Table 1 | Design Table

Question	Hypothesis	Sampling plan (e.g., power analysis)	Analysis Plan	Interpretation given to different outcomes	Outcome
Analysis Step 1: Are exploration and exploitation horizon-modulated? task effect replication	1.1.a: Less exploitation in the long horizon: High-value bandit frequency: SH > LH 1.1.b: More value-free random exploration in the long horizon: Low value bandit frequency: SH < LH 1.1.c: More novelty exploration in the long horizon: Novel bandit frequency: SH < LH	Pilot data lowest effect size: 0.410. To detect with 95% power, a similar effect size requires a sample of N = 83.	Paired samples t-test (or Wilcoxon signed-rank test if the Shapiro normality assumption is violated).	1.1.a: No effect: no evidence that our horizon manipulation modulated overall exploitation. Opposite effect: overall exploitation is increased in the long horizon. 1.1.b: No effect: no evidence that our horizon manipulation modulated value-free random exploration. Opposite effect: value-free random exploration is increased in the short horizon. 1.1.c: No effect: no evidence that our horizon manipulation modulated novelty exploration. Opposite effect: novelty exploration is increased in the short horizon.	1.1.a: Hypothesis confirmed: Less exploitation in the long horizon. 1.1.b: Hypothesis confirmed: More value-free random exploration in the long horizon. 1.1.c: Hypothesis confirmed: More novelty exploration in the long horizon. Answer to research question: Yes, exploration and exploitation are horizon-modulated.
Is exploration beneficial for participants?	1.2.a: Lower initial reward in the long horizon: Reward of first sample: SH > LH 1.2.b: Higher reward overall in the long horizon: Reward averaged over samples: SH < LH	Pilot data lowest effect size: 0.835. To be detected with 95% power, a similar effect size requires a sample of N = 21.		1.2.a: No effect: no evidence that participants sacrificed a higher initial outcome in the long horizon. Opposite effect: participants optimised initial reward in the long horizon. 1.2.b: No effect: no evidence that participants took advantage of the information gained in the long horizon. Opposite effect: information gain negatively impacted reward.	1.2.a: Hypothesis confirmed: Lower initial reward in the long horizon. 1.2.b: Hypothesis confirmed: Higher reward overall in the long horizon. Answer to research question: Yes, exploration is beneficial for participants.
Do participants use exploration heuristics?	1.3: Average BIC score: complex model + ϵ + η > other models	Pilot data effect size: 1.304. To be detected with 95% power, a similar effect size requires a sample of N = 10.		1.3: Hypothesis confirmed: The average BIC score was higher for the complex models with both heuristics compared to other models. Opposite effect: Participants are not using complex models with both heuristics.	1.3: Hypothesis confirmed: The average BIC score was higher for the complex models with both heuristics compared to other models. Answer to research question: Yes, participants use exploration heuristics.
Are exploration heuristics used more in the long horizon?	1.4.a: Value-free random exploration is used more in the long horizon: ϵ -greedy parameter: SH < LH 1.4.b Novelty exploration is used more in the long horizon: Novelty bonus η : SH < LH	Pilot data lowest effect size: 0.446. To be detected with 95% power, a similar effect size requires a sample of N = 71.		1.4.a: No effect: no evidence that ϵ was modulated by the horizon. Opposite effect: value-free random exploration is increased in the short horizon. 1.4.b: No effect: no evidence that η was modulated by the horizon. Opposite effect: novelty exploration is increased in the short horizon.	1.4.a: Hypothesis confirmed: Value-free random exploration is used more in the long horizon. 1.4.b: Hypothesis confirmed: Novelty exploration is used more in the long horizon. Answer to research question: Yes, exploration heuristics are used more in the long horizon.
Analysis Step 2: Is impulsivity linked to value-free random exploration? impulsivity	2.1: Value-free random exploration is positively associated to BIS: ϵ -greedy parameter and low-value bandit frequency correlates positively with the BIS total score. 2.2: Value-free random exploration is positively associated to ASRS: ϵ -greedy parameter and low-value bandit frequency correlates positively with the ASRS total score.	Previous study ²⁷ correlation for a similar measure: R = 0.26. To be detected with 95% power, a similar effect size requires a sample of N = 190.	Bivariate and partial (correcting for age and IQ) Pearson correlation and a repeated-measures ANOVA with within factor horizon and between participants variable [Impulsivity/ADHD-symptoms]	2.1: No effect: no evidence for an association between value-free random exploration and general impulsivity as measured by the BIS. Opposite effect: Value-free random exploration is negatively associated to BIS. 2.2: No effect: no evidence for an association between value-free random exploration and ADHD as measured by the ASRS total score. Opposite effect: Value-free random exploration is negatively associated to ASRS total score.	2.1: Hypothesis confirmed: Value-free random exploration is positively associated to BIS. Answer to research question: Yes, impulsivity is linked to value-free random exploration. 2.2: Hypothesis confirmed: Value-free random exploration is positively associated to ASRS. Answer to research question: Yes, ADHD symptoms are linked to value-free random exploration.
Are ADHD symptoms linked to value-free random exploration?	2.2: Value-free random exploration is positively associated to ASRS: ϵ -greedy parameter and low-value bandit frequency correlates positively with the ASRS total score.				

Summary of preregistered hypotheses from our Stage 1 Registered report (the full protocol can be found at: <https://doi.org/10.6084/m9.figshare.14346506.v1>) with an additional 'outcome' column describing the observed effect. SH: Short horizon condition, LH: Long horizon condition.

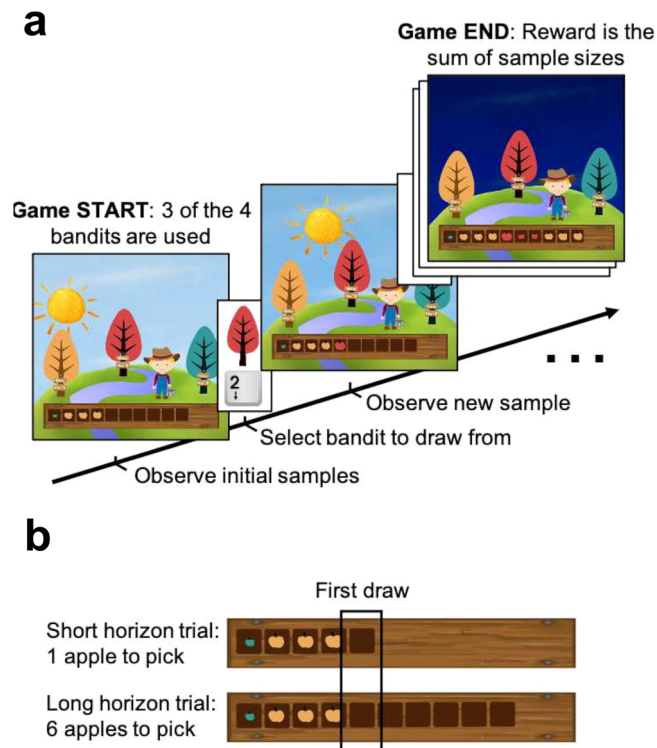


Fig. 1 | Exploration task. In the Maggie's farm task, participants had to choose from three bandits (depicted as trees) to maximise their overall reward. The rewards (apple size) of each bandit followed a normal distribution with a fixed sampling variance. **a** At the beginning of each trial, participants are provided with some initial samples (number varied depending on the bandits present on that trial) on the wooden crate at the bottom of the screen and participants had to select which bandit they want to sample from next. **b** Depending the condition, they can either perform one draw (short horizon) or six draws (long horizon). The empty spaces on the wooden crate (and the position of the sun) indicate how many draws they have left. Image adapted from our previous study²³.

$r = 0.425$; Fig. 2a; Hypothesis 1.1.b. in Table 1) as well as the novel bandit ($V = 10355$, $p < 0.001$, $r = 0.750$; Fig. 2a; Hypothesis 1.1.c. in Table 1), for which there were no initial information available. Our findings thus match our preregistered hypotheses of an increase in exploration in the long horizon.

Step 1.2. Is exploration beneficial for participants? Yes

To evaluate whether participants were able to use their exploration beneficially, we looked at their performance (i.e., the outcomes they obtained). In alignment with the above analyses, we observed that participants obtained a lower reward (i.e., apple size) in the first draw of the long horizon (i.e., when we observed increased exploration) compared to the single draw in the short horizon ($V = 131612$, $p < 0.001$, $r = 0.53$; Supplementary Fig. 3c; Hypothesis 1.2.a. in Table 1). To assess the long-term benefits of exploration, we calculated the long horizon average reward (across 6 draws) and found that this was higher than the short horizon reward ($V = 264$, $p < 0.001$, $r = 0.864$; Supplementary Fig. 3c; Hypothesis 1.2.b. in Table 1). This indicated that participants made good use of the additional information earned by exploring as observed in previous studies^{22,23} in alignment with our preregistered hypotheses. For an analysis of score per trial and per block cf. Supplementary Fig. 4.

Step 1.3. Do participants use exploration heuristics? Yes

To assess more formally which exploration strategies were being used, we turned to computational modelling, which allows us to tease apart different exploration strategies. In line with our previous studies^{22,23},

we found that participants used a mixture of computationally demanding (i.e., Thompson sampling) and two heuristic exploration strategies (i.e., value-free random exploration ϵ and novelty exploration η) as captured by the winning model (comparison of BIC average scores: Thompson+ η + ϵ vs Thompson model: $V = 3089$, $p < 0.001$, $r = 0.835$; Hypothesis 1.3. in Table 1; Thompson+ η + ϵ vs UCB+ η + ϵ model: $V = 46440$, $p < 0.001$, $r = 0.389$; Supplementary Fig. 5a). The pilot data (cf. Supplementary Information) and our preregistered hypotheses (cf. Table 1) predicted the same winning model.

Step 1.4. Are exploration heuristics used more in the long horizon? Yes

Next, we were interested to assess which exploration strategies were deployed more in the long horizon, which is why we examined the winning model's (Thompson+ η + ϵ) fitted parameters. We found an increase in the ϵ -greedy parameter in the long horizon, which captures the contribution of value-free random exploration ($V = 35367$, $p < 0.001$, $r = 0.503$; Fig. 2b; Hypothesis 1.4.a. in Table 1). Similarly, the novelty bonus η , which captures the intrinsic reward of selecting a novel option, was also increased in the long horizon ($V = 10334$, $p < 0.001$, $r = 0.76$; Fig. 2b; Hypothesis 1.4.b. in Table 1). This thus confirms our preregistered hypothesis of a flexible deployment of these exploration heuristics (cf. Table 1). In addition, we found that the prior variance, capturing complex, uncertainty-related exploration, was also increased in the long horizon (prior variance fitted parameter: $V = 54537$, $p < 0.001$, $r = 0.306$; Fig. 2b), which supports the notion that the long horizon facilitates the exploration strategies we assessed in this task.

Step 2.1. Is impulsivity linked to value-free random exploration? Yes

Next, we looked at the link between impulsivity and exploration. First, we characterised general impulsivity as a broad concept, and expected it to be linked with value-free random exploration. For this, we used the total score on the Barratt Impulsiveness Scale (BIS), the most commonly administered self-report measure for impulsiveness⁴⁹. We assessed its link to the model parameter and behavioural measure of value-free random exploration, the ϵ -greedy parameter and the low-value bandit picking frequency. As hypothesized (cf. Table 1), we found a significant association between the BIS total score and the ϵ -greedy parameter ($r(578) = 0.171$, $p < 0.001$, Fig. 3a; accounting for age and IQ: $r(573) = 0.117$, $p = 0.005$; Hypothesis 2.1. in Table 1; cf. Methods for details and Supplementary Table 1 for demographics), which was also reflected by a correlation between the BIS total score and the low-value bandit frequency ($r(578) = 0.174$, $p < 0.001$, Fig. 3b; accounting for age and IQ: $r(573) = 0.117$, $p = 0.005$; Hypothesis 2.1. in Table 1). In line with these results, when performing a repeated-measures ANOVA with the horizon as within-participant factor, we found a main effect of impulsivity on how frequently the low-value bandit was chosen (BIS main effect: $F(1,578) = 18.103$, $p < 0.001$, partial eta squared $\eta_p^2 = 0.03$; horizon main effect: $F(1,578) = 113.614$, $p < 0.001$, $\eta_p^2 = 0.164$; BIS-by-horizon interaction: $F(1,578) = 0.773$, $p = 0.380$, $\eta_p^2 = 0.001$) and on the ϵ -greedy parameter (BIS main effect: $F(1,578) = 17.454$, $p < 0.001$, $\eta_p^2 = 0.029$; horizon main effect: $F(1,578) = 125.804$, $p < 0.001$, $\eta_p^2 = 0.179$; BIS-by-horizon interaction: $F(1,578) = 0.084$, $p = 0.772$, $\eta_p^2 < 0.001$), but no significant interaction effects, suggesting that this exploration strategy was increased in both horizons.

In summary, these findings confirmed our preregistered hypothesis that value-free random exploration is linked to general impulsivity traits in this large convenience sample, and our exploratory analyses (cf. Step 4. Exploratory analyses) showed that it was not associated with any other exploration strategy. Detailed correlations can be found in Supplementary Table 5. Detailed correlations with all questionnaires can be found in Supplementary Table 10 and Supplementary Table 11.

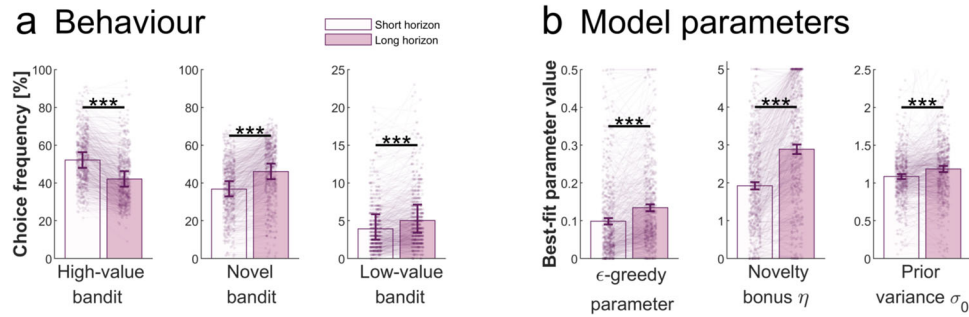


Fig. 2 | Increased exploration in the long horizon. **a** Behavioural horizon effects: in the long (versus short) horizon participants sampled less from the high-value bandit (two-sided Wilcoxon signed-rank two-tailed test: $V = 157079.5$, $p = 1.536e-81$, Wilcoxon effect size $r = 0.797$) and more from the novel ($V = 10355$, $p = 1.811e-72$, $r = 0.750$) and low-value bandit ($V = 34420$, $p = 2.817e-24$, $r = 0.425$). **b** Model parameters: in the long (versus short) horizon participants had higher value of ϵ (i.e., value-free random exploration; $V = 35367$, $p = 9.831e-34$, $r = 0.503$), η (i.e., novelty exploration; $V = 10334$, $p = 7.411e-75$, $r = 0.076$) and σ_0 (their uncertainty about a bandit's mean before seeing any samples; $V = 54537$, $p = 1.868e-13$, $r = 0.306$). The

parameters were fitted to each participant's first draw, and they were fitted to each horizon separately. $***p < 0.001$. For detailed statistics cf. Supplementary Table 2. For details about model parameters cf. Supplementary Table 4. Data are shown as mean \pm 95% CI and each dot/line represent one participant. Sample size for statistics: $N = 580$ human participants. Source data are provided as a Source Data file. Bar values: High-value bandit: Short Horizon (SH): 52.134, Long Horizon (LH): 42.088; Novel bandit: SH: 36.801, LH: 46.073; Low-value bandit: SH: 3.928, LH: 5.026; ϵ -greedy parameter: SH: 0.099, LH: 0.134; Novelty bonus η : SH: 1.919, LH: 2.884; Prior variance σ_0 : SH: 1.085, LH: 1.186.

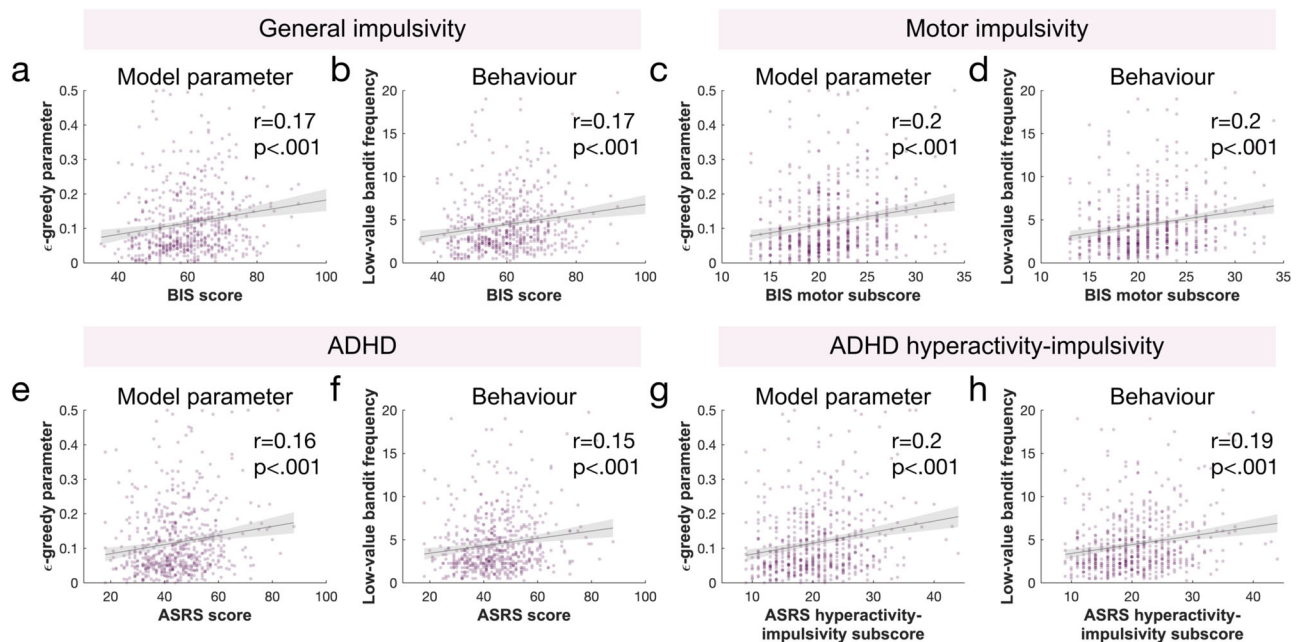


Fig. 3 | Value-free random exploration linked to impulsivity. General impulsivity (as measured by the BIS⁴⁹) was significantly associated to value-free random exploration. This was observed both in **a** the model parameter, the ϵ -greedy parameter (Pearson's correlation: $r(578) = 0.171$, $p = 3.398e-05$) and in **b** the behaviour, the frequency of picking the low-value bandit ($r(578) = 0.174$, $p = 2.442e-05$). The motor subscore of the BIS was most closely associated with value-free random exploration, both in **c** the model parameter (Bonferroni corrected ($n = 3$): $r(578) = 0.198$, $p_{\text{cor}} = 5.318e-06$, $p_{\text{unc}} = 1.772e-06$) and **d** the behaviour ($r(578) = 0.205$, $p_{\text{cor}} = 2.249e-06$, $p_{\text{unc}} = 7.495e-07$). Similarly, ADHD-

related impulsivity (as measured by the ASRS⁷²) was significantly associated to value-free random exploration, both in **e** the model parameter ($r(578) = 0.157$, $p = 1.466e-04$) and in **f** the behaviour ($r(578) = 0.151$, $p = 3.069e-04$). The hyperactivity-impulsivity subscore of the ASRS was most tightly associated to novelty exploration, both in **g** the model parameter (Bonferroni corrected ($n = 2$): $r(578) = 0.205$, $p_{\text{cor}} = 1.352e-06$, $p_{\text{unc}} = 6.764e-07$) and **h** the behaviour ($r(578) = 0.193$, $p_{\text{cor}} = 6.319e-06$, $p_{\text{unc}} = 3.159e-06$). The filled lines represent the 95%CI. Sample size for statistics: $N = 580$ human participants. Source data are provided as a Source Data file.

Step 2.2. Are ADHD symptoms linked to value-free random exploration? Yes

After having established the association between value-free random exploration and general impulsivity, we sought to investigate the association more specifically, focusing on ADHD symptoms. Based on our previous preliminary findings of a positive association between ADHD traits and value-free random exploration²² in youths, we thus investigated how the ASRS total score is linked to this form of exploration. A correlation of $r = 0.63$ (cf. Supplementary Fig. 10) between the above BIS total score and the ASRS total

score suggests that they are similar, but not entirely overlapping constructs.

As hypothesized in our preregistration (cf. Table 1), we found an association between the ASRS total score and the ϵ -greedy parameter ($r(578) = 0.157$, $p_{\text{unc}} < 0.001$, Fig. 3e; accounting for age and IQ: $r(573) = 0.115$, $p_{\text{unc}} = 0.006$; Hypothesis 2.2. in Table 1). Likewise, the same effect was present when looking at the association with its behavioural equivalent, the low-value bandit frequency ($r(578) = 0.151$, $p_{\text{unc}} < 0.001$, Fig. 3f; accounting for age and IQ: $r(573) = 0.104$, $p_{\text{unc}} = 0.012$; Hypothesis 2.2. in Table 1).

Similar to the above findings, we did not find any interaction with horizon, neither in the low-value bandit (ASRS main effect: $F(1,578) = 13.187$, $p < 0.001$, $\eta_p^2 = 0.022$; horizon main effect: $F(1,578) = 113.468$, $p < 0.001$, $\eta_p^2 = 0.164$; ASRS-by-horizon interaction: $F(1,578) = 0.025$, $p = 0.875$, $\eta_p^2 < 0.001$) nor the ϵ -greedy parameter (ASRS main effect: $F(1,578) = 14.609$, $p < 0.001$, $\eta_p^2 = 0.025$; horizon main effect: $F(1,578) = 126.34$, $p < 0.001$, $\eta_p^2 = 0.179$; ASRS-by-horizon interaction: $F(1,578) = 2.549$, $p = 0.111$, $\eta_p^2 = 0.004$). Our exploratory analyses (cf. Step 4. Exploratory analyses) showed that it was not associated with any other exploration strategy. These results thus confirmed our preregistered hypothesis that value-free random exploration is linked to ADHD symptoms. Detailed correlations can be found in Supplementary Table 5.

Step 3. Preregistered exploratory analyses

Step 3.1. Investigating subscales of impulsivity and ADHD

To explore the association between value-free random exploration and general impulsivity further, we performed an exploratory analysis of the BIS subscores (i.e., attentional, motor, and non-planning impulsivity), correcting for multiple comparisons using Bonferroni correction ($N = 3$). Those subscores allow to differentiate between attentional impulsiveness, an ‘inability to focus attention or concentrate’, motor impulsiveness, ‘acting without thinking’, and non-planning impulsiveness, a lack of ‘futures’ or ‘forethought’⁴⁹.

We found that the BIS motor subscore was associated with value-free random exploration in all indicators of that exploration heuristic (Bonferroni corrected ($n = 3$): ϵ -greedy parameter: $r(578) = 0.198$, $p_{\text{cor}} < 0.001$, $p_{\text{unc}} < 0.001$, Fig. 3c [accounting for age and IQ: $r(573) = 0.159$, $p_{\text{cor}} < 0.001$, $p_{\text{unc}} < 0.001$]; frequency of low-value bandit: $r(578) = 0.205$, $p_{\text{cor}} < 0.001$, $p_{\text{unc}} < 0.001$, Fig. 3d [accounting for age and IQ: $r(573) = 0.165$, $p_{\text{cor}} < 0.001$, $p_{\text{unc}} < 0.001$]). We did not observe any robust association with the BIS non-planning subscore when correcting for age and IQ (using the ϵ -greedy parameter: $r(578) = 0.120$, $p_{\text{cor}} = 0.012$, $p_{\text{unc}} = 0.004$ [accounting for age and IQ: $r(573) = 0.058$, $p_{\text{cor}} = 0.501$, $p_{\text{unc}} = 0.167$]; using the low-value bandit: $r(578) = 0.128$, $p_{\text{cor}} = 0.006$, $p_{\text{unc}} = 0.002$ [accounting for age and IQ: $r(573) = 0.065$, $p_{\text{cor}} = 0.364$, $p_{\text{unc}} = 0.121$]) or with the BIS attentional subscore (using the ϵ -greedy parameter: $r = 0.095$, $p_{\text{cor}} = 0.067$, $p_{\text{unc}} = 0.022$ [accounting for age and IQ: $r(573) = 0.067$, $p_{\text{cor}} = 0.331$, $p_{\text{unc}} = 0.11$]; using the low-value bandit: $r(578) = 0.086$, $p_{\text{cor}} = 0.118$, $p_{\text{unc}} = 0.039$ [accounting for age and IQ: $r(573) = 0.054$, $p_{\text{cor}} = 0.592$, $p_{\text{unc}} = 0.197$]). This suggests that it is the motor dimension of general impulsivity, i.e., acting without thinking, that is related to value-free random exploration the most. Detailed correlations can be found in Supplementary Table 6.

Next, we further explored how the value-free random exploration is associated with the two ADHD subdomains (as assessed by the ASRS), namely inattention and hyperactivity-impulsivity (Bonferroni correcting for $N = 2$ tests). We found that value-free random exploration was linked to the hyperactivity-impulsivity subscore (using the ϵ -greedy parameter: $r(578) = 0.205$, $p_{\text{cor}} < 0.001$, $p_{\text{unc}} < 0.001$, Fig. 3g [accounting for age and IQ: $r(573) = 0.152$, $p_{\text{cor}} = 0.001$, $p_{\text{unc}} < 0.001$]; using the low-value bandit: $r(578) = 0.193$, $p_{\text{cor}} < 0.001$, $p_{\text{unc}} < 0.001$, Fig. 3h [accounting for age and IQ: $r(573) = 0.136$, $p_{\text{cor}} = 0.002$, $p_{\text{cor}} = 0.001$]) but not reliably with the ASRS inattention subscore (using the ϵ -greedy parameter: $r(578) = 0.087$, $p_{\text{cor}} = 0.074$, $p_{\text{unc}} = 0.037$ [accounting for age and IQ: $r(573) = 0.061$, $p_{\text{cor}} = 0.292$, $p_{\text{unc}} = 0.146$]; using the low-value bandit: $r(578) = 0.085$, $p_{\text{cor}} = 0.082$, $p_{\text{unc}} = 0.041$ [accounting for age and IQ: $r(573) = 0.057$, $p_{\text{cor}} = 0.348$, $p_{\text{unc}} = 0.174$]). This suggests that value-free random exploration is more closely linked to the impulsivity-hyperactivity dimension of ADHD than the other subdomains. Detailed correlations can be found in Supplementary Table 6.

Step 3.2. Investigating exploration across transdiagnostic dimensions

Thus far, we exclusively focused on our hypothesised association between exploration and impulsivity / ADHD symptoms. However, to be able to explore the wider associations with other symptom dimensions, we additionally collected data from further questionnaires, in the same spirit as previous transdiagnostic dimensional approaches^{39,41,42}.

As specified in our preregistration, we conducted a factor analysis across all items of the collected questionnaires (including BIS and ASRS). This factor analysis of individual questionnaire items revealed three distinct latent factors (Fig. 4a) which we labelled as ‘anxious-depression’, ‘uncertainty-related distress’ and ‘impulsivity’ factor, in accordance with the strongest individual item loadings (cf. Fig. 4b). For correlations between questionnaires and factors cf. Supplementary Fig. 10.

As we had initially expected, our two impulsivity-related questionnaires (BIS and ASRS) primarily loaded onto one factor (labelled as impulsivity factor). We thus explored the association between this impulsivity factor and value-free random exploration. We found that value-free random exploration was more closely related with the impulsivity factor than with each questionnaire separately (i.e., BIS and ASRS). We found an association between the impulsivity factor and the ϵ -greedy parameter (correcting for multiple comparison using Bonferroni correction across 4 parameters \times 3 factors, i.e., $N = 12$; $r(578) = 0.257$, $p_{\text{unc}} < 0.001$, $p_{\text{cor}} < 0.001$, Fig. 5a; accounting for age and IQ: $r(573) = 0.204$, $p_{\text{unc}} < 0.001$, $p_{\text{cor}} < 0.001$) as well as between the impulsivity factor score and the low-value bandit frequency (correcting for multiple comparison using Bonferroni correction across 3 bandits \times 3 factors, i.e., $N = 9$; $r(578) = 0.247$, $p_{\text{unc}} < 0.001$, $p_{\text{cor}} < 0.001$, Fig. 5b; accounting for age and IQ: $r(573) = 0.191$, $p_{\text{unc}} < 0.001$, $p_{\text{cor}} < 0.001$). Together, our results suggest that value-free random exploration is associated with a general impulsivity that spans across multiple questionnaires.

Having established the link with value-free random exploration, we now explored whether impulsivity was also linked to other forms of exploration. When linking the impulsivity factor with the parameters capturing the other exploration strategies, we did not observe any significant correlation (Bonferroni correction with $N = 12$), neither with the novelty bonus η ($r(578) = 0.051$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.223$; accounting for age and IQ: $r(573) = 0.058$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.167$; Fig. 5e), nor the prior variance σ_0 ($r(578) = -0.02$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.631$; accounting for age and IQ: $r(573) = 0.01$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.816$; Fig. 5e), or the prior mean Q_0 ($r(578) = -0.006$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.891$; accounting for age and IQ: $r(573) = -0.016$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.701$; Fig. 5e). This suggests that the impulsivity is first and foremost linked with value-free random exploration.

As a second step, we explored whether exploration correlates with the other factors identified in the factor analysis. Similar to previous studies^{40–42,50,51}, we retrieved a factor, labelled anxious-depression, which was mainly capturing depression, social anxiety and trait anxiety questions (SDS, LSAS and STAI-Y2 questionnaires respectively). As for the third factor, we obtained a factor that was mainly capturing intolerance of uncertainty (IUS questionnaire), labelled as uncertainty-related distress.

First, we looked at the anxious-depression factor and all exploration strategies as captured by the model parameters (correcting for multiple comparison using Bonferroni correction across all parameters \times factors, i.e., $N = 12$). Our exploratory analysis revealed that the anxious-depression factor correlated positively with the novelty bonus η ($r(578) = 0.14$, $p_{\text{cor}} = 0.008$, $p_{\text{unc}} < 0.001$, Fig. 5c; accounting for age and IQ: $r(573) = 0.126$, $p_{\text{cor}} = 0.03$, $p_{\text{unc}} = 0.002$). None of the other parameters was linked to the anxious-depression factor (ϵ : $r(578) = -0.047$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.262$; accounting for age and IQ: $r(573) = -0.078$, $p_{\text{cor}} = 0.73$,

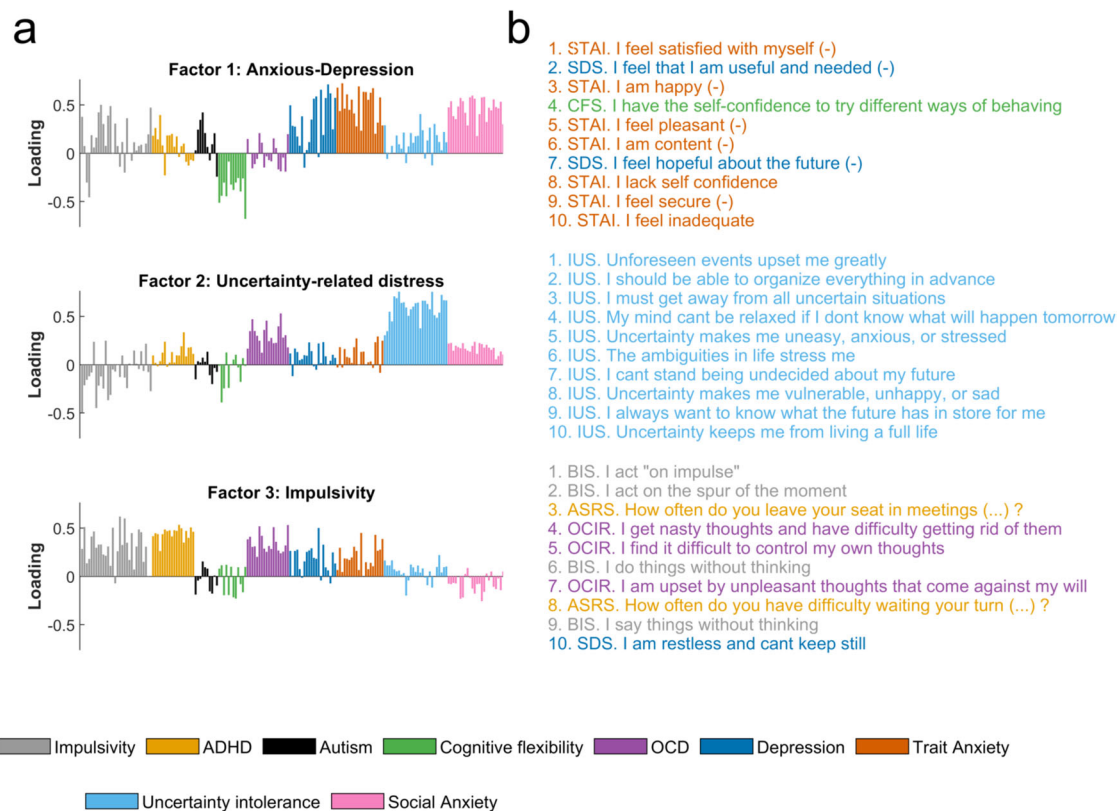


Fig. 4 | Transdiagnostic parcellation of symptoms. **a** Three latent factors were identified when performing a factor analysis on individual questionnaire items. **b** The 10 most loading items for each factor illustrate the aspects that contribute to each dimension. Abbreviations: ASRS: Adult ADHD Self-Report Scale, BIS: Barratt Impulsiveness Scale, LSAS: Liebowitz Social Anxiety Scale, STAI: State-Trait Anxiety Inventory, IUS: Intolerance of Uncertainty, OCIR:

Obsessive-Compulsive Inventory-Revised, SDS: Zung's Self-rating Depression Scale, CFS: Cognitive Flexibility Scale, AQ10: Autism spectrum Quotient. (-) indicates reversed items. For correlations between each questionnaire total score cf. Supplementary Fig. 10. For loading weights of individual items on each factor cf. Supplementary Table 10–18. Source data are provided as a Source Data file.

$p_{\text{unc}} = 0.061$; σ_0 : $r = 0.099$, $p_{\text{cor}} = 0.203$, $p_{\text{unc}} = 0.017$; accounting for age and IQ: $r(573) = 0.107$, $p_{\text{cor}} = 0.121$, $p_{\text{unc}} = 0.01$; Q_0 : $r(578) = 0.073$, $p_{\text{cor}} = 0.925$, $p_{\text{unc}} = 0.077$; accounting for age and IQ: $r(573) = 0.052$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.209$; Fig. 5e).

This pattern was also reflected in the behavioural indicators when looking at the correlation between the 3 factors and the bandit picking frequencies. Correcting for multiple comparisons (Bonferroni correction with $N = 9$), we found that with the anxious-depression factor, the novel bandit frequency was increased ($r(578) = 0.19$, $p_{\text{cor}} < 0.001$, $p_{\text{unc}} < 0.001$, Fig. 5d; accounting for age and IQ: $r(573) = 0.17$, $p_{\text{cor}} < 0.001$, $p_{\text{unc}} < 0.001$), and in turn the high-value bandit frequency decreased ($r(578) = -0.174$, $p_{\text{cor}} < 0.001$, $p_{\text{unc}} < 0.001$; accounting for age and IQ: $r(573) = -0.138$, $p_{\text{cor}} = 0.008$, $p_{\text{unc}} = 0.001$; Fig. 5f). We did not observe any correlation with the low-value bandit frequency ($r(578) = -0.058$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.165$; accounting for age and IQ: $r(573) = -0.093$, $p_{\text{cor}} = 0.232$, $p_{\text{unc}} = 0.026$; Fig. 5f). Together our results demonstrate that the anxious-depression factor is associated with an increase in novelty exploration.

Lastly, we explored whether the uncertainty-related distress factor was associated with any exploration strategy. We did not observe any significant association (after correcting for multiple comparisons) in neither in the model parameters (ϵ : $r(578) = 0.107$, $p_{\text{cor}} = 0.119$, $p_{\text{unc}} = 0.01$; accounting for age and IQ: $r(573) = 0.072$, $p_{\text{cor}} = 0.997$, $p_{\text{unc}} = 0.083$; η : $r(578) = 0.001$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.99$; accounting for age and IQ: $r(573) = -0.002$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.97$; σ_0 : $r(578) = -0.006$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.877$; accounting for age and IQ: $r(573) = 0.009$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.821$; Q_0 : $r(578) = 0.054$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.197$; accounting for age and IQ: $r(573) = 0.048$, $p_{\text{cor}} = 1$,

$p_{\text{unc}} = 0.251$; Fig. 5e) nor in the behaviour (low-value bandit: $r = 0.075$, $p_{\text{cor}} = 0.625$, $p_{\text{unc}} = 0.069$; accounting for age and IQ: $r(573) = 0.035$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.399$; novel bandit: $r(578) = 0.039$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.343$; accounting for age and IQ: $r(573) = 0.037$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.371$; high-value bandit: $r(578) = -0.09$, $p_{\text{cor}} = 0.266$, $p_{\text{unc}} = 0.03$; accounting for age and IQ: $r(573) = -0.073$, $p_{\text{cor}} = 0.732$, $p_{\text{unc}} = 0.081$; Fig. 5f).

Taken together, these findings thus suggest that – as hypothesized – impulsivity is associated with value-free random exploration. In addition, we also find a non-hypothesised association between the novelty exploration heuristic and an anxious-depression factor. Detailed correlations can be found in Supplementary Table 7.

Step 3.3. Associations with cognitive flexibility and autism

We did not observe any correlation between autism and value-free random exploration (the AQ10⁵² total score with the low-value bandit frequency: $r(578) = 0.025$, $p = 0.545$; accounting for age and IQ: $r(573) = 0.022$, $p = 0.592$; with the ϵ -greedy parameter: $r(578) = 0.023$, $p = 0.584$; accounting for age and IQ: $r(573) = 0.02$, $p = 0.631$), nor between cognitive flexibility and value-free random exploration (the CFS⁵³ total score with the low-value bandit frequency: $r(578) = -0.042$, $p = 0.317$; accounting for age and IQ: $r(573) = 0.002$, $p = 0.954$; with the ϵ -greedy parameter: $r(578) = -0.038$, $p = 0.361$; accounting for age and IQ: $r(573) = 0.004$, $p = 0.927$).

Step 4. Non-preregistered exploratory analyses

The analyses mentioned below were not part of the preregistration.

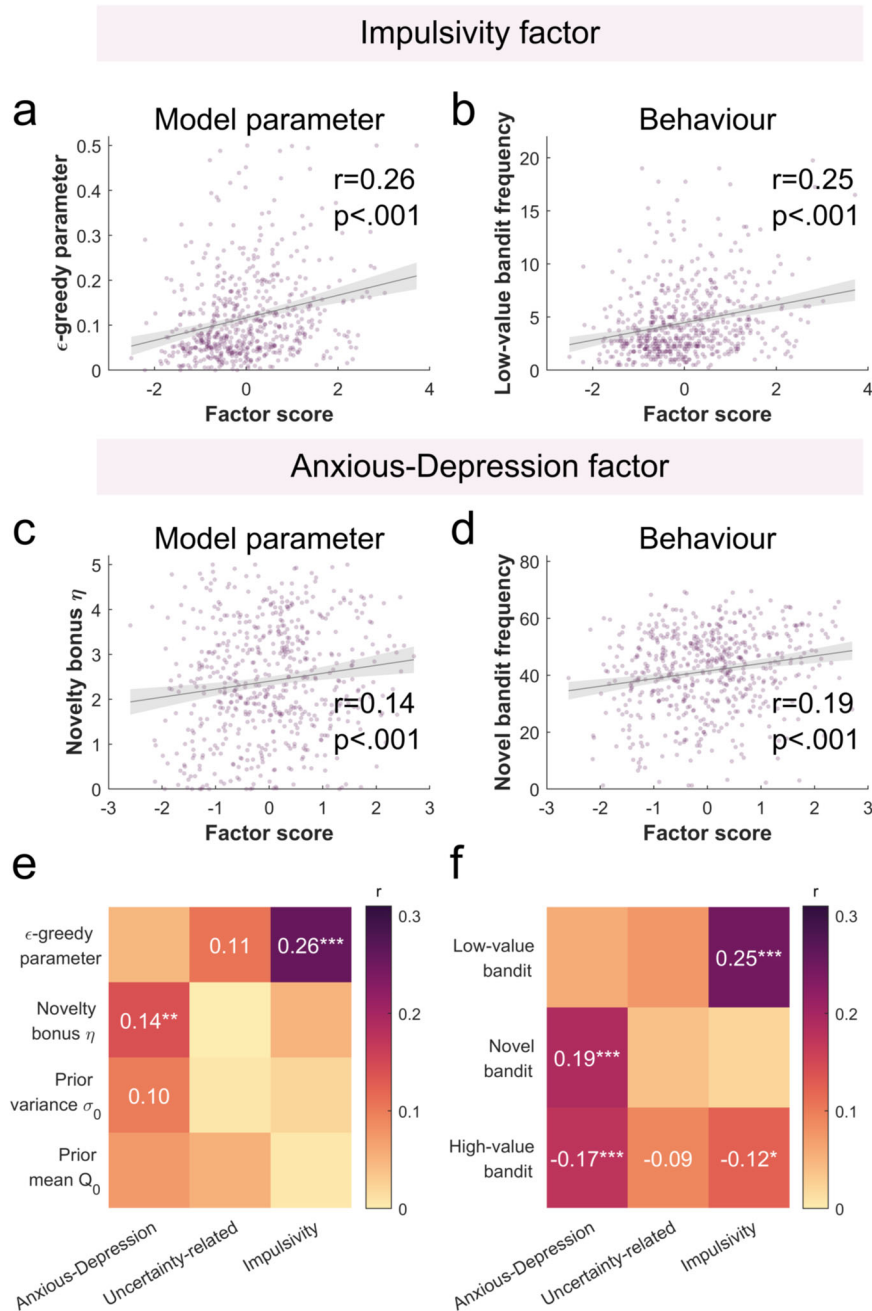


Fig. 5 | Exploration associations with transdiagnostic psychiatric factors. The factor analysis-derived impulsivity factor was significantly associated to value-free random exploration. This was observed both in **a** the model parameter, the ϵ -greedy parameter (Bonferroni corrected ($n=12$); Pearson's correlation: $r(578)=0.257$, $p_{\text{unc}}=3.352\text{e-}10$, $p_{\text{cor}}=4.023\text{e-}09$) and in **b** the behaviour, the frequency of picking the low-value bandit (Bonferroni corrected ($n=9$); $r(578)=0.257$, $p_{\text{unc}}=1.561\text{e-}09$, $p_{\text{cor}}=1.405\text{e-}08$). Similarly, the anxious-depression factor was significantly associated with the novelty exploration, both in **c** the model

parameter, the novelty bonus η (Bonferroni corrected ($n=12$): $r(578)=0.14$, $p_{\text{unc}}=7.041\text{e-}04$, $p_{\text{cor}}=0.0084$), and in **d** the behaviour (Bonferroni corrected ($n=9$): $r(578)=0.19$, $p_{\text{unc}}=4.084\text{e-}06$, $p_{\text{cor}}=3.675\text{e-}05$), the frequency of picking the novel bandit. Pearson correlations for each factor and **e** model parameters as well as **f** behavioral task measures (i.e., bandit picking frequencies). Significant uncorrected correlations are displayed. Significant corrected correlations are indicated with asterisks (* $p<0.05$, ** $p<0.01$, *** $p<0.001$). Sample size for statistics: $N=580$ human participants. Source data are provided as a Source Data file.

Step 4.1. Improved performance following exploration is not strategy-specific

To investigate whether the improved performance following exploration (cf. Supplementary Fig. 3c) was specific to an exploration strategy, we split the data by their first choice (i.e., high-value bandit, novel bandit or low-value bandit; cf. Supplementary Fig. 3d). The higher outcome (in the long run) following exploration was irrespective of the exploration strategy used (i.e., novel or low-value bandit).

Step 4.2. No association between the BIS and other exploration strategies

We explored whether the BIS score was correlated with any of the other exploration strategies (uncertainty-driven Thompson exploration, novelty exploration; correcting for multiple parameters $N=4$). We did not observe any association with any of these parameters (η : $r(578)=0.061$, corrected $p_{\text{cor}}=0.565$, uncorrected $p_{\text{unc}}=0.141$; accounting for age and IQ: $r(573)=0.073$, $p_{\text{cor}}=0.32$, $p_{\text{unc}}=0.08$; σ_0 : $r(578)=0.012$, $p_{\text{cor}}=1$, $p_{\text{unc}}=0.767$; accounting for age and IQ:

$r(573) = 0.041$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.331$; Q_0 : $r(578) = 0.029$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.491$; accounting for age and IQ: $r(573) = 0.026$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.531$), while the correlation with value-free random exploration (as detailed above) remained significant (ϵ : $r(578) = 0.171$, $p_{\text{cor}} < 0.001$, $p_{\text{unc}} < 0.001$; accounting for age and IQ: $r(573) = 0.117$, $p_{\text{cor}} = 0.019$, $p_{\text{unc}} = 0.005$).

Step 4.3. No association between the ASRS and other exploration strategies

We explored whether the ASRS score was correlated with any of the other exploration strategies (correcting for multiple parameters $N = 4$). We did not observe any robust association with any of the other parameters (η : $r(578) = 0.083$, $p_{\text{cor}} = 0.187$, $p_{\text{unc}} = 0.047$; accounting for age and IQ: $r(573) = 0.087$, $p_{\text{cor}} = 0.152$, $p_{\text{unc}} = 0.038$; σ_0 : $r(578) = 0.015$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.713$; accounting for age and IQ: $r(573) = 0.038$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.369$; Q_0 : $r(578) = 0.03$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.466$; accounting for age and IQ: $r(573) = 0.022$, $p_{\text{cor}} = 1$, $p_{\text{unc}} = 0.603$), while the correlation with value-free random exploration remained significant (ϵ : $r(578) = 0.157$, $p_{\text{cor}} = 0.001$, $p_{\text{unc}} < 0.001$; accounting for age and IQ: $r(573) = 0.115$, $p_{\text{cor}} = 0.023$, $p_{\text{unc}} = 0.006$).

Step 4.4. Analysis of 2nd winning model

Value-free random exploration (captured by the ϵ -greedy parameter) was similar in the 1st winning model (Thompson+ ϵ + η) and in the 2nd winning model (UCB+ ϵ + η), both in the short horizon (Pearson correlation: $r(578) = 0.87$, $p < 0.001$; Supplementary Fig. 7a) and in the long horizon ($r(578) = 0.85$, $p < 0.001$; Supplementary Fig. 7b). Similarly, novelty exploration (captured by the novelty bonus η) was similar across both models, both in the short horizon ($r(578) = 0.71$, $p < 0.001$; Supplementary Fig. 8a) and in the long horizon ($r(578) = 0.72$, $p < 0.001$; Supplementary Fig. 8b).

Similar to the 1st winning model, we observed an association between value-free random exploration (i.e., ϵ -greedy parameter) as captured by 2nd winning model, and impulsivity. We observed a significant association between ϵ and the BIS total score ($r(578) = 0.155$, $p < 0.001$; controlling for age and IQ: $r(573) = 0.101$, $p = 0.015$), between ϵ and the ASRS total score ($r(578) = 0.167$, $p < 0.001$; controlling for age and IQ: $r(573) = 0.126$, $p = 0.006$), between ϵ and the impulsivity factor (cf. Fig. 4; $r(578) = 0.250$, $p < 0.001$; controlling for age and IQ: $r(573) = 0.196$, $p < 0.001$).

Additionally, we observed an association between novelty exploration (i.e., novelty bonus η) and the anxious-depression factor ($r(578) = 0.124$, $p = 0.003$; controlling for age and IQ: $r(573) = 0.101$, $p = 0.015$).

Step 4.5. Further analysis of impulsivity factor and value-free random exploration association

In line with the previous impulsivity results, when performing a repeated-measures ANOVAs with horizon as the within-participants factor, we found a main effect of impulsivity (i.e., impulsivity factor score) on the low-value bandit (impulsivity main effect: $F(1,578) = 37.664$, $p < 0.001$, $\eta_p^2 = 0.061$; horizon main effect: $F(1,578) = 113.474$, $p < 0.001$, $\eta_p^2 = 0.164$; impulsivity-by-horizon interaction: $F(1,578) = 0.059$, $p = 0.808$, $\eta_p^2 = 0$) and on the ϵ -greedy parameter (impulsivity main effect: $F(1,578) = 40.872$, $p < 0.001$, $\eta_p^2 = 0.066$; horizon main effect: $F(1,578) = 126.418$, $p < 0.001$, $\eta_p^2 = 0.179$; impulsivity-by-horizon interaction: $F(1,578) = 2.906$, $p = 0.089$, $\eta_p^2 = 0.005$), but no horizon interaction.

Discussion

In this preregistered study, we investigated how impulsivity is related to exploration, and more specifically, how a computationally light exploration heuristic, value-free random exploration, is associated with different measures of impulsivity. Using a behavioural task and computational modelling we demonstrate that inter-individual

variability in value-free random exploration usage is associated with general impulsivity in a large-sample online study.

We and others have previously shown that humans deploy a multitude of different strategies for exploration^{22–25,30,54} that all approximate an optimal exploration strategy, which is intractable in open-ended decision problems. In our current data, we confirmed that our participants utilised a mixture of resource-requiring complex strategies and computationally light heuristics. The resource-demanding strategies (such as Thompson sampling or UCB) demand keeping track of expected means and uncertainties across the different choice options. The computationally lighter heuristic strategies, namely value-free random exploration (captured by ϵ -greedy) and novelty exploration (captured using a novelty bonus η), although being less optimal, require substantially less computational power, making them very useful in practice. Using model comparison as well as model simulations, we were able to demonstrate the presence of both complex and heuristic exploration strategies. The winning model, combining complex Thompson with novelty (η) and value-free random (ϵ) exploration, was not entirely distinguishable from the 2nd winning model, combining complex UCB with novelty and value-free random exploration, but was well distinguishable from other models (cf. confusion matrix, Supplementary Fig. 6b) with relatively high confidence regarding its generative origins (cf. inversion matrix, Supplementary Fig. 6c). This suggests that the two complex exploration strategies make similar predictions in our task, preventing us to disentangle them properly. However, we capture similar amounts of value-free random exploration, irrespective of the complex model used, demonstrating the robustness of our result. Our results therefore show that participants supplemented complex strategies (UCB or Thompson sampling) with two heuristic strategies. Given that we find an association between value-free random exploration and impulsivity irrespective of the complex model used, this does not impact the conclusions in the given study.

Impulsivity is a crucial construct across both general and clinical populations, but the links to specific computational mechanisms are still far from clear⁴. Based on previous theoretical^{6,12,14–17} and some experimental work^{21,23}, exploration is believed to be increased in impulsivity¹⁴ and especially in ADHD. Here, we extend these previous studies by identifying that it is value-free random exploration specifically which is increased, whilst other forms of exploration were not found to be robustly linked. This form of exploration is the computationally least demanding as it simply ignores all existing information. This is well aligned with a notion of impulsivity as ‘acting without thinking’, which is also captured in the motor impulsivity scale of the BIS. The latter showed a much closer association with value-free random exploration than the other attentional and non-planning impulsivity BIS subscores, which capture the inability to concentrate or a lack of forethought. We did not find a significant association between this form of exploration and a measure of global cognitive flexibility (cf. Supplementary Information), supporting the idea that cognitive flexibility and planning inability might be different neurocognitive constructs. However, it would be interesting to investigate whether value-free random exploration is related to more specific tasks, such as set shifting, inhibition or other decision making and learning tasks, given that cognitive flexibility in itself is a relatively heterogeneous construct⁵⁵.

From our results, it remains unclear which brain processes exactly mediate value-free random exploration. Interestingly, we have previously found that this form of exploration is modulated by noradrenaline functioning²³, a neurotransmitter which plays an important role in impulsivity-related disorders such as ADHD^{6,7,12,17,34–38}, which could be a potential mechanism. Previous findings that linked noradrenaline functioning to what is traditionally seen as motor impulsivity support this notion⁵⁶. This form of exploration may be also related to brain circuits generally seen to be linked to noradrenaline functioning

(for a detailed discussion of noradrenaline and executive functions, see Chamberlain & Robbins⁵⁶). In particular, anterior cingulate cortex would be a candidate as it is heavily innervated by noradrenaline and linked and linked to similar exploratory behaviour³³. In addition, fronto-striatal loops including orbito-frontal and dorso-lateral pre-frontal cortex may also be involved, as they have often been found to be involved in tasks that are modulated by noradrenaline related to set shifting^{56,57}. However, the precise neural processes underlying value-free random exploration needs to be examined in more detail.

Given that value-free random exploration ignores all prior information, it begs the question why humans use this strategy in exploration. Interestingly, inducing randomness or noise has often been shown to benefit a system both in living species and in machines, supporting the importance of such strategies^{12,58–62}. Here, the main benefit of value-free random exploration is that it does not require demanding computations, allowing exploration even with restrained neural resources⁶³ or a limited ability/willingness to engage with mentally effortful computations⁴³. Exploring in a seemingly random way can be beneficial, either at an individual or a group level, in many different contexts. For example, in the case of an absence of prior knowledge¹⁴, increased stochasticity can help to speed up learning. Additionally, in a case of imprecise or even inaccurate prior knowledge, random exploration ignores such erroneous priors and prevents them from penalizing future decision-making. Introducing stochasticity can also be beneficial in the case of dynamic environments e.g., where values can change drastically and thus agents should not rely solely on their expectations⁶². Our findings of such exploration heuristics are also well aligned with recent findings showing that limiting cognitive resources impacts the use of exploration strategies⁶⁴, and shifts in exploration strategies can be induced by applying constraints such as time pressure⁶⁵. Overall, our findings suggest at least two roles for exploration in impulsivity: a more flexible way of exploration which does not rely on (potentially wrong) prior knowledge and a way to circumvent mental effort. Importantly, value-free random exploration is used by all participants in a goal-directed manner (i.e., they used it more when exploration was beneficial). This means that participants adapt their usage of value-free random exploration to the demands of the task.

Because impulsivity is a feature of multiple psychiatric disorders, we investigated it in a transdiagnostic, population-based dimensional manner. This approach allowed us to capture a more general dimension of impulsivity rather than a sub-trait of a specific disorder. To obtain a transdiagnostic impulsivity factor, we performed a factor analysis similar to previous studies^{39–42}. Such an approach also helps to reduce the noise that is present when investigating individual questionnaires. We identified an impulsivity factor, capturing both impulsivity questionnaires (BIS, ASRS) as well as some aspects of OCD (as captured by OCI-R that was also related to value-free random exploration, cf. Supplementary Fig. 11). Interestingly, this factor was associated to value-free random exploration to an even stronger degree than the individual impulsivity questionnaires.

In addition to the impulsivity factor, we also identified an anxious-depression factor, but unlike previous studies we did not find a separate compulsivity factor. This is most probably due to the fact that we did not use the exact same set of questionnaires (previous studies included more compulsivity-related questionnaires). As some previous studies have suggested that depression and anxiety are associated to abnormal exploration^{45,47,48}, we explored these possible links in our dataset. After controlling for multiple comparisons, we indeed observed an association between the anxiety-depression factor and our parameter capturing the intrinsic value of novelty, the novelty bonus η . We did not observe any association between the third, uncertainty-related factor, nor any specific exploration strategy. Our findings suggest that those with increased anxiety-depression traits deployed the novelty-related exploration heuristic more eagerly. This

is aligned with previous findings showing increased exploration in participants with higher levels of anxiety^{66,67}. It is believed that this is because exploration aids in overcoming long-term uncertainty, and an uncertainty aversion is commonly reported in anxiety⁶⁸. Targeting novelty in exploration might be a way to save cognitive resources as one does not need to compute expected values and uncertainties of the other options, but instead can be simply guided by what has not been encountered before. This strategy thus seems deployable even under increased stress and anxiety. Even though we have rigorously controlled for multiple comparisons, we believe an independent replication of this somewhat unexpected result would be desirable. Moreover, it would be interesting to assess whether the deployment of such novelty exploration is more closely linked to apathy or anhedonia, as they are both important features of depression.

We did not find any direct association between the transdiagnostic factors and our complex exploration strategy (here: Thompson sampling). It needs to be noted that our task was optimised to detect the exploration heuristics. As a consequence, the complex exploration strategies make relatively similar predictions (cf. Supplementary Information). It is thus possible that in other tasks (e.g., by varying the generative bandit variance^{25,65,69}; or larger decision spaces³⁰), the coexistence of Thompson and UCB exploration is clearer and may be more directly linked to one of the trans-diagnostic dimensions. However, this is unlikely to impact the impulsivity findings presented here, as we find them irrespective of the complex strategy we are using in our computational models (cf. Supplementary Information). In addition, alternative exploration strategies, such as repeating one's previous choice could provide additional insight⁶⁵.

In this registered report, we demonstrated that transdiagnostic impulsivity is associated with value-free random exploration. By pre-registering and peer-reviewing our specific hypotheses using a previously-validated task^{22,23} and a well-defined dimensional approach^{39–41,70}, we were able to demonstrate this specific association. Our results aid in understanding the adaptivity of impulsivity and are important for the understanding of behaviour in the general and in clinical populations given the high prevalence of impulsivity. Nonetheless, future studies should investigate the validity of those effects in clinically diagnosed patient populations.

Methods

Ethics information

The study has been approved by the UCL research ethics committee (REC No 15301/001) and written informed consent was obtained from all participants. Participants were reimbursed for their participation on an hourly basis and received a bonus according to their performance (proportional to the sum of obtained rewards). The total compensation was bound between £8.25 and £12.0 per hour.

Design

Task. Participants were recruited online on Prolific Academic (www.prolific.ac), which manages the participant allocation and their reimbursement. Participants signed an online consent form and were redirected to the task.

We deployed a multi-armed bandit task which we have recently developed²³, and which allows us to capture different forms of exploration. On each trial, participants had to choose between different bandits (depicted as trees; cf. Figure 1) which one they want to draw a sample (i.e., pick an apple) from and therefore obtain a reward (the apple's size). Participants were instructed to maximise their score (i.e., sum of apple sizes) in order to maximise their overall reimbursement (i.e., they were instructed that they will receive a cash bonus proportional to their performance). Prior to the participants' first choice, bandits display varying levels of information about the plausible rewards they carry. Information is given in the form of 'initial samples', i.e., apples that have been picked before. We varied the

number of initial samples that were displayed for each bandit (identifiable by colour) to dissociate different forms of exploration (cf. below). The initial samples of each bandit are drawn from their generative normal distributions (cf. Supplementary Information for detail), meaning that initial samples carry important information for future choices as the mean of already observed bandits can be estimated²⁴.

To induce changes in exploration, similar to the horizon task²⁴ and our previous studies^{22,23}, we manipulated the number of samples they could draw from a given set of bandits²⁴. This decision horizon varied between two conditions (intermixed trials): they could either perform one draw (short horizon condition) or six draws (long horizon condition). The long horizon promotes exploration as obtained information can subsequently be used²⁴. Although there would be no interest for an optimal agent to explore in the short horizon, humans still show signs of exploration even when it is not beneficial, though to a much lesser extent^{14,58}. In fact, exploration in the short horizon has previously been observed in humans^{23,24,71}.

We constructed the reward and information of each bandit to be able to assess the contributions of different exploration strategies that have previously been put forward²³⁻²⁵. Each bandit i is from one of four generative groups characterised by different means μ_i and number of initial samples, following the same procedure as other studies²⁴. The size of the apple is determined by its radius (cf. Supplementary Fig. 1). Manipulating the amount of information participants have before they make their choice (i.e., initial samples) avoids a potential reward-information confound²⁴. The samples of each bandit are then sampled from a normal distribution with a fixed sampling variance $\mathcal{N}(\mu, 0.8)$, truncated to $[2, 10]$, and rounded to the closest integer. Each mean μ was sampled from $\mathcal{N}(\mu_{\text{overall}}, 1.4)$, with an “overall mean” μ_{overall} specific to each bandit type. The overall mean was computed similarly to previous studies:²⁴ On each trial we set the overall mean for one of the bandits, the ‘certain-standard bandit’, to be either 4.5 or 6.5. We determine the overall mean of the ‘standard bandit’ by adding a number sampled uniformly from $[-2, -1, +1, +2]$ to the certain-standard bandit overall mean. Similarly, we determine the overall mean of the ‘novel’ bandit by adding a number sampled uniformly from $[-2, -1, +1, +2]$ to either the certain-standard bandit overall mean or the standard bandit overall mean. By doing this, we make sure that the means of those 3 bandits are comparable. This results in the means of the standard bandit and novel bandit spanning a slightly larger range compared to the certain-standard bandit means (cf. Supplementary Table 3). To make sure that the ‘low-value’ bandit mean was always the smallest, it’s overall mean is computed by subtracting 1 to the minimum of the above-mentioned average means. Bandits also carry different amounts of information: The certain-standard bandit provides 3 initial samples, the standard bandit provides 1 initial sample, the novel bandit does not provide any initial samples and the low-value bandit provides 1 initial sample. Even though the absolute range of reward is set, randomly scaling each reward mean around the certain-standard bandits’ reward mean allows to maintain uncertainty about the overall average reward on each trial similarly to previous studies^{24,27}. On each trial, the average value of the certain-standard bandit initial samples is compared to the value of the standard bandit initial sample. The bandit with such a higher value is referred to as the (expected) ‘high-value’ bandit. For detailed comparison between those average rewards cf. Supplementary Table 19. At the beginning of each trial, the initial samples of the presented bandits are sampled from their respective distributions. We ensured that the initial sample from the low-value bandit is the smallest by resampling from this bandit in the trials where it is not the case, similar to our previous study. For detailed information about the value of initial samples, first draw and later draws cf. Supplementary Table 3. The order of all initial samples is then permuted to avoid biases. Additionally, to be able to compute choice consistency which is specifically reduced in value-free random

exploration²³, each trial is duplicated. Overall, each participant is asked to play 400 trials (200 in each horizon condition). The trees’ positions (left, middle or right) as well as their colour (8 sets of 3 different colours) where shuffled between trials.

The task has originally been developed in a lab setting²³ and has now been adapted for online use. We have adjusted the instructions, making them as clear as possible while keeping the participants’ attention. Following the initial task instructions, to make sure that they understood what they need to do, they were asked to answer 5 questions. Similar to previous online studies⁴⁰⁻⁴², failing to correctly answer these questions guided the participant back to the instructions until all correct answers are given. To make sure that participants understood that the apples from the same tree are always of similar size (generated following a normal distribution), participants additionally performed several training trials. In this training, based on three displayed apples of similar size, they had to guess, between two options, which apple is the most likely to come from the same tree and receive feedback about their choice. If participants gave a wrong answer in at least 3 of the 10 trials, they were asked to restart the training. Task pilot data (N = 61, cf. below) demonstrated comparable effects and effect sizes (cf. Data analysis, Step 1) to our previous lab-based data²³.

Behavioural analysis nomenclature. For the behavioural analysis, we categorized each bandit according to the number and size of initial samples (apples shown before the first draw). The bandit with the highest sampling mean, carrying either a lot or some prior information (i.e., 3 or 1 initial samples; for further split cf. Supplementary Information), is referred to as the ‘high-value bandit’. The bandit for which no prior sample was shown is named the ‘novel bandit’, and the bandit with one initial sample from a substantially lower generative mean (trials were constructed to have sufficient number of such trials²³) is called the ‘low-value bandit’. The high-value bandit is an evident signature of exploitation (choosing maximal expected value), the novel bandit is captured amongst other by ‘novelty exploration’ which is biased towards options for which nothing is known, and the ‘low-value bandit’ appeals to the value-free random exploration alone as it is the only strategy which does not take expected values into account²³.

Assessing psychiatric symptoms. After completing the task, participants were asked to fill-in several self-report questionnaires. To assess impulsivity, our key dimension of interest, we used the Adult ADHD Self-Report Scale⁷² (ASRS) and the Barratt Impulsiveness Scale⁴⁹ (BIS). In addition, we collected further questionnaires to investigate additional psychiatric dimensions (cf. Data analysis, Step 3). These entail the Liebowitz Social Anxiety Scale⁷³ (LSAS), the State-Trait Anxiety Inventory⁷⁴ (STAI-Y2), Intolerance of Uncertainty Scale⁷⁵ (IUS), Obsessive-Compulsive Inventory-Revised⁷⁶ (OCI-R), and Zung’s Self-rating Depression Scale⁷⁷ (SDS), in accordance with similar previous approaches⁴⁰⁻⁴², as well as the Cognitive Flexibility Scale⁵³ (CFS) and the Autism spectrum Quotient⁵² (AQ-10). To control for confounding factors, such as intelligence and medication, participants additionally completed the International Cognitive Ability Resource sample test⁷⁸ (ICAR) and were asked whether they take psychoactive medication and/or medication to increase attention/concentration on a regular basis. As a measure of data quality, attention checks are added to every questionnaire to make sure that participants read the questions⁷⁹. Failure in 1 or more attention check resulted in the participants’ exclusion from data analysis.

Blinding and randomisation do not apply for this study. The full code (written using the open source React JavaScript library) of the task can be found online (<https://github.com/MagDub/MFweb-app>).

Sample

Power analyses

The analysis consisted of two preregistered steps addressing separate research questions. Step 1 consisted of expanding our pilot data (cf. Data analysis) and replicating the main characteristics of the previously lab-based task²³ in an online setting. In Step 2 we assessed our main research questions and looked at associations between exploration measures and impulsivity traits. Lastly, an exploratory factor analysis was conducted at Stage 2 and is reported subsequently. In Step 3 the factor analysis across all questionnaires allowed us to explore the relationship between psychiatric dimensions and exploration measures more broadly.

For Step 1's sample size estimation, in which we attempted to replicate the task main effects, we collected online pilot data ($N=61$ after exclusion). A total of 4 hypotheses were tested (hypothesis 1.1 to 1.4; cf. Table 1 and Data analysis for details). The lowest effect size across all tests in the pilot study (Wilcoxon signed-rank effect size = 0.410) was used for our power analysis, which suggested that a sample of $N=83$ is sufficient to reach 95% power for all hypotheses. For a summary of the statistics performed on all measures on the pilot data cf. Supplementary Table 20. Importantly, Wilcoxon signed-rank tests was used instead of paired t-tests if the Shapiro normality assumption was violated.

For Step 2's sample size estimation, where the link between exploration and impulsivity was investigated, the correlation coefficient of our previous study using the same task²² was used for our power analysis. In this prior study, a Pearson correlation of $r=0.26$, $p<0.001$ was observed between an impulsivity measure⁸⁰ (the Conners ADHD questionnaire) and value-free random exploration in youths. Assuming a similar correlation in adults, our power analysis suggested that a sample of $N=190$ is sufficient to reach 95% power (G*Power analyses suggest a similar sample size of $N=186$). This moderate size correlation factor is in line with previous studies linking BIS-measured impulsivity to behaviour (e.g., with delay discounting^{81–84}). Similarly, for Step 2's Stage 2 exploratory analysis in which we looked at the correlation between value-free random exploration and the three subdomains of BIS, G*Power analyses suggested that a sample size of $N=228$ is sufficient to reach a 95% power at a significance corrected for multiple comparisons using Bonferroni correction.

However, assuming a lower association strength and taking into account previous dimensional analyses using exploratory factor analysis (similar to our exploratory Step 3), we additionally considered the correlation coefficients obtained from these previous big data dimensional studies. These previous studies have observed correlations from $r=0.15$ (negative association between dogmatism and metacognitive sensitivity³⁹) up to correlations of $r=0.25$ (association between confidence and compulsivity⁴¹). The relatively small effect sizes can be explained by the higher noise associated with large online samples as well as the lack of precision of behavioural and questionnaire measures. In this study we account for these facets and consider the study as a first step to establish associations between measures by conducting thorough effect size and power calculations. The lowest correlation ($r=0.15$) was used to extend our power analysis, which suggested that a sample of $N=580$ is sufficient to reach 95% power (cf. Supplementary Fig. 12; G*Power analyses suggested a similar sample size of $N=571$). Taking all steps together, to reach at least 95% power across all measures, we collected a total sample of $N=580$ participants. A sensitivity power analysis (performed in G*Power) predicted that with such a sample size, we would be able to detect an effect size (Minimal Detectable Effect, MDE) of 0.15 with 95%. Importantly, the lower bound of the 95% Confidence Interval of each pilot data measures' effect size was above this MDE (cf. Supplementary Table 20), ensuring a detectable effect.

Effect sizes as well as hypothetical sample sizes to reach a 95% power can be found in Table 1 (details about each measure in the pilot data can be found in Supplementary Table 20). For power analysis, we used the G*Power Software⁸⁵ for the t-tests in Step 1 (using the pilot data). Power analysis for the correlations in the further steps was performed using simulations in MATLAB. The 'matter' gallery of the 'cmocean' colourmap was used for the figures^{86,87}. We obtained summary statistic scores or correlations from previous studies and used bootstrapping to simulate data. Concretely, for t-tests, n simulated participants were sampled from each group normal distribution: $N(m1, std1)$ and $N(m2, std2)$ and significance was assessed using paired t-test on those 2 data sets. For correlations, n simulated participants were taken from the bivariate distribution of mean = [0, 0] and covariance matrix = [1, R; R, 1]. To assess power of a given sample size, we assessed the number of significant tests ($p<0.05$) of a total number of $N=10000$ simulations. The summary statistics for Step 1 were taken from our pilot data (cf. Supplementary Information), for Step 2 from our previous study in youths²², and for Step 3 from previous big data studies^{39,41}.

Participant recruitment. To take part in the study, participants had to be above 18 years of age and have their current residence in the UK. To ensure data quality, participants were excluded according to the following criteria: data was incomplete, the mean score (i.e., apple size) was lower than 5.5 indicating participants were performing at chance level⁴⁰ (cf. Supplementary Fig. 2b), the first draw mean reaction time was faster than 1500 ms (based on our pilot data and previous study²³) indicating participants were not allocating much thought to their choice (cf. Supplementary Fig. 2c) and if participants failed at least one attention check during the questionnaires meaning that they were not reading the questions^{40,41,79}. According to these exclusion criteria, $N=77$ participants were excluded (cf. Supplementary Fig. 2) and replaced prior to data analysis in order to reach a final sample of $N=580$ ($N=3$ participants were excluded out of $N=64$ in the pilot data).

Data analysis

Step 1

This step aims at replicating the main characteristics of the previously lab-based task²³ in an online setting. Here, we report results from our pilot data set ($N=61$), which we collected online using the exact same online task to estimate the effect sizes. The analysis follows the pipeline which we have successfully used in our previous studies^{22,23}. In line with previous studies investigating horizon-dependent exploration^{23,24,71}, we only investigated the first draw of each horizon in the main analysis. This allowed us to compare between horizon conditions preventing biases of collected reward and unequal variance.

Participants explore more when it is worth it. To assess whether the horizon manipulation promoted exploration, we analysed whether participants explored more in the long (versus short) horizon condition, in which additional information can inform later choices. To this end, we assessed which bandit participants chose on their first draw. Replicating our previous studies^{22,23}, we expected several exploration markers to differ. We predicted that participants would choose bandits with a lower expected value (computed as the mean of the bandits' initial samples) in the long horizon (pilot data: $t(60)=3.585$, $p=0.001$, 95% confidence interval of the mean: $CI_M=[0.047, 0.165]$, effect size: Cohen's $d=-0.459$, 95% confidence interval of the effect size $CI_{ES}=[-0.727, -0.195]$). This is reflected by the frequency of picking the high-value bandit, which we predicted to decrease in the long horizon (pilot data: $t(60)=8.45$, $p<0.001$, 95% $CI_M=[6.92, 11.211]$, $d=-1.082$, 95% $CI_{ES}=[-1.407, -0.769]$). Similarly for the frequency of picking the low-value bandit, we predicted it to increase in the long horizon (pilot

data: $t(60) = -3.446$, $p = 0.001$, $95\%CI_M = [-1.568, -0.416]$, $d = 0.441$, $95\%CI_{ES} = [0.178, 0.708]$). We predicted this exploration to be goal-directed, with participants choosing bandits they know less about (lower number of initial samples, i.e., more informative) in the long horizon (pilot data: $t(60) = 9.625$, $p < 0.001$, $95\%CI_M = [0.184, 0.281]$, $d = -1.232$, $95\%CI_{ES} = [-1.576, -0.903]$). This is largely reflected by the frequency of the novel bandit, which we predicted to increase in the long horizon (pilot data $t(60) = -8.586$, $p < 0.001$, $95\%CI_M = [-11.178, -6.954]$, $d = 1.099$, $95\%CI_{ES} = [0.784, 1.427]$).

Participants use exploration beneficially. To evaluate whether participants were able to use exploration beneficially, we looked at their performance (i.e., the outcomes they obtained). We first compared the reward (i.e., apple size) obtained in the short horizon with the first reward obtained in the long horizon. As the latter is driven by exploration, we expected it to be lower (pilot data: $t(60) = 6.522$, $p < 0.001$, $95\%CI_M = [0.059, 0.112]$, $d = -0.835$, $95\%CI_{ES} = [-1.134, -0.545]$). As observed in previous studies^{22,23}, we expected them to make good use of the additional information earned by exploring, and therefore their long horizon average reward (across 6 draws) was expected to be higher than the short horizon reward (pilot data: $t(60) = -16.096$, $p < 0.001$, $95\%CI_M = [-0.245, -0.191]$, $d = 2.061$, $95\%CI_{ES} = [1.626, 2.524]$).

Participants explore using heuristics. To formally assess which exploration strategies were being used, we turned to computational modelling. Similar to the behavioural analysis, only the first draw of each trial was analysed. We compared 16 models that make different predictions about the usage of exploration strategies (cf. Supplementary Information). Similar to our previous studies^{22,23}, we expected participants to use a mixture of computationally demanding (i.e., Thompson sampling and/or UCB) and heuristic exploration strategies (i.e., value-free random exploration and novelty exploration) captured by the winning model (pilot data: BIC average score: Thompson + $\eta + \epsilon$ vs Thompson model: $t(60) = -10.187$, $p < 0.001$, $95\%CI_M = [-72.866, -48.946]$, $d = 1.304$, $95\%CI_{ES} = [0.967, 1.657]$). Model comparison was computed using the commonly used Bayesian Information Criterion (BIC). The winning model, i.e., the model with the lowest BIC score, was used for subsequent analyses. All models that were not significantly different than the 1st winning model would have been used for subsequent analysis to demonstrate the generalisability of the effect (similar to previous studies²²). Model fitting was performed using the maximum a posteriori probability (MAP) estimate, which allows incorporation of prior beliefs. All the parameters besides participants' initial estimate of a bandit's mean (Q0; prior mean) and the contribution of each model in the hybrid model (w) were free to vary as a function of the horizon as they capture different exploration forms (cf. Supplementary Information for details).

Participants rely more on heuristics in the long horizon. To assess the changes in exploration strategy, we examined the winning model's fitted parameters. Those parameters were fitted to the first draw of all trials of each participant. We expected the ϵ -greedy parameter, which captures the contribution of value-free random exploration, to be increased in the long (versus short) horizon (pilot data: $t(60) = -3.23$, $p = 0.002$, $95\%CI_M = [-0.058, -0.014]$, $d = 0.413$, $95\%CI_{ES} = [0.152, 0.679]$). Similarly, we expected the novelty bonus η , which captures the intrinsic reward of selecting a novel option, to be increased in the long horizon (pilot data: $t(60) = -9.43$, $p < 0.001$, $95\%CI_M = [-1.265, -0.822]$, $d = 1.207$, $95\%CI_{ES} = [0.881, 1.548]$).

Step 2

In this step we tested our main hypothesis about value-free random exploration being linked to impulsivity and ADHD traits. Our key measure of interest is the mean ϵ parameter²² - measuring value-free

random exploration - and how it is related to our specific questionnaire measures. For the correlations, we used the Pearson correlation coefficient and we performed both a bivariate correlation as well as a partial correlation to control for age and IQ²². The IQ score was computed as the sum of the correct answers on the ICAR sample test⁷⁸. Additionally, we also performed repeated-measures ANOVAs with within factor horizon and a between participants variable [impulsivity/ADHD-symptoms] to assess these effects further.

Step 2.1. First, we looked at impulsivity within a broad spectrum, and expected it to be linked with value-free random exploration. For this, we used the total score on the Barratt Impulsiveness Scale (BIS). The BIS is the most commonly administered self-report measure for assessment of impulsiveness⁴⁹, and has already been used in online studies^{40,41}. We looked at the correlation between the BIS total score and the low-value bandit frequency, and between the BIS total score and the ϵ -greedy parameter. These associations allowed us to conclude that value-free random exploration is linked to impulsivity traits in general, which has implications for impulsivity disorders beyond ADHD.

Considering that impulsivity is a broad heterogeneous construct¹³⁻⁵, in Stage 2 we performed an exploratory analysis of the three subdomains of BIS (i.e., attentional, motor, and non-planning behaviour⁸⁸) similarly to previous studies⁸⁹. We investigated whether value-free random exploration is linked to a specific subdomain by looking at the correlations with each of them. Specifically, we looked at the correlation (corrected for multiple comparisons using Bonferroni correction) between the low-value bandit frequency and the BIS subdomains: attentional, motor and non-planning, as well as the correlation (corrected for multiple comparisons using Bonferroni correction) between the ϵ -greedy parameter and the BIS subdomains: attentional, motor and non-planning.

Step 2.2. Second, we looked at ADHD symptoms across our sample and expected to find an association of higher ADHD scores being related to increased value-free random exploration. This analysis extends our previous preliminary findings showing a positive association in youths (9–18 year olds) between ADHD traits (the Conners ADHD questionnaire⁸⁰) and value-free random exploration²². We looked at the correlation between the ASRS total score and the low-value bandit frequency, and between the ASRS total score and the ϵ -greedy parameter. It allows a definitive answer to the hypothesis whether ADHD symptoms are linked to value-free random exploration^{12,21,23}. The ADHD measure we used was the total score on the Adult ADHD Self-Report Scale (ASRS), a questionnaire which was developed by the World Health Organization and is used for screening ADHD in the general population⁷². In Stage 2 we additionally performed an exploratory analysis of the sub-scales of the ASRS (i.e., inattention, hyperactivity-impulsivity). Specifically, we looked at the correlation (corrected for multiple comparisons using Bonferroni correction) between the low-value bandit frequency and the ASRS sub-scales: inattention and hyperactivity-impulsivity. We also examined the correlation (corrected for multiple comparisons using Bonferroni correction) between the ϵ -greedy parameter and the ASRS sub-scales: inattention and hyperactivity-impulsivity.

Step 3 (Stage 2)

In Stage 2, we performed a further exploratory step. In order to investigate whether there exists a latent trans-diagnostic structure which can help to explain exploration differences, we performed a factor analysis. First, we used the raw scores from all questionnaire items as variables to reduce their dimensionality similarly as previous studies⁴⁰⁻⁴². Factor analysis was conducted using the $fa()$ function from the Psych package in R, with an oblique rotation (oblimin; we draw the reader's attention to the fact that the $factanal()$ function, which does

not allow for such rotation, was erroneously mentioned in the Stage 1 protocol). The number of factors was based on the Cattell's criterion⁹⁰, using the Cattell-Nelson-Gorsuch test (nFactors package in R). Factors were labelled based on the items which loaded the most strongly in a consensus discussion among the authors.

First, we expected our two impulsivity questionnaires (BIS and ASRS) to primarily load onto one factor, and we expected this factor to be at least as much associated with value-free random exploration as the impulsivity/ADHD questionnaires alone (cf. above). In addition to the hypothesized increase in value-free random exploration, we investigated using multiple comparison whether impulsivity correlates with other forms of exploration (e.g., complex strategies).

As a second step, we investigated whether exploration correlates with other factors. In particular, similar to previous studies^{40–42,50,51}, we expected to retrieve a depression / anxiety dimension, on which depression, social anxiety and anxiety would load onto (SDS, LSAS and STAI-Y2 questionnaires respectively) and a compulsivity dimension, on which OCD and uncertainty intolerance traits load onto (OCIR and IUS questionnaires) respectively. Indeed, previous research has found that impulsivity and compulsivity only show a modest overlap⁶³, which is also why previous studies that used factor analyses have found that these items load onto different factors^{40,42}. Previous studies have demonstrated increases in exploration in OCD patients^{44,91}, but it is not clear which exploration strategy is concerned. We therefore looked at the correlation (corrected for multiple comparisons using Bonferroni correction) between the compulsivity dimension and each exploration free parameter (depending on the model). Similarly, studies have demonstrated abnormality in exploration in patients with depression⁴⁷, anxiety⁴⁸ and other disorders related to avoidance of uncertainty⁴⁵. However, different exploration strategies have not been tested. We therefore looked at the correlation (corrected for multiple comparisons using Bonferroni correction) between the depression/anxiety dimension and each exploration free parameter (depending on the model). We also investigated two separate questions. First, we looked at the correlation between the autism scale, AQ-10 total score and value-free random exploration, as autism has overlapping symptoms with ADHD⁹². An association between the autism score and our impulsivity measure would have resulted in further analysis using partial correlations. Second, we looked at the correlation between the cognitive flexibility scale (CFS) and value-free random exploration, as cognitive flexibility is thought to play a role in the exploration-exploitation trade-off^{93,94}.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The raw (anonymized) and processed data are available at Github <https://github.com/MagDub/Mfweb-data> and Zenodo: <https://doi.org/10.5281/zenodo.6522060>⁹⁵. The pilot data are available at Github https://github.com/MagDub/Mfweb-pilot_data and Zenodo: <https://doi.org/10.5281/zenodo.6522062>⁹⁶. Source data are provided with this paper.

Code availability

Code for power simulations, computational modelling and data analysis can be found on Github: https://github.com/MagDub/MFweb-data_analysis and Zenodo: <https://doi.org/10.5281/zenodo.6445661>⁹⁷.

References

1. Evenden, J. L. Varieties of impulsivity. *Psychopharmacol. (Berl.)* **146**, 348–361 (1999).

2. Barratt, E. S. Anxiety and impulsiveness related to psychomotor efficiency. *Percept. Mot. Skills*. <https://doi.org/10.2466/pms.1959.9.3.191> (1959).
3. Caswell, A. J., Bond, R., Duka, T. & Morgan, M. J. Further evidence of the heterogeneous nature of impulsivity. *Pers. Individ. Dif.* **76**, 68–74 (2015).
4. Dalley, J. W. & Robbins, T. W. Fractionating impulsivity: neuropsychiatric implications. *Nat. Rev. Neurosci.* **18**, 158–171 (2017).
5. Dalley, J. W., Everitt, B. J. & Robbins, T. W. Impulsivity, compulsivity, and top-down cognitive control. *Neuron* **69**, 680–694 (2011).
6. Robbins, T. W., Gillan, C. M., Smith, D. G., de Wit, S. & Ersche, K. D. Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends Cogn. Sci.* **16**, 81–91 (2012).
7. Robinson, E. S. J. et al. Similar effects of the selective noradrenaline reuptake inhibitor atomoxetine on three distinct forms of impulsivity in the rat. *Neuropsychopharmacology* **33**, 1028–1037 (2008).
8. Benn, A. & Robinson, E. S. J. Differential roles for cortical versus sub-cortical noradrenaline and modulation of impulsivity in the rat. *Psychopharmacol. (Berl.)* **234**, 255–266 (2017).
9. Besson, M. et al. Dissociable control of impulsivity in rats by dopamine D2/3 receptors in the core and shell subregions of the nucleus accumbens. *Neuropsychopharmacology*. <https://doi.org/10.1038/npp.2009.162> (2010).
10. Costa, A. et al. Impulsivity is related to striatal dopamine transporter availability in healthy males. *Psychiatry Res. - Neuroimaging* **211**, 251–256 (2013).
11. Economidou, D., Theobald, D. E. H., Robbins, T. W., Everitt, B. J. & Dalley, J. W. Norepinephrine and dopamine modulate impulsivity on the five-choice serial reaction time task through opponent actions in the shell and core sub-regions of the nucleus accumbens. *Neuropsychopharmacology* **37**, 2057–2066 (2012).
12. Hauser, T. U., Fiore, V. G., Moutoussis, M. & Dolan, R. J. Computational Psychiatry of ADHD: Neural Gain Impairments across Marrian Levels of Analysis. *Trends Neurosci.* **39**, 63–73 (2016).
13. Loos, M. et al. Dopamine receptor D1/D5 gene expression in the medial prefrontal cortex predicts impulsive choice in rats. *Cereb. Cortex*. <https://doi.org/10.1093/cercor/bhp167> (2010).
14. Williams, J. & Taylor, E. The evolution of hyperactivity, impulsivity and cognitive diversity. *J. R. Soc. Interface* **3**, 399–413 (2006).
15. Williams, J. & Dayan, P. *Dopamine, Learning, and Impulsivity: A Biological Account of Attention-Deficit/Hyperactivity Disorder*. (2005).
16. Sonuga-Barke, E. J. S., Cortese, S., Fairchild, G. & Stringaris, A. Annual Research Review: Transdiagnostic neuroscience of child and adolescent mental disorders - Differentiating decision making in attention-deficit/hyperactivity disorder, conduct disorder, depression, and anxiety. *J. Child Psychol. Psychiatry Allied Discip.* **57**, 321–349 (2016).
17. Luman, M., Tripp, G. & Scheres, A. Identifying the neurobiology of altered reinforcement sensitivity in ADHD: a review and research agenda. *Neurosci. Biobehav. Rev.* **34**, 744–754 (2010).
18. Scheres, A., Tontsch, C., Thoeny, A. L. & Kaczurkin, A. Temporal reward discounting in attention-deficit/hyperactivity disorder: the contribution of symptom domains, reward magnitude, and session length. *Biol. Psychiatry* **67**, 641–648 (2010).
19. Sadeghiyeh, H. et al. Temporal discounting correlates with directed exploration but not with random exploration. *Sci. Rep.* **10**, (2020).
20. Moutoussis, M., Dolan, R. J. & Dayan, P. How people use social information to find out what to want in the paradigmatic case of inter-temporal preferences. *PLoS Comput. Biol.* **12**, 1–17 (2016).
21. Hauser, T. U. et al. Role of the medial prefrontal cortex in impaired decision making in juvenile attention-deficit/hyperactivity disorder. *JAMA Psychiatry* **71**, 1165–1173 (2014).

22. Dubois, M. et al. Exploration heuristics decrease during youth. *Cogn Affect Behav Neurosci* (2022) <https://doi.org/10.3758/s13415-022-01009-9> (2022).
23. Dubois, M. et al. Human complex exploration strategies are enriched by noradrenaline-modulated heuristics. *Elife* **10**, 1–34 (2021).
24. Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A. & Cohen, J. D. Humans use directed and random exploration to solve the explore–exploit dilemma. *J. Exp. Psychol. Gen.* **143**, 2074–2081 (2014).
25. Gershman, S. J. Deconstructing the human algorithms for exploration. *Cognition* **173**, 34–42 (2018).
26. Carpentier, A., Lazaric, A., Ghavamzadeh, M., Munos, R. & Auer, P. Upper-confidence-bound algorithms for active learning in multi-armed bandits. *Lect. Notes Comput. Sci.* **6925 LNAI**, 189–203 (2011).
27. Wu, C. M., Schulz, E., Garvert, M. M., Meder, B. & Schuck, N. W. Similarities and differences in spatial and nonspatial cognitive maps. *PLoS Comput. Biol.* **16**, 1–28 (2020).
28. Schulz, E. & Gershman, S. J. The algorithmic architecture of exploration in the human brain. *Curr. Opin. Neurobiol.* **55**, 7–14 (2019).
29. Stojic, H., Schulz, E., Analytis, P. P. & Speekenbrink, M. It’s New, but Is It Good? How generalization and uncertainty guide the exploration of novel options. *J. Exp. Psychol. Gen.* <https://doi.org/10.1037/xge0000749> (2020).
30. Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D. & Meder, B. Generalization guides human exploration in vast decision spaces. *Nat. Hum. Behav.* **2**, 915–924 (2018).
31. Krebs, R. M., Schott, B. H., Schütze, H. & Düzel, E. The novelty exploration bonus and its attentional modulation. *Neuropsychologia*. <https://doi.org/10.1016/j.neuropsychologia.2009.01.015> (2009).
32. Sutton, R. S. & Barto, A. G. Introduction to Reinforcement Learning. MIT Press Cambridge (1998). 10.1.1.32.7692
33. Tervo, D. G. R. et al. Behavioral variability through stochastic choice and its gating by anterior cingulate cortex. *Cell* **159**, 21–32 (2014).
34. Arnsten, A. F. T. & Pliszka, S. R. Catecholamine influences on prefrontal cortical function: relevance to treatment of attention deficit/hyperactivity disorder and related disorders. *Pharmacol. Biochem. Behav.* **99**, 211–216 (2011).
35. Berridge, C. W. & Devilbiss, D. M. Psychostimulants as cognitive enhancers: the prefrontal cortex, catecholamines, and attention-deficit/hyperactivity disorder. *Biol. Psychiatry* **69**, (2011).
36. Del Campo, N., Chamberlain, S. R., Sahakian, B. J. & Robbins, T. W. The roles of dopamine and noradrenaline in the pathophysiology and treatment of attention-deficit/hyperactivity disorder. *Biol. Psychiatry* **69**, e145–e157 (2011).
37. Frank, M. J., Santamaria, A., O’Reilly, R. C. & Willcutt, E. Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. *Neuropsychopharmacology* **32**, 1583–1599 (2007).
38. Pattij, T. & Vanderschuren, L. J. M. J. The neuropharmacology of impulsive behaviour. *Trends Pharmacol. Sci.* **29**, 192–199 (2008).
39. Rollwage, M., Dolan, R. J. & Fleming, S. M. Metacognitive failure as a feature of those holding radical beliefs. *Curr. Biol.* **28**, 4014–4021.e8 (2018).
40. Rouault, M., Seow, T., Gillan, C. M. & Fleming, S. M. Psychiatric symptom dimensions are associated with dissociable shifts in metacognition but not task performance. *Biol. Psychiatry* **84**, 443–451 (2018).
41. Seow, T. X. F. & Gillan, C. M. Transdiagnostic phenotyping reveals a host of metacognitive deficits implicated in compulsivity. *Sci. Rep.* **10**, 1–11 (2020).
42. Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A. & Daw, N. D. Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *Elife* **5**, 1–24 (2016).
43. Patzelt, E. H., Kool, W., Millner, A. J. & Gershman, S. J. The transdiagnostic structure of mental effort avoidance. *Sci. Rep.* **9**, 1–10 (2019).
44. Kanen, J. W., Ersche, K. D., Fineberg, N. A., Robbins, T. W. & Cardinal, R. N. 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. *Psychopharmacol. (Berl.)*. **236**, 2337–2358 (2019).
45. Morris, L. S. et al. Biases in the explore-exploit tradeoff in addictions: the role of avoidance of uncertainty. *Neuropsychopharmacology* **41**, 940–948 (2016).
46. Addicott, M. A., Pearson, J. M., Sweitzer, M. M., Barack, D. L. & Platt, M. L. A primer on foraging and the explore/exploit trade-off for psychiatry research. *Neuropsychopharmacology* **42**, 1931–1939 (2017).
47. Blanco, N. J., Otto, A. R., Maddox, W. T., Beevers, C. G. & Love, B. C. The influence of depression symptoms on exploratory decision-making. *Cognition* **129**, 563–568 (2013).
48. Browning, M., Behrens, T. E., Jocham, G., O’Reilly, J. X. & Bishop, S. J. Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat. Neurosci.* <https://doi.org/10.1038/nn.3961> (2015).
49. Stanford, M. S. et al. Fifty years of the Barratt Impulsiveness Scale: an update and review. *Pers. Individ. Dif.* **47**, 385–395 (2009).
50. St Clair, M. C. et al. Characterising the latent structure and organisation of self-reported thoughts, feelings and behaviours in adolescents and young adults. *PLoS One* **12**, 1–27 (2017).
51. Polek, E. et al. How do the prevalence and relative risk of non-suicidal self-injury and suicidal thoughts vary across the population distribution of common mental distress (the p factor)? Observational analyses replicated in two independent UK cohorts of young people. *BMJ Open* **10**, 1–9 (2020).
52. Allison, C., Auyeung, B. & Baron-Cohen, S. Toward brief ‘red flags’ for autism screening: the short Autism Spectrum Quotient and the short Quantitative Checklist in 1,000 cases and 3,000 controls. *J. Am. Acad. Child Adolesc. Psychiatry* **51**, 202–212.e7 (2012).
53. Martin, M. M. & Rubin, R. B. A new measure of cognitive flexibility. *Psychol. Rep.* **76**, 623–626 (1995).
54. Dezza, I. C., Noel, X., Cleeremans, A. & Yu, A. J. Distinct motivations to seek out information in healthy individuals and problem gamblers. *Transl. Psychiatry* **11**, (2021).
55. Friedman, N. P. et al. Not all executive functions are related to intelligence. *Psychol. Sci.* **17**, 172–179 (2006).
56. Chamberlain, S. R. & Robbins, T. W. Noradrenergic modulation of cognition: therapeutic implications. *J. Psychopharmacol.* **27**, 694–718 (2013).
57. Zhang, R., Geng, X. & Lee, T. M. C. Large-scale functional neural network correlates of response inhibition: an fMRI meta-analysis. *Brain Struct. Funct.* **222**, 3973–3990 (2017).
58. Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S. & Wyart, V. Computational noise in reward-guided learning drives behavioral variability in volatile environments. *Nat. Neurosci.* **22**, 2066–2077 (2019).
59. Pouget, A., Beck, J. M., Ma, W. J. & Latham, P. E. Probabilistic brains: knowns and unknowns. *Nat. Neurosci.* **16**, 1170–1178 (2013).
60. Drugowitsch, J., Wyart, V., Devauchelle, A. D. & Koehlin, E. Computational precision of mental inference as critical source of human choice suboptimality. *Neuron* **92**, 1398–1411 (2016).
61. Schmidt, M. & Lipson, H. Learning noise. *Proc. GECCO 2007 Genet. Evol. Comput. Conf.* 1680–1685 <https://doi.org/10.1145/1276958.1277289> (2007).

62. Gureckis, T. M. & Love, B. C. Learning in noise: dynamic decision-making in a variable environment. *J. Math. Psychol.* **53**, 180–193 (2009).
63. Ziegler, G. et al. Compulsivity and impulsivity traits linked to attenuated developmental frontostriatal myelination trajectories. *Nat. Neurosci.* **22**, 992–999 (2019).
64. Dezza, I. C., Cleeremans, A. & Alexander, W. Should we control? The interplay between cognitive control and information integration in the resolution of the exploration-exploitation dilemma. *J. Exp. Psychol. Gen.* **148**, 977–993 (2019).
65. Wu, C. M., Schulz, E., Pleskac, T. J. & Speekenbrink, M. Time pressure changes how people explore and respond to uncertainty. *Sci. Rep.* **12**, 1–14 (2022).
66. Bennett, D., Sutcliffe, K., Tan, N. P. J., Smillie, L. D. & Bode, S. Anxious and obsessive-compulsive traits are independently associated with valuation of noninstrumental information. *J. Exp. Psychol. Gen.* **150**, 739–755 (2020).
67. Aberg, K. C., Toren, I. & Paz, R. A neural and behavioral trade-off between value and uncertainty underlies exploratory decisions in normative anxiety. *Mol. Psychiatry* **27**, 1573–1587 (2022).
68. McEvoy, P. M. & Mahoney, A. E. J. Achieving certainty about the structure of intolerance of uncertainty in a treatment-seeking sample with anxiety and depression. *J. Anxiety Disord.* **25**, 112–122 (2011).
69. Tomov, M. S., Schulz, E. & Gershman, S. J. Multi-task reinforcement learning in humans. *Nat. Hum. Behav.* **5**, 764–773 (2021).
70. Gillan, C. M. & Daw, N. D. Taking psychiatry research online. *Neuron* **91**, 19–23 (2016).
71. Somerville, L. H. et al. Charting the expansion of strategic exploratory behavior during adolescence. *J. Exp. Psychol. Gen.* **146**, 155–164 (2017).
72. Kessler, R. C. et al. The World Health Organization adult ADHD self-report scale (ASRS): a short screening scale for use in the general population. *Psychol. Med.* **35**, 245–256 (2005).
73. Liebowitz, M. R. Liebowitz Social Anxiety Scale. *Mod. Probl. Pharmacopsychiatry* (1987).
74. Spielberger, C. D., Gorsuch, R. L. & Lushene, R. E. STAI manual for the state-trait anxiety inventory. Self-Evaluation Questionnaire. MANUAL <https://doi.org/10.1037/t06496-000> (1970).
75. Buhr, K. & Dugas, M. J. The intolerance of uncertainty scale: Psychometric properties of the English version. *Behav. Res. Ther.* [https://doi.org/10.1016/S0005-7967\(01\)00092-4](https://doi.org/10.1016/S0005-7967(01)00092-4) (2002).
76. Foa, E. B. et al. The obsessive-compulsive inventory: development and validation of a short version. *Psychol. Assess.* <https://doi.org/10.1037/1040-3590.14.4.485> (2002).
77. Zung, W. W. K. A Self-rating depression scale. *Arch. Gen. Psychiatry* <https://doi.org/10.1001/archpsyc.1965.01720310065008> (1965).
78. Condon, D. M. & Revelle, W. The international cognitive ability resource: development and initial validation of a public-domain measure. *Intelligence* **43**, 52–64 (2014).
79. Oppenheimer, D. M., Meyvis, T. & Davidenko, N. Instructional manipulation checks: detecting satisficing to increase statistical power. *J. Exp. Soc. Psychol.* **45**, 867–872 (2009).
80. Conners, C. K. *Conners 3rd Edition (Conners 3)*. Journal of Psychoeducational Assessment <https://doi.org/10.1177/0734282909360011> (2008).
81. Reynolds, B., Ortengren, A., Richards, J. B. & de Wit, H. Dimensions of impulsive behavior: personality and behavioral measures. *Pers. Individ. Dif.* **40**, 305–315 (2006).
82. Mobini, S., Grant, A., Kass, A. E. & Yeomans, M. R. Relationships between functional and dysfunctional impulsivity, delay discounting and cognitive distortions. *Pers. Individ. Dif.* **43**, 1517–1528 (2007).
83. Baumann, A. A. & Odum, A. L. Impulsivity, risk taking, and timing. *Behav. Process.* **90**, 408–414 (2012).
84. Dougherty, D. M., Marsh, D. M. & Mathias, C. W. Immediate and delayed memory tasks: a computerized behavioral measure of memory, attention, and impulsivity. *Behav. Res. Methods Instrum. Comput.* **34**, 391–398 (2002).
85. Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191 (2007).
86. Thyng, K. M., Greene, C. A., Hetland, R. D., Zimmerle, H. M. & DiMarco, S. F. True colors of oceanography: guidelines for effective and accurate colormap selection. *Oceanography* **29**, 9–13 (2016).
87. Cramer, F., Shephard, G. E. & Heron, P. J. The misuse of colour in science communication. *Nat. Commun.* **11**, 1–10 (2020).
88. Patton, J. H., Stanford, M. S. & Barratt, E. S. Factor structure of the barratt impulsiveness scale. *J. Clin. Psychol.* **51**, 768–774 (1995).
89. Frey, R., Pedroni, A., Mata, R., Rieskamp, J. & Hertwig, R. Risk preference shares the psychometric structure of major psychological traits. *Sci. Adv.* **3**, 1–14 (2017).
90. Cattell, R. B. Multivariate behavioral translator disclaimer the scree test for the number of factors. *Multivar. Behav. Res.* **1**, 245–276 (1966).
91. Hauser, T. U. et al. Increased fronto-striatal reward prediction errors moderate decision making in obsessive-compulsive disorder. *Psychol. Med.* <https://doi.org/10.1017/S0033291716003305> (2017).
92. Mayes, S. D., Calhoun, S. L., Mayes, R. D. & Molitoris, S. Autism and ADHD: Overlapping and discriminating symptoms. *Res. Autism Spectr. Disord.* **6**, 277–285 (2012).
93. Good, D. & Michel, E. J. Individual ambidexterity: Exploring and exploiting in dynamic contexts. *J. Psychol. Interdiscip. Appl.* **147**, 435–453 (2013).
94. Laureiro-Martínez, D., Brusoni, S. & Zollo, M. The neuroscientific foundations of the exploration-exploitation dilemma. *J. Neurosci. Psychol. Econ.* **3**, 95–115 (2010).
95. Dubois, M. MagDub/MFweb-data: (v1.0). Zenodo <https://doi.org/10.5281/ZENODO.6522060> (2022).
96. Dubois, M. MagDub/Mfweb-pilot_data: (v1.0). Zenodo <https://doi.org/10.5281/ZENODO.6522062> (2022).
97. Dubois, M. MagDub/MFweb-data_analysis: (v1.0.0). Zenodo <https://doi.org/10.5281/zenodo.6445661> (2022).

Acknowledgements

We thank Vasilisa Skvortsova for her help with implementing the exploration task online. M.D. is a predoctoral fellow of the International Max Planck Research School on Computational Methods in Psychiatry and Ageing Research. The participating institutions are the Max Planck Institute for Human Development and the University College London (UCL). T.U.H. is supported by a Wellcome Sir Henry Dale Fellowship (211155/Z/18/Z), a grant from the Jacobs Foundation (2017-1261-04), the Medical Research Foundation, and a 2018 NARSAD Young Investigator Grant (27023) from the Brain and Behaviour Research Foundation. T.U.H. has also 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).

Author contributions

Conceptualization, M.D., and T.U.H.; Methodology, M.D., and T.U.H.; Software, M.D.; Formal Analysis, M.D., and T.U.H.; Investigation, M.D.; Data Curation, M.D.; Writing – Original Draft, M.D., and T.U.H.; Writing – Review & Editing, M.D. and T.U.H.; Supervision, T.U.H.; Funding Acquisition, M.D., and T.U.H.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-31918-9>.

Correspondence and requests for materials should be addressed to Magda Dubois.

Peer review information *Nature Communications* thanks Christopher Chambers, Charley Wu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022