Value-free random exploration is linked to impulsivity

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 differences. Impulsivity is often described as ‘acting without thinking’ and is traditionally assessed using self-report questionnaires. It is a broad and heterogenous construct 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 disorder (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 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 Taylor 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 level. Several theoretical accounts have embraced this concept and demonstrated how increased exploration can arise due to catecholamine imbalance, how this may be implemented neurally, and how an excessive exploration can explain other behaviours observed with impulsivity, such as delay discounting.

Outside of theoretical work, however, empirical evidence so far is still sparse. Only the first few studies using ADHD patients or looking at ADHD symptoms in youths 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 bound\(^{26}\) (UCB), which take the expectation as well as the uncertainty of all possible choice options into account\(^{22\text{-}27,30}\), 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\(^{21,26,31}\). In addition, there is evidence for exploration strategies that deliberately omit all existing knowledge to choose all options equally likely, termed value-free random exploration\(^{21}\). 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\(^{2\text{-}27}\). Even though such mechanisms are suboptimal, their low computational demand has made them popular in artificial intelligence\(^{21\text{-}27}\) (i.e., \(\varepsilon\)-greedy) and we have found clear signatures in humans\(^{24}\).

Whether exploration mechanisms (and if so, which ones) are altered in impulsivity remains unknown. However, this is of critical importance as recent animal\(^{24}\) and human\(^{23}\) 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) functioning\(^{24}\), which is a neurotransmitter that has been repeatedly suggested to be critical in impulsivity disorders\(^{4,21,27,34\text{-}36}\).

In this study, we put the large body of theoretical work to test and explored 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\(^{8\text{-}10}\). We ran a big data dimensional study using our recently developed exploration task\(^{23}\), 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 effort\(^{2\text{-}4}\) and that ADHD is associated to increased value-free random exploration\(^{23}\), we tested our hypothesis that it is specifically value-free random exploration (captured by our model parameter \(e\)) 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\(^{44\text{-}46}\), to depression, anxiety and anhedonia\(^{46\text{-}49}\)) using data-driven methods.

**Results**
To capture different forms of exploration, we used our previously lab-validated Maggie’s Farm task\(^{25}\) (cf. Fig. 1), which is essentially a 3-armed variant of the Horizon task\(^{24}\). 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 initials 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 computational 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 Task\(^{24}\). 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\(^{24\text{-}25}\). 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: \(W = 110037, p < 0.001\), 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 forgo 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 (\(V = 34420, p < 0.001\),
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.
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 and novelty exploration) as captured by the winning model (comparison of BIC average scores: Thompson+ε greedy vs Thompson model: V = 30589, p < 0.001, r 2 = 0.835; Hypothesis 1.3. in Table 1; Thompson+ε greedy vs UCB+ε greedy model: V = 46440, p < 0.001, r 2 = 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+ε greedy) 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 = 33567, p < 0.001, r 2 = 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 impulsiveness 51. 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 η p 2 = 0.03; horizon main effect: F(1,578) = 113.614, p < 0.001, η p 2 = 0.164; BIS-by-horizon interaction: F(1,578) = 0.773, p = 0.380, η p 2 = 0.001) and on the ε-greedy parameter (BIS main effect: F(1,578) = 17.454, p < 0.001, η p 2 = 0.029; horizon main effect: F(1,578) = 125.804, p < 0.001, η p 2 = 0.179; BIS-by-horizon interaction: F(1,578) = 0.848, p = 0.772, η 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.
Step 2.2. Are ADHD symptoms linked to value-free random exploration?

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 exploration 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 \( \epsilon \)-greedy parameter (\( r(578) = 0.157, \ p_{unc} < 0.001 \), Fig. 3e; accounting for age and IQ; \( r(573) = 0.115, \ p_{unc} = 0.006 \); 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 (\( r(578) = 0.134, \ p_{unc} = 0.012 \); Hypothesis 2.2. in Table 1).
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 = 8). 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 ‘futuring’ or ‘foreshotting’.

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): $r(578) = 0.198$, $p < 0.001$, $\eta^2 = 0.01$, $p_{unc} < 0.001$). When correcting for age and IQ: $r(573) = 0.159$, $p < 0.001$, $\eta^2 = 0.01$; frequency of low-value bandit: $r(578) = 0.205$, $p < 0.001$, $\eta^2 = 0.01$; accounting for age and IQ: $r(573) = 0.165$, $p < 0.001$, $\eta^2 = 0.01$). We did not observe any robust association with the BIS non-planning subscore when correcting for age and IQ (using the $\epsilon$-greedy parameter: $r(578) = 0.120$, $p_{corr} = 0.012$, $p_{unc} = 0.004$ [accounting for age and IQ: $r(573) = 0.058$, $p_{corr} = 0.051$, $p_{unc} = 0.167$]; using the low-value bandit: $r(578) = 0.128$, $p_{corr} = 0.006$, $p_{unc} = 0.002$ [accounting for age and IQ: $r(573) = 0.065$, $p_{corr} = 0.364$, $p_{unc} = 0.121$]) or with the BIS attentional subscore (using the $\epsilon$-greedy parameter: $r(578) = 0.095$, $p_{corr} = 0.067$, $p_{unc} = 0.022$ [accounting for age and IQ: $r(573) = 0.067$, $p_{corr} = 0.331$, $p_{unc} = 0.11$]; using the low-value bandit: $r(578) = 0.086$, $p_{corr} = 0.118$, $p_{unc} = 0.039$ [accounting for age and IQ: $r(573) = 0.054$, $p_{corr} = 0.592$, $p_{unc} = 0.197$]). 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 $\epsilon$-greedy parameter: $r(578) = 0.205$, $p_{corr} < 0.001$, $p_{unc} < 0.001$, Fig. 3g [accounting for age and IQ: $r(573) = 0.152$, $p_{corr} = 0.001$, $p_{unc} < 0.001$]; using the low-value bandit: $r(578) = 0.193$, $p_{corr} < 0.001$, $p_{unc} < 0.001$, Fig. 3h [accounting for age and IQ: $r(573) = 0.136$, $p_{corr} < 0.002$, $p_{corr} < 0.001$]) but not reliably with the ASRS inattention subscore (using the $\epsilon$-greedy parameter: $r(578) = 0.087$, $p_{corr} = 0.074$, $p_{unc} = 0.037$ [accounting for age and IQ: $r(573) = 0.061$, $p_{corr} = 0.292$, $p_{unc} = 0.146$]; using the low-value bandit: $r(578) = 0.085$, $p_{corr} = 0.082$, $p_{unc} = 0.041$ [accounting for age and IQ: $r(573) = 0.057$, $p_{corr} = 0.348$, $p_{unc} = 0.174$]). 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.

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 $\epsilon$-greedy parameter (correcting for multiple comparison using Bonferroni correction across 4 parameters x 3 factors, i.e., N = 12; $r(578) = 0.257$, $p_{unc} < 0.001$, $p_{corr} < 0.001$, Fig. 5a; accounting for age and IQ: $r(573) = 0.204$, $p_{unc} < 0.001$, $p_{corr} < 0.001$) as well as between the impulsivity factor score and the low-value bandit frequency (correcting for multiple comparison using Bonferroni correction across 3 bandits x 3 factors, i.e., N = 9; $r(578) = 0.247$, $p_{unc} < 0.001$, $p_{corr} < 0.001$, Fig. 5b; accounting for age and IQ: $r(573) = 0.191$, $p_{unc} < 0.001$, $p_{corr} < 0.001$). 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 $\eta$ ($r(578) = 0.051$, $p_{corr} = 0.223$; accounting for age and IQ: $r(573) = 0.058$, $p_{corr} = 0.167$; Fig. 5e), nor the prior variance $\epsilon_0$ ($r(578) = 0.02$, $p_{corr} = 0.631$; accounting for age and IQ: $r(573) = 0.01$, $p_{corr} > 0.816$; Fig. 5e), or the prior mean $Q_0$ ($r(578) = -0.006$, $p_{corr} = 0.891$; accounting for age and IQ: $r(573) = -0.016$, $p_{corr} < 0.701$; 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, 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 (UIS 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 x factors, i.e., N = 12). Our exploratory analysis revealed that the anxious-depression factor correlated positively with the novelty bonus $\eta$ ($r(578) = 0.14$, $p_{corr} = 0.008$, $p_{unc} < 0.001$, Fig. 5c; accounting for age and IQ: $r(573) = 0.126$, $p_{corr} = 0.03$, $p_{unc} = 0.002$). None of the other parameters was linked to the anxious-depression factor (e.g. $r(578) = -0.047$, $p_{corr} = 1$, $p_{unc} = 0.262$; accounting for age and IQ: $r(573) = -0.078$, $p_{corr} = 0.73$).
accounting for age and IQ: $r(573) = 0.17$, $p_{unc} = 0.017$; accounting for age and IQ: $r(573) = 0.107$, $\rho_{cor} = 0.121$, $\rho_{unc} = 0.01$; $Q_0$: $r(578) = 0.073$, $\rho_{cor} = 0.925$, $\rho_{unc} = 0.077$; accounting for age and IQ: $r(573) = 0.052$, $\rho_{cor} = 1$, $\rho_{unc} = 0.209$; 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 ($r(578) = 0.19$, $p_{unc} < 0.001$; Fig. 5d) nor in the behaviour (low-value bandit: $r(578) = 0.054$, $\rho_{cor} = 0.09$, $\rho_{unc} = 0.03$; accounting for age and IQ: $r(573) = 0.037$, $\rho_{cor} = 0.732$, $\rho_{unc} = 0.081$; 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 AQ10 total score with the low-value bandit frequency: $r(578) = 0.025$, $p = 0.545$; accounting for age and IQ: $r(573) = 0.022$, $p = 0.392$; with the $e$-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 CFS 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 $e$-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.
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 = 0.565, uncorrected p = 0.141; accounting for age and IQ: r(573) = 0.073, p = 0.08; σ0: r(578) = 0.012, p = 0.767; accounting for age and IQ: r(578) = 0.257, p < 0.001).
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 (r; r(573) = 0.083, p < 0.001; accounting for age and IQ: r(573) = 0.072, p < 0.001; accounting for age and IQ: r(573) = 0.029, p = 0.019; punc < 0.001; accounting for age and IQ: r(573) = 0.101, p < 0.001).

Step 4.4. Analysis of 2nd winning model

Value-free random exploration (captured by the ɛ-greedy parameter) was similar in the 1st winning model (Thompson+ɛ+η) and in the 2nd winning model (UCB+ɛ+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., ɛ-greedy parameter) as captured by 2nd winning model, and impulsivity. We observed a significant association between ɛ and the BIS total score (r(578) = 0.155, p < 0.001; controlling for age and IQ: r(573) = 0.101, p < 0.001), between η 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 ɛ 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, ηp² = 0.061; horizon main effect: F(1,578) = 113.474, p < 0.001, ηp² = 0.164; impulsivity-by-horizon interaction: F(1,578) = 0.059, p = 0.808, ηp² = 0) and on the ɛ-greedy parameter (impulsivity main effect: F(1,578) = 40.572, p < 0.001, ηp² = 0.066; horizon main effect: F(1,578) = 126.418, p < 0.001, ηp² = 0.179; impulsivity-by-horizon interaction: F(1,578) = 2.906, p = 0.089, ηp² = 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\(^2\) that all approximate an optimal exploration strategy, which is intractable in open-ended decision problems. In our current data, we confirmed that participants utilized a multitude of strategies during 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 ɛ-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 and novelty-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 clear. Based on previous theoretical\(^2\),\(^3\),\(^4\) and some experimental work\(^2\),\(^5\),\(^6\), exploration is believed to be increased in impulsivity 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 abilities might of 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 construct\(^8\). 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 norepinephrine functioning\(^9\), a neurotransmitter which plays an important role in impulsivity-related disorders such as ADHD\(^1\),\(^2\),\(^4\),\(^5\),\(^9\),\(^10\), which could be a potential mechanism. Previous findings that linked norepinephrine functioning to what is traditionally seen as motor impulsivity support this notion\(^9\). This form of exploration may be also related to brain circuits generally seen to be linked to norepinephrine functioning.
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. Here, the main benefit of value-free random exploration is that it does not require demanding computations, allowing exploration even with restrained neural resources or a limited ability/willingness to engage with mentally effortful computations. 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 knowledge, 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 expectations. Our findings of such exploration heuristics are also well aligned with recent findings showing that limiting cognitive resources impacts the use of exploration strategies, and shifts in exploration strategies can be induced by applying constraints such as time pressure. 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. 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 OC-RR 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, 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. It is believed that this is because exploration aids in overcoming long-term uncertainty, and an uncertainty aversion is commonly reported in anxiety. 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 or larger decision spaces), 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 insights.

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 and a well-defined dimensional approach, 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 developed, 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 estimated.

To induce changes in exploration, similar to the horizon task and our previous studies, we manipulated the number of samples they could draw from a given set of bandits. 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 used. 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. In fact, exploration in the short horizon has previously been observed in humans.

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 forward. Each bandit is from one of four generative groups characterised by different means μ and number of initial samples, following the same procedure as other studies. 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 confound. The samples of each bandit are then sampled from a normal distribution with a fixed sampling variance \(\sigma^2(\mu, 0.5)\), truncated to \([2, 10]\), and rounded to the closest integer. Each mean \(\mu\) was sampled from \(\mu_{\text{overall}} \sim \mathcal{N}(\mu_{\text{overall}}-1.4, 1)\), with an ‘overall mean’ \(\mu_{\text{overall}}\) specific to each bandit type. The overall mean was computed similarly to previous studies. 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 similar to previous studies, one a certain bandit provides 1 initial sample. Even though the absolute range of reward is similar to previous studies, one a certain bandit does not provide any initial samples and the low-value bandit does not take expected values into account.

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 Scale \(2(\text{ASRS})\) and the Barratt Impulsiveness Scale \(2(\text{BIS})\). In addition, we collected further questionnaires to investigate additional psychiatric dimensions (cf. Data analysis, Step 3). These entail the Liebowitz Social Anxiety Scale \(2(\text{LSAS})\), the State-Trait Anxiety Inventory \(2(\text{STAI-Y2})\), Intolerance of Uncertainty Scale \(2(\text{IUS})\), Obsessive-Compulsive Inventory-Revised \(2(\text{OCI-R})\), and Zung’s Self-Rating Depression Scale \(2(\text{SDS})\), in accordance with similar previous approaches \(2(\text{ICAR})\), as well as the Cognitive Flexibility Scale \(2(\text{CFS})\) and the Autism spectrum Quotient \(2(\text{AQ-10})\). To control for confounding factors, such as intelligence and medication, participants additionally completed the International Cognitive Ability Resource sample test \(2(\text{ICAR})\) 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 questions. 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/MWeb-app).

Registered report | https://doi.org/10.1038/s41467-022-31918-9
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 Software\textsuperscript{85} 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 'mrcolourmap' was used for the figures\textsuperscript{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: \( N(\text{m1, std1}) \) and \( N(\text{m2, std2}) \) 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 = \( \begin{bmatrix} 1 & R; R,1 \end{bmatrix} \). 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 youths\textsuperscript{22}, and for Step 3 from previous big data studies\textsuperscript{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 level\textsuperscript{46} (cf. Supplementary Fig. 2b), the first draw mean reaction time was faster than 1500 ms (based on our pilot data and previous study\textsuperscript{23}) 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\textsuperscript{47,48}. 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 task\textsuperscript{23} 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\textsuperscript{22,23}. In line with previous studies investigating horizon-dependent exploration\textsuperscript{22,23,34}, 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\textsuperscript{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: \( \text{CIM} = [0.047,0.165] \), effect size: Cohen’s \( d = 0.459, 95\% \) confidence interval of the effect size \( \text{CIS} = [-0.727, -0.195] \). This is reflected by the frequency of picking the high-value bandit, which we predicted to decrease in the long horizon (pilot data: \( t(60) = 8.45, p < 0.001, 95\% \text{CIM} = [-6.92, 11.211], d = -1.082, 95\% \text{CIS} = [-1.407, -0.769] \). 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_{BM} = [-1.568, -0.416], d = 0.441, 95%CI_{EF} = [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_{BM} = [10.184, 0.281], d = -1.232, 95%CI_{EF} = [-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_{BM} = [-11.78, -6.954], d = 1.099, 95%CI_{EF} = [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_{BM} = [0.059, 0.112], d = -0.835, 95%CI_{EF} = [-1.134, -0.543]). As observed in previous studies22,23, 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_{BM} = [-0.245, -0.191], d = 2.061, 95%CI_{EF} = [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 studies22, 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_{BM} = [-72.866, -48.946], d = 1.304, 95%CI_{EF} = [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 \eta-\text{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_{BM} = [-0.058, -0.014], d = 0.413, 95%CI_{EF} = [0.152, 0.679]). Similarly, we expected the novelty bonus q, 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_{BM} = [-1.265, -0.822], d = 1.207, 95%CI_{EF} = [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 \epsilon 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 test28. 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 impulsiveness22, and has already been used in online studies30,31. We looked at the correlation between the BIS total score and the low-value bandit frequency, and between the BIS total score and the \epsilon-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 heterogenous construct22,31, in Stage 2 we performed an exploratory analysis of the three subdomains of BIS (i.e., attentional, motor, and non-planning behaviour) similarly to previous studies30. 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 \epsilon-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 \epsilon-greedy parameter. It allows a definitive answer to the hypothesis whether ADHD symptoms are linked to value-free random exploration12,21,23. 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 population79. 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 \epsilon-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 studies22,31. Factor analysis was conducted using the fa() function from the Psych package in R, with an oblique rotation (oblimin); we drew 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 Cattel's criterion\textsuperscript{10}, using the Cattel-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\textsuperscript{40–42,50,51}, we expected to retrieve a depression / anxiety dimension, on which OCD and uncertainty intolerance traits load onto (OCIR and IUS questionnaires) respectively. Indeed, previous research has found that impulsive and compulsivity only show a modest overlap\textsuperscript{63}, which is also why previous studies that used factor analyses have found that these items load onto different factors\textsuperscript{60,62}. Previous studies have demonstrated increases in exploration in OCD patients\textsuperscript{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 impulsivity dimension and each exploration free parameter (depending on the model). Similarly, studies have demonstrated abnormality in exploration in patients with depression\textsuperscript{11}, anxiety\textsuperscript{12} and other disorders related to avoidance of uncertainty\textsuperscript{45}. 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 ADHD\textsuperscript{16}. An association between the autism scale 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\textsuperscript{13,14}.

**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\textsuperscript{39}]. The pilot data are available at Github [https://github.com/MagDub/Mfweb-pilot-data and Zenodo: https://doi.org/10.5281/zenodo.6522062\textsuperscript{39}]. 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\textsuperscript{39}].

**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, J. M. 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. Neuropsychiatric 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. Neuropsy. 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).
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 fronto-striatal myelination trajectories. Nat. Neurosci. 22, 992–999 (2019).

64. Deza, 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. Bennett, D., Sutcliffe, K., Tan, N. P. J., Smillie, L. D. & Bode, S. A neural and behavioral trade-off between value and uncertainty underlies exploratory decisions in obsessive-compulsive disorder. J. Exp. Psychol. Gen. 150, 739–755 (2020).

66. Aberg, K. C., Toren, I. & Paz, R. Should we control? The neuroscientific foundations of the exploration-exploitation dilemma. J. Psychol. Interdiscip. Appl. 147, 435–453 (2013).

67. 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).

68. 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).

69. Hill, G. D. & Michel, E. J. Individual ambidexterity: Exploring and exploiting in dynamic contexts. J. Psychol. Interdiscip. Appl. 147, 435–453 (2013).

70. Laureiro-Martínez, D., Brusoni, S. & Zollo, M. The neuroscientific foundations of the exploration-exploitation dilemma. J. Neurosci. Psychol. Econ. 3, 95–115 (2010).

71. Dubois, M. MagDub/MFweb-data: (v.1.0). Zenodo https://doi.org/10.5281/ZENODO.6522060 (2022).

72. Dubois, M. MagDub/MFweb-pilot_data: (v1.0). Zenodo https://doi.org/10.5281/ZENODO.6522062 (2022).

73. 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)/17/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