Novelty-seeking impairment in addiction

Irene Cogliati Dezza[1,2,*], Xavier Noel[3], Axel Cleeremans[1], Angela J. Yu[4]

[1]Centre for Research in Cognition & Neurosciences, ULB Neuroscience Institute, Université Libre de Bruxelles, Belgium

[2]Department of Experimental Psychology, Faculty of Brain Sciences, University College London, London, UK

[3]Faculty of Medicine, Université Libre de Bruxelles, Belgium

[4] Department of Cognitive Science, University of California San Diego, United States

KEY WORDS

Novelty-seeking, Information-seeking, Exploration, Problem Gambling, Addiction, Reinforcement learning

Supplemental Information: Yes

ABSTRACT

Information-seeking is an important aspect of human cognition. Despite its adaptive role, we have rather limited understanding on the mechanisms that subtend information-seeking in healthy individuals and in psychopathological populations. Here, we aim to formalize the computational basis of healthy human information behavior, as well as how those components may be compromised in behavioral addiction. We focus on gambling disorder, a form of addiction without the confound of substance consumption. We investigate and model human behavior using a novel decision-making task and a novel reinforcement learning model. Our results indicate that healthy information behavior is motivated by both novelty and general knowledge (or information). In contrast, problem gamblers have a specific deficit in novelty processing in choice behavior, but not in general information. This finding sheds light both on the computational mechanisms underlying healthy human information behavior, and on how they can go awry in behavioral addiction.

INTRODUCTION

Humans spend considerable time in seeking information. Information-seeking is therefore an essential aspect of human cognition that supports healthy decision-making and goal-directed processing [1 2 3]. Here, we argue that human information behaviors can be driven by both the desire for novelty and the level of general knowledge or information: how much the decision maker wants to know about novel information available in the environment and how many pieces (or bits) of information she has already acquired in the past. This distinction can be relevant in understanding psychopathologies such addiction where individuals are trapped into the same behavioral routines (e.g., gambling, substance intake, binge eating) despite the negative consequences associated with them (e.g., financial loss, healthy problems [4]). Engaging in these repeated behaviors may be due to an inability to either represent and implement novel behavioral patterns (novelty) or to higher weights given to already known ones (general information). Here, we aim to formalize the computational basis of healthy human information behavior and how it may be compromised in gambling disorder, a form of addiction without the confound of substance consumption.

Information-seeking behaviors are often contraposed to the human tendency of maximizing immediate benefits. A decision-maker who is trying to find out the best restaurant in town may try out all different available options in order to obtain information on the potential benefit of each restaurant, but this information search may be costly or result in unpleasant experiences. Yet, information-seeking has a significant role in human daily activities (e.g., exploring, reading, searching, asking questions etc..). A number of studies [5], including our own [6 7], have suggested that healthy humans finely balance the urge for immediate reward vs. longer-term information gain during sequential decision-making (i.e., exploration-exploitation trade-off). Appropriately balancing this tension is a necessary tool for navigating in a world fraught with uncertainty and changeable dynamics. Resolving this tension plays a key role, for instance, in foraging problems [8], complex decisions in the human daily life [9], and even boosting the performance of artificial agents [10 11]. And, deficiency in its resolution has been observed in psychopathological conditions such as addiction [12] and depression [13]. Previous work, however, has not specifically considered the importance of *novelty* seeking and general *information* seeking in driving human information behavior under repeated scenarios. And, whether/how alterations in these components can independently contribute to certain pathological conditions such addiction. Investigating novelty seeking and general information seeking in addiction could help in elucidating the mechanisms underlying decision-making impairments often observed in this population [14 15] and ultimately developing more efficient clinical interventions (e.g., behavioral therapies, neural stimulations or pharmacological interventions may specifically target the implementation of novel behavioral routines or devaluation of preexisting ones).

To study information behavior under repeated scenarios, we adopt a modified version of a popular task (i.e., the multi-armed bandit) often used to study sequential learning and decision-making behavior. In

the bandit task, the decision-maker must make repeated choices among options characterized by initially unknown reward distributions. Each choice can be driven either by a more myopic desire to maximize immediate gain (based on knowledge gained from previous choices and outcomes) or by a more long-term goal of being more informed about all the options. In these repeated scenarios, however, the more the decision-maker tends to choose the most rewarding options, the more those rewarding options tend to be (anti-) correlated with the amount of (remaining) information that can be obtained [16] [5]. Accordingly, these classical decision-making tasks make it difficult to quantify exactly how much reward and information each contribute independently to choices [5]. Here, we therefore adopt a novel variant of the bandit task [6], inspired by [5], which has an initial phase of forced choices that carefully controls for reward and information associated with each option. Thus, it dissociates the relative contribution of reward and information as motivating factors in choice behavior.

To investigate the role that novelty-seeking and general information-seeking play in human information behavior, we introduce a new learning and decision-making model for the bandit task that makes it possible to quantitatively separate out the importance of novelty versus general information in driving human information behavior.

To explore the role that novelty and general information-seeking play in addictive disorders, we compare a pathological gambling group to a healthy control group. The focus on problem gambling, as opposed to substance abuse, allows us to target the behaviors underlying addiction without the confounding effects of chronic substance use and abuse [17]. The inclusion of the problem gambling group not only allows us to identify the processes and mechanisms that are altered in behavioral addiction, but also to reveal modular processes that operate semi-autonomously in the healthy brain and thus can independently break down in pathological conditions. Indeed, as our study will demonstrate, problem gamblers have a specific deficit in novelty processing in choice behavior, but not in general information seeking or in reward seeking.

METHODS AND MATERIAL

Participants

Forty (40) unmedicated problem gamblers (PG's; mean age = 30.1, mean 4 female) and twenty-two (22) healthy controls (mean age = 29, 4 female) were recruited from the local communities. The sample size of both groups was based on previous studies [18] [6]. We excluded participants having co-morbidity with substance abuse and alcohol use disorder or undergoing psychological and pharmacological treatment and with injuries involving the brain (Table 1; Supplement). Gamblers were selected among those who were gambling at least once per week, while healthy controls were those without gambling experience in the year preceding experimental participation (Table 1; Supplement). The two groups statistically differed only in terms of gambling severity and years of education (years of education did not correlate with any of the behavioral measures considered in this study and removing problem gamblers with lower years of education did not change the main results reported in the text).

	Problem Gamblers n=40	**Healthy Control** n=22	**Test Statistic**
Gender (M/F)	36|4	18|4	$p = 0.601$
Age	30.1(9.3)	29(6.6)	$p = 0.982$
Years of Education	14.7(2)	16.2(2.2)	$p = 0.037$ *
IQ (WAIS block)	8.4(2.6)	9.3(1.9)	$p = 0.131$
Gambling Severity (CPGI)	8.8(6.1)	0	$p < 10^{-10}$ *
Alcohol use (AUDIT)	4.6(3.9)	5.3(3.1)	$p = 0.48$
Drug use (DAST)	0.225(0.423)	0.227(0.429)	$p = 0.992$
Smoking dependence (FTND)	n=4	n=1	NA
Memory Capacity (WAIS)	10.3(3.5)	9.7(4.1)	$p = 0.483$
Attentional Control (ACS)	35.4(9)	37.5(7)	$p = 0.312$
Depression (BDI)	5.6(4.9)	4.2(4.8)	$p = 0.137$
Anxiety (STAI-S)	35.1(10.9)	37.9(9.5)	$p = 0.173$
Anxiety (STAI-T)	39.6(12.4)	43.1(11)	$p = 0.2$

Positive Mood (PANAS)	35.4(6.3)	36.3(5.3)	$p = 0.701$
Negative Mood (PANAS)	21.1(7.9)	19.8(4.8)	$p = 0.808$

Table 1. ***Demographic information.*** Mean and standard deviations are shown for each measure. For each comparison, we ran a two-sampled t test, except for gender comparison where chi-squared test was used. The two groups differ only in terms of gambling severity (with no gambling problems reported in the control group) and years of education as often reported in the literature [18]. Note: WAIS IV-Wechsler Adult Intelligence Scale (the block-design component of the WAIS is the subset that best predicts performance IQ [19]); CPGI- Canadian Problem Gambling Index ; AUDIT - Alcohol Use Disorders Identification Test; DAST - Drug Abuse Screening Test; FTND - Fagerström Test for Nicotine Dependence; ACS - Attentional Control Scale; BDI- Beck Depression Inventory; STAI-S - State version of the State-Trait Anxiety Inventory; STAI-T - Trait version of the State-Trait

Behavioral Task

Participants performed 162 games of a decision-making task [6] which allow to orthogonalize the influence of reward and information on sequential choices [5] (Fig. 1a, Supplement). On each game, participants were initially instructed about which option (deck of cards) to choose from on each trial (*forced-choice task*; Fig. 1b) for six consecutive trials, after which they were free to choose from any of the options (*free-choice task*; Fig. 1c) so as to maximize their final gain. When selected, each deck provides a reward (from 1 to 100 points) generated from a truncated Gaussian distribution with mean set to either 30 (Low Reward Context) or 50 points (High Reward Context; Supplement). The generative mean for each option was stable within a game, and varied across games. The "true", generative mean reward value of the three decks had statistically similar values in 50% of the games.

In order to perform the decision task, participants should alternate exploration vs. exploitation choices. During exploration, participants can either choose at random (undirected exploration) or they can direct their exploration toward the most informative or novel alternative (information-driven exploration) [5]. In order to dissociate between these two behavioral patterns, we implemented two conditions in the forced-choice task [5]. Participants were either forced to choose each deck 2 times (*equal information condition*), or to choose one deck 4 times, another 2 times, and the third 0 times (*unequal information condition*). Information-driven exploration was then defined as choosing options that had never been sampled during the forced-choice task (in the unequal information condition), while undirected exploration was defined as choosing options associated with the lowest gain (in the equal information condition; Supplement) [6]. In 50% of the games, participants played the unequal information condition. The order of card selection was randomized in both information conditions, as was the occurrence of the equal and unequal information conditions.

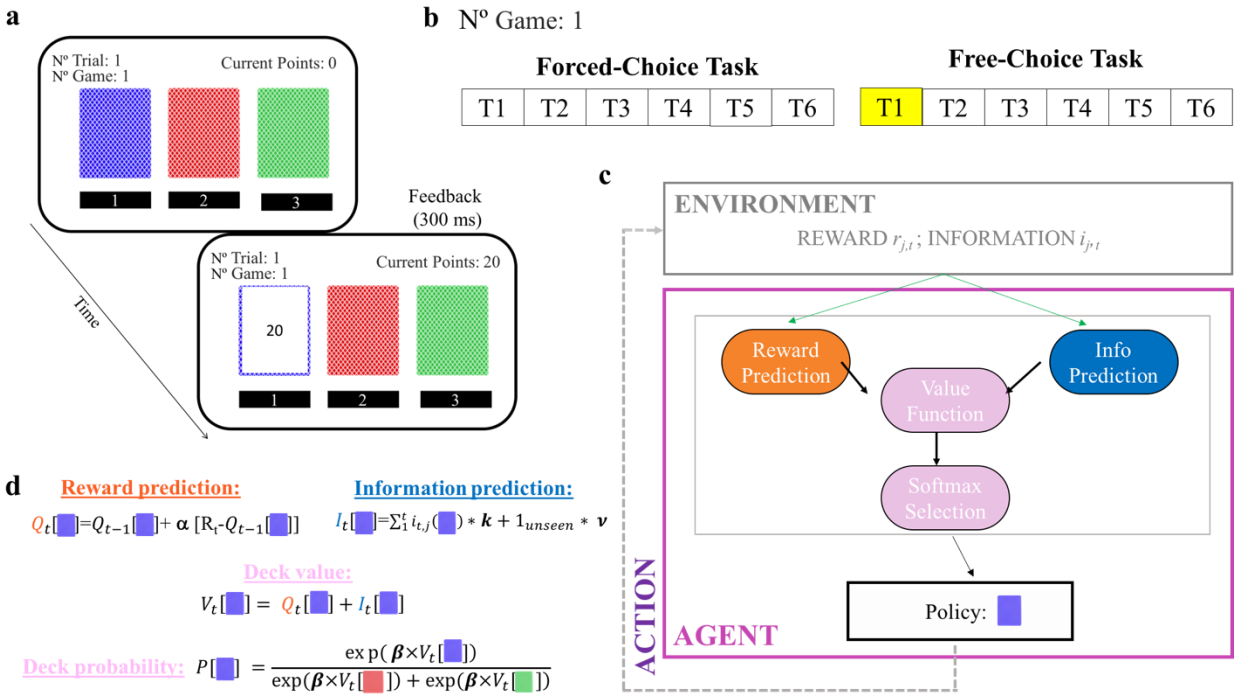


Figure 1. ***Behavioral paradigm and RL model.*** **a)** On each trial participants make choices between 3 decks of cards. After selecting a deck, the card turned and revealed the points associated with the selected option, between 1 and 100 points. Participants were instructed to attempt to maximize the total points earned at the end of the experiment. **b**) On each game, participants faced two phases: the forced-choice task (6 consecutive trials) and the free-choice task (between 1 and 6 trials; participants were not aware of the task length). In the forced-choice task, participants were forced to choose a preselected deck. In the free-choice task, participants made their own choices among the same three decks of cards displayed during the forced-choice task. At this stage, the points displayed on the screen were added to the participants' total score after each choice. In the first free-choice trial (in yellow), reward and information are orthogonalized, so enabling the distinction between undirected and information-driven exploration [5]. **c**) The reinforcement learning (RL) model uses reward and information in the environment to learns expected reward values (reward prediction) and information prediction on a trial-basis. Next, it combines these predictions into a value function in order to determine choices stochastically using a softmax function. **d)** On each trial, novelty-knowledge RL (nkRL) model updates its reward prediction using the delta learning rule. Using this rule, novel reward predictions are made by integrating new outcomes (i.e., the points obtained after a trial) to previous reward predictions. The degree by which this integration is achieved is controlled by the learning rate α. With small α the model slowly updates its estimate in response to new outcomes, whereas with higher values the model integrates new reward outcomes more rapidly. Next, it computes information prediction as sum of general information and novelty term. The general information term describes the level of general information participants have about the selected option. In other words, it defines how many information bits have been acquired by the decision-maker. The novelty term, on the contrary, captures the desire for novelty. In other words, it defines how much the decision-maker wants to know about a novel option. Reward and information predictions are then combined in the deck value; and a choice is made by entering the deck values into the softmax function. Under lower β, choices are more random, thus the decision maker's choices are driven by decision noise. Model's parameters are shown in bold.

Computational modelling

During the execution of the decision-making task, humans generate both reward and information predictions which jointly guide choices [6]. This can be formalized using a computational model which follows a reinforcement-learning (RL) routine (Fig. 1c). In particular, the RL model learns expected reward values and information on a trial-basis, and combines them into a value function in order to determine choices stochastically [6] (Supplement). In order to investigate the nature of information valuation in healthy participants and problem gamblers, we implement a novel computational model which combine reward and information evaluation as previous formulations [6], but also dissociates between novelty and general information during information valuation. Our model, the "novelty-knowledge RL" (nkRL) computes expected reward values using the delta learning rule [20] (Fig. 1d). In addition, as in previous RL versions, the predicted value of a deck is a combination of reward and information prediction (Eq. S3). However, the information term is split into two components resulting in the following value function:

$$V_{t,j}(c) = Q_{t+1,j}(c) + \sum_1^t i_{t,j}(c) * k + 1_{unseen} * \nu \quad (1)$$

where $Q_{t,j}(c)$ is the expected reward value for trial *t* and game *j* . $\sum_1^t i_{t,j}(c)$ is the number of information bits acquired until trial t ($i_{t,j}$ is either 1, if the option has been selected in a trial, or 0 otherwise). k is the knowledge (or general information) parameter which defines the weight toward previous acquired information. $1_{unseen} * \nu$ captures the desire for novelty, where ν is the novelty bonus given uniquely to options never selected in the past trial history. Lastly, a choice is made by entering deck values into the softmax function [21] (Eq. S4), where the decision policy is controlled by the inverse temperature β (Fig. 1d). NKRL can shed light on the processes that underpin information valuation in both healthy participants and problem gamblers by distinguishing the effects of reward-seeking and information valuation on choices (β vs. k, ν), and of novelty-seeking and general information-seeking (ν vs. k). The model's parameters were estimated by fitting nkRL to trial-by-trial participants' free choices (Supplement).

RESULTS

Information behavior in healthy individuals and problems gamblers

We first investigated information-seeking behavior in both groups by looking at the first free-choice trials (namely, the one trial where we can be sure that information and experienced reward are uncorrelated [5]; Supplement). We focused on those games where information-driven exploration can be quantified i.e., unequal information condition. We classified choices as *information-driven* when participants selected options that had never been sampled during the forced-choice task, as *reward-driven* when participants chose the experienced decks with the highest average of points (regardless of the number of times that deck had been selected during the forced-choice task) and *other* when the classification did not meet the previous

criteria (e.g., choosing a familiar deck with a relatively low average reward value). In particular, we calculated the number of trials in which information-driven, reward-driven and other strategy was adopted by each subject and we averaged those estimates across trials. This gave the averaged probability of each strategy adopted by each subject during the task. A 2 (Group: PGs, Controls) X 2 (Strategy: Reward-driven, Information-driven, Other) between-subject ANOVA revealed an effect of strategy $F(2,180) = 65.47$, $p < 10^{-15}$ and a strategy X group interaction $F(2,180) = 8.92$, $p < 10^{-3}$, whereas the effect of group was not significant ($p > 0.05$). A post-hoc comparison revealed a decrease in information-driven exploration in problem gamblers compared to controls (M_{PG} =0.377, SD_{PG} = 0.19; M_{CON} =0.524, SD_{CON} = 0.258; p = 0.025) and an increase in reward-driven choices in the gambling group compared to healthy group (M_{PG} =0.476, SD_{PG} = 0.151; M_{CON} =0.364, SD_{CON} = 0.187; p = 0.021; Figure 2a). *Other* choices did not differ between the two groups ($p = 0.08$). The decrease in information-driven exploration was independent of the reward context (Supplement; Figure S1) and on the reward condition participants were in (Supplement). Moreover, the impairment in exploration was absent when the exploratory strategies were not motivated by an informative drive (i.e., undirected exploration; Figure 2b, Supplement). Further, the increase in reward-driven choices only arise when information was delivered unequally across options (Figure 2b; Supplement).

Next, we extended the analysis to all free-choice trials by investigating participants' preferences toward most-, mid-, and least-sampled options. Here, we considered both equal and unequal information conditions altogether. In particular, we computed the relative frequency of choosing options whose outcome was experienced the most (Most-Sampled), the least (Least-Sampled) and intermediate (Mid-Sampled) number of times during previous trial history. A 2 (Group: PGs, Controls) X 3 (Choice: Most-Sampled, Least-Sampled, Mid-Sampled) between-subject ANOVA revealed an effect of Choice $F(2,180) = 338.6$, $p < 10^{-15}$ and a Choice X Group interaction $F(2,180) = 8.45$, $p < 10^{-3}$, whereas the effect of Group was not significant ($p > 0.05$). A post-hoc comparison revealed an increase in Most-Sampled choices in problem gamblers compared to controls (M_{PG} =0.569, SD_{PG} = 0.115; M_{CON} =0.507, SD_{CON} = 0.068; $p = 0.008$) and a decrease in Least-Sampled choices (M_{PG} =0.273, SD_{PG} = 0.092; M_{CON} =0.332, SD_{CON} = 0.065; $p = 0.006$; Figure 2c), while Mid-Sampled choices did not differ between the two groups ($p > 0.05$). Overall, problem gamblers showed increased preferences toward frequently selected options in the previous trial history, regardless of the associated outcome, at the expense of novel alternatives, compared to healthy controls. However, both groups exhibited a higher overall preference toward frequently selected options compared to options they were more ignorant about ($p < 10^{-4}$; Figure 2c, S2; Supplement). Interestingly, this tendency underwent a reversal in the control group, which preferred novel options over familiar options during the first free-choice trials (M_{least}= 0.536, SD_{least}= 0.258; M_{most}= 0.262, SD_{most}= 0.157; $p < 10^{-3}$; Figure 2e). This result suggests

that healthy controls are driven by novelty under certain conditions. This "novelty-familiarity" shift, however, was absent in problem gamblers where novel (M= 0.386, SD = 0.193) and frequently selected options (M= 0.344, SD= 0.131) were chosen at the same rate during the first free-choice trial ($p > 0.05$; Figure 2f).

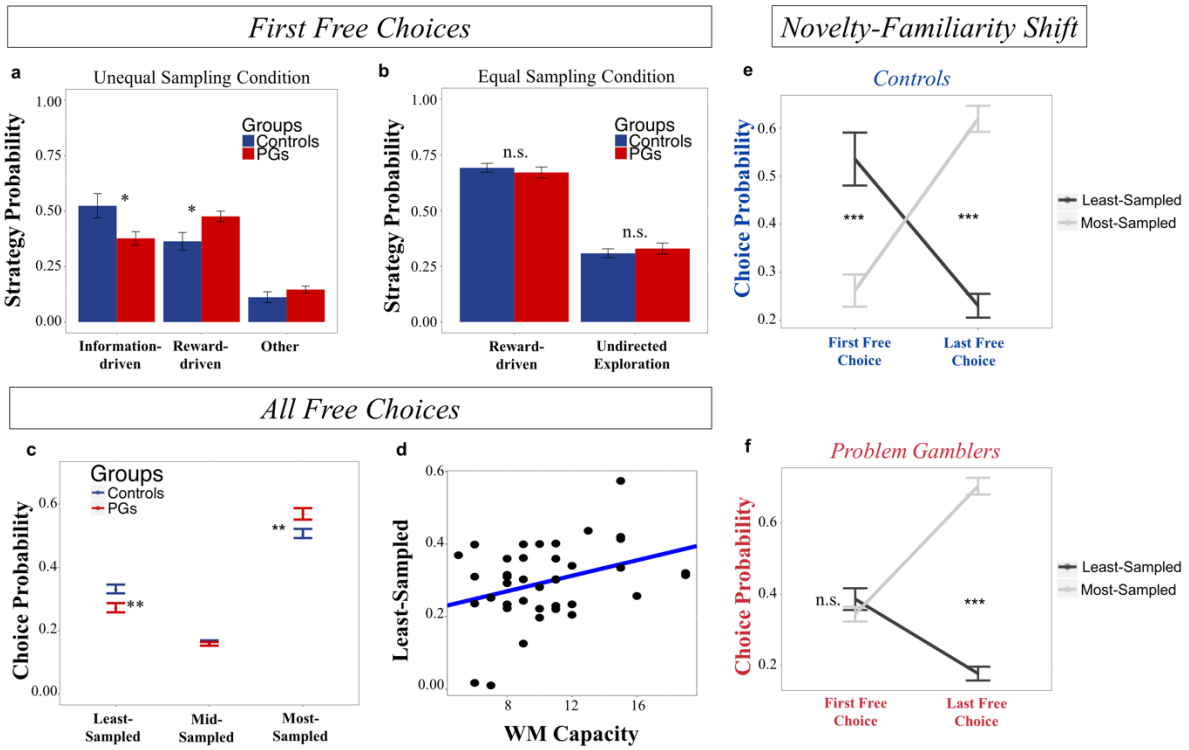


Figure 2. ***Model-Free analysis*****. a**) Probability to choose *information-driven*, *reward-driven* and *other* strategies in the first free-choice trial of the unequal information condition (i.e., when options are sampled unequally during the forced-choice task). Probabilities were averaged across games. Information-driven exploration decreases and reward-driven strategy increases in problem gamblers (PGs) compared to healthy controls. **b**) Probability of choosing *reward-driven* and *undirected exploration* in the first free-choice trial of the equal information condition (i.e., when options are sampled equally during the forced-choice task). No differences were observed between the two groups. **c**) Probability of selecting options the outcome of which was experienced the most (Most-Sampled), the least (Least-Sampled) and the mid (Mid-Sampled) number of times during previous trial history. Probabilities were averaged across games. Problem gamblers frequently selected Most-Sampled options at the expense of Least-Sampled alternatives compared to controls. Both groups showed higher preference for familiar options (Most-Sampled). **d**) In problem gamblers, the probability of selecting Least-Sampled options positive correlated with working memory (WM) capacity (computed using the working memory subset of the Wechsler Adult Intelligence Scale -WAIS IV; $r = 0.347$, $p = 0.028$, $n = 40$). As reported in Table 1, working memory capacity did not differ between problem gamblers and healthy controls as a group. **e**) Controls showed a novelty-familiarity shift by showing an increased preference toward novel options (Least-Sampled) in the first free-choice trial and an increased preference for familiar alternatives (Most-Sampled) in the last free-choice trial. d) Problem gamblers showed no preference between novel and familiar alternatives in the first free-choice trial, while in the last free-choice trial higher preferences toward familiar options were exhibited. In all the figures, error bars are represented as standard error of the mean (s.e.m).

Novelty-seeking in healthy individuals and novelty failure in problem gamblers

In order to elucidate the mechanisms underlying information valuation in healthy controls and problem gamblers, we turn to model-based analyses.

We first checked whether reward and information combination during learning could better explain participants' behavior compared to only computing reward predictions (Supplement). We replicated our previous findings [6] by showing that combining reward and information prediction better accounts for learning processes in both groups. Next, we compared the ability of our model – nkRL- to explain participants' behavior compared to a previous RL model of the same task (i.e., gkRL; Supplement). Despite the two models are equally good in explaining participants' choices across all trials (BIC_{gkRL}: M_{PGs} = 876.7, SD_{PGs} = 246; M_{CON} = 880.9, SD_{CON} = 157.9; BIC_{nkRL} : M_{PGs} = 894.6, SD_{PGs} = 231.4; M_{CON} = 879.9, SD_{CON} = 163; p_{PGs} = 0.86 and controls p_{CON} = 0.88; Figure 3a, b; Supplement), only nkRL was able to reproduce the behavioral pattern observed in problem gamblers' data in the first free choice trial (Figure 3c, d). We then adopted nkRL to better investigate the process underlying the absence of "novelty-familiarity" shift in gamblers. To do so, we compared the parameter estimates between the two groups. A Wilcoxon Signed Rank Test showed smaller novelty parameter ν in problem gamblers (M = - 1.6., SD = 15.9) compared to controls (M = 5.15, SD = 28.2), p = 0.012, while the knowledge parameter k did not differ between gamblers (M = - 0.304, SD = 2.34) and controls (M = 0.482, SD = 2.59), $p > 0.05$ (Figure 4 a, b). These results suggest that the absence of novelty-familiarity shift in gamblers' behavior is due to a failure in either computing or utilizing a novelty bonus early on in the free-choice period, while weights given to already acquired information were represented in the same way as in healthy controls. Problem gamblers also showed a smaller learning rate α (M = 0.349, SD = 0.253) compared to controls (M = 0.51, SD = 0.237; p = 0.022), suggesting a slower integration of new available reward outcomes from the environment (Figure 4c). Moreover, the analysis showed no differences in the parameter β between the two groups (M_{PG} = 0.484, SD_{PG} = 0.977; M_{CON} = 0.183, SD_{CON} = 0.188; $p > 0.2$) suggesting that the behavioral alterations observed in problem gamblers were not related to alterations in the decision noise (or policy). This latter result additionally confirms that exploratory impairments in problem gamblers were specifically driven by novelty-related information valuation without affecting other undirected or unexplained exploratory components (e.g., softmax parameter). Overall, the model-based analyses appear to suggest that healthy subjects integrate both general information and novelty during information valuation, while in gamblers the integration of novelty is specifically impaired alongside the coding scheme of reward learning experiences.

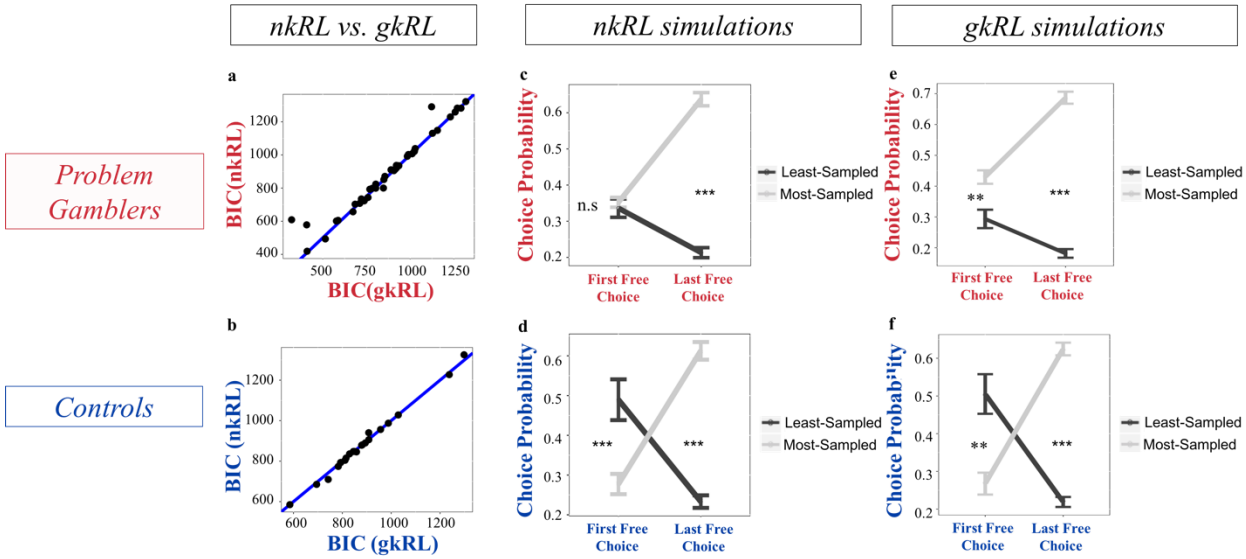


Figure 3. ***nkRL vs. gkRL.*** Comparitative fit of nkRL and gkRL model in problem gamblers (**a**) and controls (**b**). The comparative fit is based on BIC of both models computed by fitting the models to all participants' free choices. Each point is one participant. Almost all participants line on the identity line suggesting that overall the two models equally explain participants' behavior. Simulations of the nkRL model using the estimated individual parameters for problem gamblers (**c**) and controls (**d**). The model correctly predicts the novelty-familiarity shift in the healthy sample, and the absence of preference toward novel and familiar options for the gambling groups. Simulations of the gkRL model using the estimated individual parameters for problem gamblers (**e**) and controls (**f**). The model predicts higher preferences for familiar options in the first free-choice trials for the gambling group, while it correctly predicts the novelty-familiarity shift in the healthy sample. Error bars are represented as s.e.m.

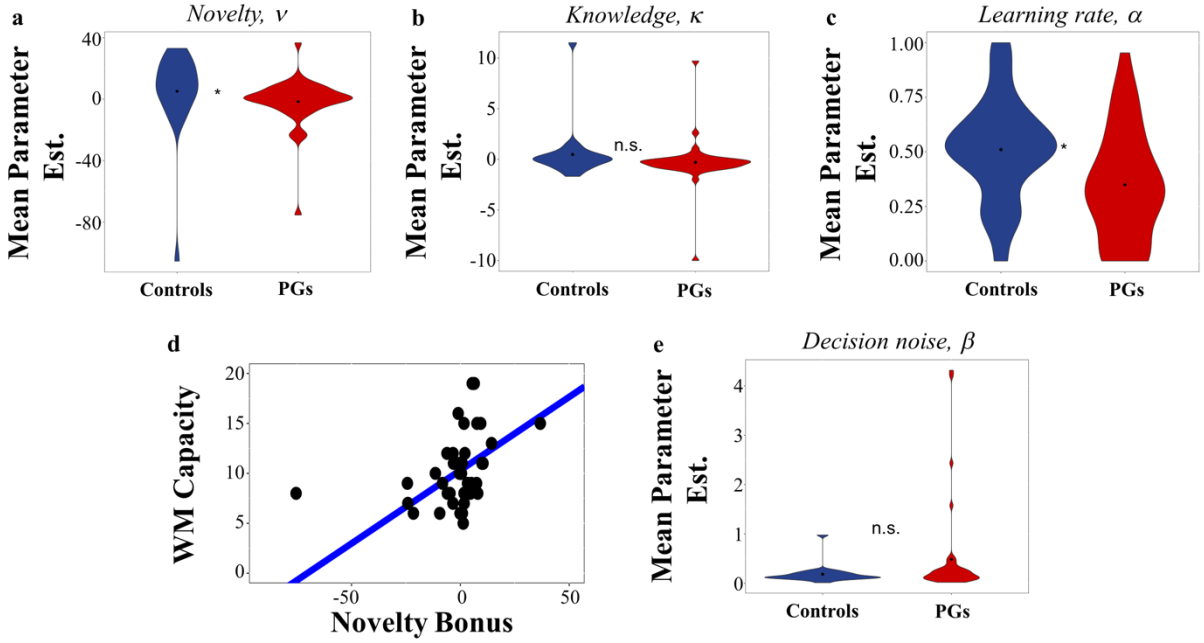


Figure 4. ***nkRL's estimated parameters.*** Model fit on all free-choices revealed a decrease in the novelty parameter ν (**a**) in problem gamblers (PGs) compared to controls, while the knowledge parameter κ did not differ between the two groups (**b**). Learning rate

290 α is reduced in problem gamblers compared to controls (**c**), while the decision noise did not differ between the two groups (**e**). **d**)
291 In problem gamblers, the novelty bonus increased as a function of working memory (WM) capacity ($r = 0.371$, $p = 0.018$, $n = 40$).
292

DISCUSSION

In this study, we adopted behavioral, self-reported, and computational measures to individuate the processes underlying human information behavior in healthy individuals and problem gamblers. We observed alternations in problem gamblers' information behavior compared to the healthy control group, as a consequence of a failure to represent novelty but not general information. By showing functional dissociations between novelty-induced exploration and general information, this study not only sheds light on the novelty-seeking impairment in addiction, but also highlights the importance of both novelty and general information-seeking in human information behavior, and their likely functional and biological dissociation in the human brain.

Information-seeking is a crucial component of adaptive behavior observed both in healthy humans [1] and animals [22]. Integrating information during learning appears to be a (inhibitory) control signal, that 'neurotypical' subjects use to adapt to the surrounding environment [7]. Defective information-seeking can indeed evolve in or contribute to certain psychopathologies [23] [24-26] [27]. By showing information-based exploratory impairments in problem gamblers, our findings suggest that the search for information in problem gamblers is compromised under certain conditions. Reduced exploratory behaviors in addiction has already been documented in past research [12]. By focusing on problem gambling, the results of this study clarify that exploratory impairments in addiction [12] are the results of modifications in decision-making processes related to addictive behaviors *per se*, and not by long-term intake of chemical compounds – although our study does not rule out the possibility that neurophysiological alterations in the brain could pre-date or even induce problem gambling. Furthermore, by using a task and a model that can specifically identify different components of exploratory drive, we showed that this impairment is specific to the information component of exploratory behaviors, and in particular to the elimination of the novelty bonus. The reduced ability to represent novelty may explain perseveration and impaired flexibility observed in both gambling [28,29] and substance abuse disorder [30,31]. In other words, our results suggest that the reduced ability to represent novel behavioral patterns may freeze addicted individuals' decision processes and trap them into the same behavioral routines. Further work, however, is necessary to confirm this suggestion and to test whether novelty-seeking may be target during clinical intervention to reduce the impact of perseveration in addicted individuals.

In contrast, we found that there was no difference between healthy controls and problem gamblers in terms of the softmax decision policy's temperature parameter fitted to the two subject groups. This temperature parameter has sometimes been interpreted as representing "undirected exploration" or decision "noise" in the literature [32] [5]. However, we note that any trial-to-trial variability in decision choice that is not accounted for explicitly by our model (in terms of reward or information value) would be absorbed into this decision stochasticity parameter. It is important to recognize that such apparent decision stochasticity could

either reflect true "noise" (or "undirected" exploration) in human choice behavior, or it could simply reflect the fact that our model insufficiently capture trial-by-trial variability. In any case, our results suggest that healthy controls and problem gamblers differ in the information-related component of sequential choice behavior, and not in other undirected or unexplained aspects of stochasticity in it. Therefore, only information-driven exploration could constitute a promising future target for the early diagnosis or clinical intervention of problem gambling [33]. Additionally, by showing information-based exploratory impairments in problem gamblers, the results of this study are in line with recent findings that assign different behavioral roles and neurocognitive mechanisms to informative and undirected component of exploration [5-7,34-36].

While our results showed novelty-seeking impairment in problem gamblers, different explanations may account for this impairment. One possibility is that problem gamblers may quickly jump to conclusion. After seeing the outcome of 2 out of 3 options, they might be highly confident in their representation of the environment and the search for novel information results "unnecessary". Abnormalities in confidence judgements and metacognitive capacities have been reported in both gambling [37] and substance abuse disorders [38]. However, it may also be possible that the absence of novelty bonus is due to an inability of dynamically represent the surrounding environment. Problem gamblers might be unable to represent changes in the environment, as when new options are available for selection. Model-based impairments have also been found to be associated with addictive disorders [39 40], and in particular with problem gambling [41]. Future experiments should explicitly test these alternative hypotheses.

Although our study adds additional insight on information behavior in behavioral addiction, some limitations may influence the scope of our results. Firstly, in order to have a control group as similar as possible to the gambling group (Table 1), the number of control participants we were able to include in the study after pre-screening was 22 (Supplement). The behavioral pattern observed in the control group (Figure 2a, S3a), however, replicates our previous findings on healthy humans playing with the behavioral task adopted in the current study [6,7]. Furthermore, although testing problem gamblers appears relevant for minimizing the confounding effects of chemical compounds, most of gambling games involve exploration/exploitation problems. Therefore, the observed behavioral alterations might have been affected by excessive gambling experience. However, we observed no differences between strategic and non-strategic gamblers (who usually play with games that employ different decision strategies, Supplement[42]). Moreover, our findings on alterations in exploratory behaviors are consistent with previous work on substance addiction where gambling experience was absent. Therefore, it is unlikely that our findings are an artifact resulting from more gambling experience. However, we cannot rule out the possibility that the observed differences might have already been present prior to the onset of gambling disorder instead of being directly induced by it.

Our results on novelty-seeking impairment in addiction not only provide insights on information-seeking alterations in addiction, but also deepens our understanding of information-seeking in general. We showed that healthy human behavior is motivated by reward gain and information gain, and the later decomposes into a novelty-specific component, and a general information component. By showing problem gamblers inappropriately represent novelty but not general information, our results appear to suggest that both components are not only functional but also biologically dissociable. Information-seeking behaviors are controlled by an interconnected cortico-basal ganglia network [43] and novelty-seeking is believed to motivate the brain's reward system [2 44]. However, the biological markers of both novelty and general information within the information-seeking network are still unknown. Further work is needed to individuate the mechanisms underlying novelty and general information and whether they are indeed biologically dissociable.

In summary, our findings extend the scientific understanding of human information behavior in healthy individuals and its impairment in behavioral addiction. Healthy information behavior was motivated by both novelty and general information. In contrast, problem gamblers showed impairment in novelty-seeking but not general information-seeking. Our results suggest that information-driven behaviors could be a promising target for clinical diagnosis and intervention (e.g., behavioral training and therapies, neurostimulation, etc.). However, whether our framework can be effectively implement in pre-existing clinical interventions, generalized to each individual patient [45], generalized to other types of addiction or other pathologies involving behavioral flexibility (see Supplement) has to be tested in future research. Methodologically, this work offers promising novel experimental and computational approaches for studying the mechanisms underlying information valuation in sequential choice behavior in both healthy and pathological populations.

ACKNOWLEDGEMENT AND DISCLOSURE

ICD is a researcher supported by F.R.S.-FNRS grant (Belgium). XN is a researcher supported by F.R.S.-FNRS (Belgium). AC is a research director with the F.R.S.-FNRS (Belgium). This work was in part funded by ERC Advanced Grant RADICAL to AC. AJY is funded in part by an NIH/NIDA CRCNS grant (R01 DA050373-01). We would like to thank Pauline Deroubaix for helping I.C.D in recruiting participants and collecting their data and Tali Sharot for comments on previous versions of this manuscript. The authors report no competing financial interests.

REFERENCES

1 Gottlieb, J., Oudeyer, P. Y., Lopes, M. & Baranes, A. Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends Cogn Sci* **17**, 585-593, doi:10.1016/j.tics.2013.09.001 (2013).

2 Kidd, C. & Hayden, B. Y. The Psychology and Neuroscience of Curiosity. *Neuron* **88**, 449-460, doi:10.1016/j.neuron.2015.09.010 (2015).

3 Sakaki, M., Yagi, A. & Murayama, K. Curiosity in old age: A possible key to achieving adaptive aging. *Neurosci Biobehav Rev* **88**, 106-116, doi:10.1016/j.neubiorev.2018.03.007 (2018).

4 APA. *Diagnostic and Statistical Manual of Mental Disoders*. (American Psychiatry Publishing, 2013).

5 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. *Journal of experimental psychology. General* **143**, 2074-2081, doi:10.1037/a0038199 (2014).

6 Cogliati Dezza, I., Yu, A. J., Cleeremans, A. & Alexander, W. Learning the value of information and reward over time when solving exploration-exploitation problems. *Sci Rep* **7**, 16919, doi:10.1038/s41598-017-17237-w (2017).

7 Cogliati Dezza, I., Cleeremans, A. & Alexander, W. Should we control? The interplay between cognitive control and information integration in the resolution of the exploration-exploitation dilemma. *Journal of experimental psychology. General*, doi:10.1037/xge0000546 (2019).

8 Stephens, D. W. & Krebs, J. R. *Foraging Theory*. (Princeton University Press, 1986).

9 Schulz, E. *et al.* Structured, uncertainty-driven exploration in real-world consumer choice. *Proceedings of the National Academy of Sciences of the United States of America* (2019).

10 Smirnov, Y., Koening, S., Veloso, M. M. & Simmons, R. G. Efficient Goal-Directed Exploration. *AAAI-96 Proceedings* (1996).

11 Chung, J. J., Lawrance, N. R. J. & Sukkarieh, S. Learning to soar: Resource-constrained exploration in reinforcement learning. *The international journal of robotics research* **34**, 158-172 (2015).

12 Morris, L. S. *et al.* Biases in the Explore-Exploit Tradeoff in Addictions: The Role of Avoidance of Uncertainty. *Neuropsychopharmacology* **41**, 940-948, doi:10.1038/npp.2015.208 (2016).

13 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, doi:10.1016/j.cognition.2013.08.018 (2013).

14 Bechara, A. Decision making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective. *Nat Neurosci* **8**, 1458-1463, doi:10.1038/nn1584 (2005).

15 Noel, X., Brevers, D. & Bechara, A. A neurocognitive approach to understanding the neurobiology of addiction. *Curr Opin Neurobiol* **23**, 632-638, doi:10.1016/j.conb.2013.01.018 (2013).

16 Hertwig, R. & Erev, I. The description-experience gap in risky choice. *Trends Cogn Sci* **13**, 517-523, doi:10.1016/j.tics.2009.09.004 (2009).

17 Leeman, R. F. & Potenza, M. N. Similarities and differences between pathological gambling and substance use disorders: a focus

435 on impulsivity and compulsivity. *Psychopharmacology* **219**, 469-490 (2012).
436 18 Lorains, F. K. *et al.* Strategic and non-strategic problem gamblers differ on decision-
437 making under risk and ambiguity. *Addiction* **109**, 1128-1137, doi:10.1111/add.12494
438 (2014).
439 19 Joy, S., Fein, D., Kaplan, E. & Freedman, M. Quantifying qualitative features of Block
440 Design performance among healthy older adults. *Arch Clin Neuropsychol* **16**, 157-170
441 (2001).
442 20 Rescorla, R. A. & Wagner, A. R. A theory of Pavlovian conditioning: Variations in the
443 effectiveness of reinforcement and nonreinforcement. *Classical conditioning: Current*
444 *research and theory*, 64-99 (1972).
445 21 Bishop, C. M. *Pattern recognition and machine learning*. (2006).
446 22 Bromberg-Martin, E. S. & Hikosaka, O. Midbrain dopamine neurons signal preference
447 for advance information about upcoming rewards. *Neuron* **63**, 119-126,
448 doi:10.1016/j.neuron.2009.06.009 (2009).
449 23 Camp, C. J. I am curious - grey: Information seeking and depression across the adult
450 lifespan. *Educational Gerontology* **12**, 375-384 (1986).
451 24 Gray, K. M. & Tonge, B. J. Are there early features of autism in infants and preschool
452 children? *J Paediatr Child Health* **37**, 221-226 (2001).
453 25 Aderka, I. M., Haker, A., Marom, S., Hermesh, H. & Gilboa-Schechtman, E. Information-
454 seeking bias in social anxiety disorder. *J Abnorm Psychol* **122**, 7-12,
455 doi:10.1037/a0029555 (2013).
456 26 Van den Driessche, C., Chevrier, F., Cleeremans, A. & Sackur, J. Lower Attentional Skills
457 predict increased exploratory foraging patterns. *Sci Rep* **9**, 10948,
458 doi:10.1038/s41598-019-46761-0 (2019).
459 27 Hauser, T. U., Moutoussis, M., Consortium, N., Dayan, P. & Dolan, R. J. Increased
460 decision thresholds trigger extended information gathering across the compulsivity
461 spectrum. *Transl Psychiatry* **7**, 1296, doi:10.1038/s41398-017-0040-3 (2017).
462 28 Goudriaan, A. E., Oosterlaan, J., de Beurs, E., van den Brink, W. Decision making in
463 gambling disoder: a comparison between pathological gamblers, alcohol dependents,
464 persons with Tourette syndrome, and normal controls. *Cogn. Brain. Res.*, 137-151,
465 doi:doi:10.1016/j.cogbrainres.
466 2005.01.017 (2005).
467 29 Marazziti, D. *et al.* Executive function abnormalities in pathological gamblers. *Clin*
468 *Pract Epidemiol Ment Health* **4**, 7, doi:10.1186/1745-0179-4-7 (2008).
469 30 Ersche, K. D., Roiser, J. P., Robbins, T. W. & Sahakian, B. J. Chronic cocaine but not
470 chronic amphetamine use is associated with perseverative responding in humans.
471 *Psychopharmacology (Berl)* **197**, 421-431, doi:10.1007/s00213-007-1051-1 (2008).
472 31 Camchong, J. *et al.* Frontal hyperconnectivity related to discounting and reversal
473 learning in cocaine subjects. *Biol Psychiatry* **69**, 1117-1123,
474 doi:10.1016/j.biopsych.2011.01.008 (2011).
475 32 Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An introduction*. (MIT Press, 1998).
476 33 Addicott, M. A., Pearson, J. M., Sweitzer, M. M., Barack, D. L. & Platt, M. L. A Primer on
477 Foraging and the Explore/Exploit Trade-Off for Psychiatry Research.
478 *Neuropsychopharmacology* **42**, 1931-1939, doi:10.1038/npp.2017.108 (2017).

34 Somerville, L. H. *et al.* Charting the expansion of strategic exploratory behavior during adolescence. *Journal of experimental psychology. General* **146**, 155-164, doi:10.1037/xge0000250 (2017).
35 Krueger, P. M. Strategies for exploration in the domain of losses. *Judgement and Decision Making* **12**, 104-117 (2017).
36 Zajkowski, W. K., Kossut, M. & Wilson, R. C. A causal role for right frontopolar cortex in directed, but not random, exploration. *Elife* **6**, doi:10.7554/eLife.27430 (2017).
37 Brevers, D. *et al.* Impaired Metacognitive Capacities in Individuals with Problem Gambling. *J Gambl Stud* (2012).
38 Hoven, M. *et al.* Abnormalities of confidence in psychiatry: an overview and future perspectives. *Transl Psychiatry* **9**, 268, doi:10.1038/s41398-019-0602-7 (2019).
39 Simon, D. A. & Daw, N. D. in *Computational Neuroscience of Drug Addiction* Vol. 10 (eds B. Gutkin & S.H. Ahmed) (Springer Series in Computational Neuroscien, 2012).
40 Sebold, M. *et al.* Model-based and model-free decisions in alcohol dependence. *Neuropsychobiology* **70**, 122-131, doi:10.1159/000362840 (2014).
41 Wyckmans, F. *et al.* Reduced model-based decision-making in gambling disorder . *Sci Rep* (2019).
42 Grant, J. E., Odlaug, B. L., Chamberlain, S. R. & Schreiber, L. R. Neurocognitive dysfunction in strategic and non-strategic gamblers. *Prog Neuropsychopharmacol Biol Psychiatry* **38**, 336-340, doi:10.1016/j.pnpbp.2012.05.006 (2012).
43 White, J. K. *et al.* A neural network for information seeking. *Nat Commun* **10**, 5168, doi:10.1038/s41467-019-13135-z (2019).
44 Kakade, S. & Dayan, P. Dopamine: generalization and bonuses. *Neural Netw* **15**, 549-559 (2002).
45 Fisher, A. J., Medaglia, J. D. & Jeronimus, B. F. Lack of group-to-individual generalizability is a threat to human subjects research. *Proc Natl Acad Sci U S A* **115**, E6106-E6115, doi:10.1073/pnas.1711978115 (2018).

SUPPLEMENTARY MATERIAL

Novelty-seeking impairment in addiction

Irene Cogliati Dezza[1,2,*], Xavier Noel[3], Axel Cleeremans[1], Angela J. Yu[4]

[1] Centre for Research in Cognition & Neurosciences, ULB Neuroscience Institute, Université Libre de Bruxelles, Belgium

[2] Department of Experimental Psychology, Faculty of Brain Sciences, University College London, London, UK

[3] Faculty of Medicine, Université Libre de Bruxelles, Belgium

[4] Department of Cognitive Science, University of California San Diego, United States

SUPPLEMENTARY METHODS

Clinical and demographic characteristics

Inclusion/exclusion criteria were examined the day before the experiment by conducting a short telephone interview as well as on the day of the experiment by filling self-reported questionnaires presented in a random order during the last part of the experimental session. The telephone interview was adopted as pre-screening for both gamblers and controls. We specifically asked for information concerning age, gender, frequency of gambling per week (problem gamblers) or last gambling experience (for controls), consumption of alcohol per week or substance (including legal and illegal drugs), inability to stop drinking alcohol, undergoing psychological treatments, and possible brain surgeries underwent in the past. In the following two sections, we describe the clinical and demographic characteristics of problem gamblers and healthy controls.

Problem gamblers

Gambling severity was evaluated using the Canadian Problem Gambling Index (CPGI [1]). Eight gamblers were classified as low level of problem gambling with $1 \leq$ GPCI ≤ 3, thirteen gamblers with moderate level of problem gambling (leading to some negative consequences; $4 \leq$ GPCI ≤ 7), and nineteen as exhibiting pathological problem gambling (with negative consequences and possible loss of control; GPCI$\geq$8). We also interviewed participants using DSM-V (French translation) and we observed that 52.4% of problem gamblers met the DSM-V criteria for gambling disorder [2]. The relatively low level of gambling addiction presented in this population is the result of selecting participants who showed no co-morbidities with substance abuse or alcohol use disorder. Specifically, to be able to tell apart effects of addictive behaviors *per se* on decision-making from effects of long-term intake of chemical compound, we tested problem gamblers with no use (N= 31, Drug Abuse Screening Test [3]- DAST =0) or non-problematic use (N=9, DAST =1) of legal and illegal substances and with absence of alcohol addiction (Alcohol Use Disorders Identification Test [4]- AUDIT- <12 in men and AUDIT < 11 in women, M = 4.625, SD = 3.868; N=30 did not show any misuse of alcohol AUDIT< 8). We also controlled for smoking addiction using the Fagerström Test for Nicotine Dependence- FTND [5]. Seven participants reported to smoke, but only 2 were classified with a mid-dependence and 2 with a weak-dependence, the other 3 were not dependent. Given that the main statistical results remained unchanged after removing those participants, we decided to include them in all the analyses. Additionally, to avoid the scenario that participants under psychological treatment may have developed a certain type of cognitive strategy over their decision processes, we included only participants who were not undergoing or seeking for psychological treatment. Moreover, we only included regular gamblers that were gambling at least once per week. Finally, we recruited both strategic problem gamblers (sport betting, poker, black jack; N=22) and non-strategic problem gamblers (bingo, lotto, slot machine, roulette; N=18) [6]. Given that no behavioral difference was found between the two sub-types (in line with

[7]), we combined strategic and non-strategic gamblers in the same gambling group in all analyses reported in this manuscript.

Healthy controls

The inclusion criteria for the healthy control group were as follow: CPGI=0 and no gambling experience in the past 12 months. 40% of control participants reported to have gambled in the past years, whereas the rest of the group reported to have never gambled in their life. As for the problem gambling group, we only included participants who scored DAST < 2 (with 17 subjects DAST = 0) and AUDIT < 12 (for the men), 11 (for the women) (with 17 subjects scored AUDIT< 8; M = 5.3, SD = 3.1). Three participants reported to smoke, two of them showed no sign of addiction (FTND = 0 ; 2) and one showed mid-level of addiction (FTND = 7). Removing this participant did not change the main statistical results, therefore the participant was included in all the analyses.

Behavioral Task

In the decision-making task, the influence of reward and information on choices is orthogonalized in the first free-choice trial (since after the commencement of the first free-choice trial, subjects tend to choose the more rewarding options more often, thus reward and information become anti-correlated). In particular, adding a forced-choice task before the actual decision task allows to control for available information and the reward magnitude associated with each option (i.e., options associated with the lowest amount of information were least associated with experienced reward values)[8]. This procedure allows to dissociate between information-driven exploration and undirected exploration. For instance, in the unequal sampling condition, the deck never selected during the forced choice task has highest informative value (it is completely unknown to participants) but it has no reward value associated with. By choosing that deck, participants are engaging in information-driven exploration. On the contrary, in the equal information condition, no differences are observed in terms of information. Therefore, whenever participants choose to explore, this strategy is not driven by an informative drive but only by decision noise[8].

As reported in the main text, the goal of the decision-making task is to maximize the final gain. The final gain is represented as the amount of points earned throughout the experiment. The total gain is showed to participants at the end of the experiment and converted in a monetary payoff (0.01 euros every 60 points). We adopted the same conversion procedure for both groups. However, because gamblers play regularly with higher amounts of money that those offered in our study, their contribution in the study was payed 2.5 more than in healthy controls. This modification was introduced in order to minimize the differences in motivation between the two groups during the experiment.

Furthermore, the reward was generated from a truncated Gaussian distribution with a fixed standard deviation of 8 points, and then rounded to the nearest integer. The generative mean for each deck was set to 30 (Low Reward Context) and 50 (High Reward Context) points and adjusted by +/- 0, 4, 12, & 20 points

(i.e., the generative means ranged from 10 to 70 points), to avoid the possibility that participants might be able to discern the generative mean for a deck after a single observation. The 3 decks of cards had the same generative means in 50% of the games (*equal reward*) and different means in the rest of the games (*unequal reward*). In the unequal reward condition (50% of the total games), the generative means differed so that two options had higher means compared to the third one in 25% of the total games (High Reward Context) and had lower means in the remaining games of the unequal reward condition (Low Reward Context). The appearance of the equal and unequal reward conditions was randomized. Participants were told that during the forced-choice task of certain games they could sample options at different rates, and that the decks of cards did not change during the same game, but were replaced by new decks at the beginning of each new game. However, they were not informed of the details of the reward manipulation or the underlying generative distribution adopted during the experiment. Contrary to our previous versions of this task [9 10], in half of the games of the equal reward-equal information condition, we introduced an unusually high reward outcome (with respect of the deck mean in that game) for a specific option (e.g., 90 points) the first time that this option was selected in the forced-choice task (subsequently the mean of the deck was set to its original value). This manipulation was introduced as a control condition in order to test whether gamblers' perseverate in choosing a generally poor option that they initially have a good experience with (the 'big win' hypothesis for gambling addiction [11]).

Computational Modelling

In this section, we first describe the RL model already validated for our task (gamma knowledge RL model-gkRL [9]). Next, we describe our novel implementation (nkRL). Lastly, we report the information concerning model fitting, selection and parameter recovery.

gkRL

The gkRL model learns reward values using the delta learning rule [12]:

$$Q_{t+1,j}(c) = Q_{t,j}(c) + \alpha \times \delta_{t,j}$$

$$\text{where, } \delta_{t,j} = R_{t,j}(c) - Q_{t,j}(c) \quad \text{(S1)}$$

where $Q_{t,j}(c)$ is the expected reward value for trial t and game j and $\delta_{t,j}$ is the *prediction error*, which quantifies the discrepancy between the previous predicted outcome $Q_{t,j}(c)$ and the actual outcome $R_{t,j}$ obtained at trial t and game j. Since participants were told that games were independent from one another, Q_0 is initialized at the beginning of each game to the global estimate of the expected reward values for each deck. We previously showed that this initialization was better able to capture healthy participants' behaviour than learning Q_0 on a trial-by-trial basis [9]. In addition, gkRL accumulates information over time

(i.e., the amount of observations toward each deck) where the importance of already acquired information is tuned by the information magnitude γ :

$$I_{t,j}(c) = \left(\sum_{1}^{t} i_{t,j}(c)\right)^{\gamma}$$

$$\text{where, } i_{t,j}(c) = \begin{cases} 0, & choice \neq c \\ 1, & choice = c \end{cases} \quad \text{(S2)}$$

γ defines both the degree of non-linearity in the amount of observations obtained from options after each observation and its related importance. Under high γ the information already gained is highly relevant, whereas the information to be acquired is less relevant or penalized. γ is constrained to be > 0. Next, in the gkRL the predicted value of a deck is a combination of reward and information where the importance of information relative to experienced reward value is controlled by the information integration ω:

$$V_{t,j}(c) = Q_{t+1,j}(c) - I_{t,j}(c) * \omega \quad \text{(S3)}$$

ω modulates the devaluation of previous experiences. With large ω the devaluation of experienced rewards increases favoring the selection of unknown options. Note that the effect of $I_{t,j}(c)$ is subtractive, because as more knowledge is gained about an option, less knowledge remains to be gained, thus the choice value decreases as $I_{t,j}(c)$ increases. Therefore, gkRL allows the exploration of unknown options by dynamically devaluating the selection of previous alternatives. Finally, a choice is made by entering choice values into the softmax function [13], as follows:

$$P(c/V_{t,j}(c_i)) = \frac{\exp(\beta \times V_{t,j}(c))}{\sum_i \exp(\exp \beta \times V_{t,j}(c_i))} \quad \text{(S4)}$$

where β is the inverse temperature that determines the degree to which choices are randomized by decision stochasticity (or choice variability).

nkRL

If, on one hand gkRL allows to solve exploration and exploitation problems on our decision-making task [9], on the other hand it cannot individuate the relative influence of general information and novelty on choices. Indeed, in gkRL information-driven exploration is allowed by dynamically devaluating previous options in order to explore new alternatives and by tuning the information magnitude; however, this formulation cannot differentiate the relative influence of general information and novelty on choices, since they are completely anti-correlated in gkRL. Therefore, we dissociate general information from novelty by utilizing a novel RL model i.e., nkRL.

As described in the main text nkRL computes expected reward values as in Eq. S1. Our novel model, however computes choice values as follows:

$$V_{t,j}(c) = Q_{t+1,j}(c) + \sum_1^t i_{t,j}(c) * k + 1_{unseen} * \nu \quad \text{(S5)}$$

Model fitting and Model selection

The models' parameters were estimated by fitting the model to trial-by-trial participants' free choices (~600 choices for each subject). The fitting procedure was performed using MATLAB function *fminsearchbnd* and iterated for 15 randomly chosen multiple starting points in order to minimize the chance of finding a local optimum instead of a global one. The fitting procedure was validated by running a recovery analysis: the nkRL (gkRL) model was simulated on the task using the retrieved parameter estimates to generate synthetic behavioral data and then the fitting procedure was applied to the synthetic data in order to check whether previously estimated parameters were indeed recovered. For model comparisons, negative log likelihoods obtained during the fitting procedure were used to compute model evidence (the probability of obtaining the observed data given a particular model). We adopted an approximation to the (log) model evidence, namely the Bayesian Information Criterion (BIC) [14]. We checked the model comparison outcome by computing a confusion matrix and checking whether data generated from a model was indeed best explained by that model. In order to inspect the fitting procedure for overfitting we adopted cross validation procedure [15]. We fitted nkRL and gkRL to 70% of the trials and we tested their ability to predict choices on future data (30% of the trials) compared to a simpler nested model (i.e., standard RL model –sRL- which enters directly enters reward values $Q_{t+1,j}$ into Eq. S4 without integrating information). We then adopted the likelihood ratio test to determine if the better fit of complex models (i.e., nkRL and gkRL) was due to noise captured in the data. Both nkRL and gkRL were correctly able to predict future choices.

Statistical analysis

Statistical analysis was performed using RStudio (https://www.rstudio.com/). When violations of parametric tests were indicated, non-parametric tests were performed. *P*-values < .05 were considered significant.

SUPPLEMENTARY RESULTS

Information behavior in healthy individuals and problems gamblers

Reward Context

We investigated whether differences in reward-driven and information-driven strategies between the two groups could be explained by reward contexts. To investigate the effect of reward context, we computed *reward-driven*, *information-driven* and *other* in the first free choice trial of the unequal information condition under High Reward Context (i.e., the generative mean of each deck was set to 30 points and adjusted by - 0, 4, 12, & 20 points) and Low Reward context (i.e., the generative mean of each deck was set to 30 points and adjusted by + 0, 4, 12, & 20 points). We conducted a 2 (Group: PGs, Controls) X 2 (Reward Context: Low Reward, High Reward) X 3 (Strategy: Reward-driven, Information-driven, Other) between-subject ANOVA. Besides the effects of strategy ($p < 10^{-15}$) and strategy X group ($p < 10^{-3}$) already reported in the main text, the results showed an interaction of strategy X Reward Context $F(2, 360) = 33.1$, $p < 10^{-13}$, whereas we did not find a general effect of Reward Context, Group X Reward Context, Group X Reward Context X Strategy (all $p > 0.958$). Post-hoc comparisons in the High Reward Context revealed an increase in *reward-driven* in problem gamblers (M = 0.554, SD = 0.165) compared to controls (M = 0.437, SD = 0.21), $p = .04$; a decrease in *information-driven* in problem gamblers (M = 0.276, SD = 0.194) compared to controls (M = 0.419, SD = 0.271), $p = .043$; whereas *other* did not differ between groups $p > 0.2$. Additionally, post-hoc comparisons in the Low Reward Context revealed an increase in *reward-driven* in problem gamblers (M = 0.339, SD = 0.164) compared to controls (M = 0.29, SD = 0.181), $p = .007$; a decrease in *information-driven* in problem gamblers (M = 0.479, SD = 0.217) compared to controls (M = 0.63, SD = 0.261), $p = .0075$; and an increase in *other* in problem gamblers (M = 0.123, SD = 0.112) compared to controls (M = 0.08, SD = 0.117), $p = .047$ (Figure S1). Moreover, both groups showed a decrease in *information-driven* and an increase in both *reward-driven* and *other* in High Reward Context compared to Low Reward Context (all p values < .05). These results replicate our previous findings on reward context and the resolution of the exploration-exploitation trade-off [9,10].

Reward Condition

In this section, we investigate whether differences in reward-driven and information-driven strategies between the two groups could be explained by distinct reward conditions (whether the Gaussian means of the decks were equal or unequal). A 2 (Group: PGs, Controls) X 2 (Reward Condition: Equal Reward, Unequal Reward) X 3 (Strategy: Reward-driven, Information-driven, Other) between-subject ANOVA. Besides the effects of strategy ($p < 10^{-15}$) and strategy X group ($p < 10^{-6}$) already reported in the main text, the results showed an interaction of strategy X Reward Condition $F(2, 360) = 46.83$, $p < 10^{-15}$, whereas we did not find a general effect of Reward Condition, Group X Reward Condition, Group X Reward Condition

X Strategy (all $p > 0.807$) suggesting that the observed decision-making impairments were independent on the reward conditions subjects were in.

Equal sampling scenario

In this section, we investigate whether the impairment in exploration also arises when exploration was not driven by an informative drive. To do so, we estimated the frequency of undirected exploration and reward-driven strategy in the first-free choice trials of the equal information condition i.e. when options have been seen the same number of times during the forced-choice task. Choices were classified as *reward-driven* when participants chose the deck with the highest average points and *undirected exploration* otherwise. A 2 (Group: PGs, Controls) X 2 (Strategy: Undirected Exploration, Reward-driven) between ANOVA revealed an effect of strategy $F(2,120) = 205.4$, $p < 10^{-15}$. However, the interaction effect Groups X Strategy and the Group effect were not significant ($p > .05$; Figure 2b), suggesting that, when information is deployed equally among options, problem gamblers engage in exploratory behaviors similarly to healthy controls.

Familiarity vs. Ignorance

In Figure 2c of the main text we showed that both groups preferred familiar options compared to options they were more ignorant about ($p < 10^{-4}$). In order to confirm this result, we computed participant's tendency to choose Most-Sampled and Least-Sampled alternatives with the respect of a baseline (i.e., mean tendency to choose Most-Sampled and Least-Sampled options across all subjects). A Wilcoxon Signed Rank Test showed higher tendency to choose Most-Sampled options over baseline (M_{PG} =1.35, SD_{PG} = 0.273; M_{CON} =1.21, SD_{CON} = 0.162) compared to Least-Sampled options (M_{PG} =0.648, SD_{PG} = 0.219; M_{CON} =0.791, SD_{CON} = 0.155) in both groups (all $p < 10^{-3}$; Figure S2 a). Additionally, we computed participant's tendency to choose Least-Sampled and Mid-Sampled alternatives with the respect of a baseline (i.e., mean tendency to choose Least-Sampled and Mid-Sampled options across all subjects). A Wilcoxon Signed Rank Test showed higher tendency to choose Least-Sampled options over baseline (M_{PG} =1.27, SD_{PG} = 0.48; M_{CON} =1.35, SD_{CON} = 0.263) compared to choose Mid-Sampled options (M_{PG} =0.735, SD_{PG} = 0.171; M_{CON} =0.653, SD_{CON} = 0.159) in both groups (all $p < 10^{-3}$; Figure S2 b). Therefore, both groups preferred most familiar options over most unknown over intermediate alternatives. We also computed the same tendency on first free choice trials. While both groups prefer novel options over intermediate alternatives (all $p < 10^{-3}$) (S2 d), healthy controls reversed their tendency selecting more often novel options (M =1.34, SD = 0.65) over familiar ones (M =0.656, SD = 0.393), $p < 10^{-3}$ (Figure S2 c). This novelty-familiarity shift was absent in problem gamblers who showed no preference between novel (M =1.06, SD = 0.528) and familiar options (M =0.943, SD = 0.358), $p = 0.41$ (Figure S2 c).

Novelty vs. General Information

In the main text and in the previous section we showed that both groups preferred most familiar options over novel options, and this tendency was reversed in the control group during the first free choice trials.

We further showed that both groups preferred novel options over mid-known alternatives in the first free choice trials (Figure S2 d). In order to understand whether the information-seeking impairment in problem gamblers is a result of a novelty-seeking impairment or of general information impairment, we directly compared the two groups preference over mid-known alternatives in the first free choice trial of the unequal information condition. If the information-seeking impairment is driven by general information, gamblers should also decrease the selection of mid-known alternative (which in the free choice trial of the unequal information condition corresponds to the option sampled twice during the forced choice task, and therefore less known to them). A Wilcoxon Signed Rank Test showed higher tendency to choose Mid-Sampled options in problem gamblers (M =0.27, SD = 0.084) compared to controls (M =0.203, SD = 0.112; p = 0.015) suggesting that information-seeking impairment in gamblers might be specific to the representation of novelty, rather than an impairment in general information. To better clarify this point, we turn to the model based analyses.

Novelty-seeking in healthy individuals and novelty failure in problem gamblers

Nested models

Even though the combination of reward and information was already shown in healthy human subjects[9 10], we checked if this was still the case in our experimental groups (problem gamblers and healthy controls). To do so, we compare gkRL to a simpler nested model i.e., sRL (which doesn't integrate information into the value function). A Wilcoxon Signed Rank Test showed a decrease in BIC_{gkRL} (M_{PGs} = 876.7, SD_{PGs} = 246; M_{CON} = 880.9, SD_{CON} = 157.9) compared to BIC_{sRL} (M_{PGs} = 925.3, SD_{PGs} = 216.5; M_{CON} = 941.9, SD_{CON} = 156.8) both in problem gamblers ($p < 10^{-7}$) and controls ($p < 10^{-5}$). The results suggest that the combination of reward and information better accounts for the learning process in both groups compared to only computing reward predictions, essentially replicating our previous findings [9].

gkRL's parameters

We compared the estimates of the gkRL's parameters obtained by fitting the model to participants' free-choices. A Wilcoxon Signed Rank Test showed a decrease in information integration parameter ω in problem gamblers (M = -14.2, SD = 95.4) compared to controls (M = 2.93, SD = 36.5), p = .003 (Figure S3a), and an increase in the information magnitude parameter γ in gamblers (M = 1.15, SD = 1.38) compared to controls (M = 0.145, SD = 0.352), p = .005 (Figure S3b). Additionally, the analysis showed a smaller learning rate α in problem gamblers (M = 0.391, SD = 0.264) compared to controls (M = 0.534, SD = 0.206), p = .038, suggesting a slower integration of new available reward outcome from the environment (Figure S3c). Lastly, the analysis showed no difference in the softmax parameter β between the two groups, p = .302 (M_G = 0.599, SD_G = 1.199; M_C = 0.139, SD_C = 0.073, (Figure S3d).

nkRL vs. gkrl

We compared the ability of nkRL to explain participants' behavior compared to a previous RL model of the same task (i.e., gkRL; Supplement). A Wilcoxon Signed Rank Test showed no difference between the BIC of either model in each group (BIC_{gkRL}: M_{PGs} = 876.7, SD_{PGs} = 246; M_{CON} = 880.9, SD_{CON} = 157.9; BIC_{nkRL} : M_{PGs} = 894.6, SD_{PGs} = 231.4; M_{CON} = 879.9, SD_{CON} = 163; p_{PGs} = 0.86 and controls p_{CON} = 0.88). Thus, gkRL and nkRL are equally good in explaining overall participants' choices (Figure 3a, b). However, in order to increase the probability of accurately estimating individuals' model parameters, both models were fitted to all free choices. Being the first free trials the trials were the differences between the two groups are mostly observed, we simulated both models using the parameters estimated. The rationale is that the differences between the two models in predicting participants' behavior may occur in the first free choice trial only. And, fitting the models to all free choices may average out the differences, if any, in the first free choice trials. In order to tell a part potential differences between gkRL and nkRL in the first free choice trial, we simulated both models using the parameters estimated during the fitting procedure and analyzed their predictions in both first and last free trials. In particular, we analyzed the models' tendency to choose Most-Sampled and Least-Sampled alternatives in the first and last free-choice trials. Both models predict a "novelty-familiarity" shift for the control groups ($p < 0.001$, as shown in participants' data Figure 2e; Figure 3d, f). However, gkRL predicts increased preferences toward frequently selected options for gamblers ($p < 0.01$; Figure 3e), while nkRL predicts no difference between familiar and novel options ($p > 0.05$, Figure 3c). Therefore, only our novel model was able to reproduce the behavioral pattern observed in problem gamblers' data during the first free choice trials.

Personality traits

In this section, we explore the correlations between model parameters and personal traits and the individual differences between problem gamblers and controls.

We observed significant correlations in problem gamblers between the novelty bonus ν and working memory capacity (r = 0.371, p = 0.018, p_{fdr} = 0.037, n =40; Figure 4d), sensation seeking (measured using sensation-seeking scale –SSS; [22]; r = 0.321, p = 0.043, p_{fdr} = 0.043, n =40) and sensitivity to reward (measured using punishment and reward sensitivity questionnaire- SPSRQ; [23]; r = 0.354, p = 0.025, p_{fdr} = 0.037, n =40). p_{fdr} is the p-value after correcting for multiple comparison. Additionally, the knowledge parameter κ positively correlated with sensation seeking (r = 0.344, p = 0.03, n =40). Also, we observed a negative correlation between softmax parameter β and positive mood (computed using PANAS- Table1; r = - 0.314, p = 0.049, n =40), suggesting that the more gamblers were feeling enthusiastic and active the more they were flexible in their decision policy.

We conclude our analyses by comparing individual differences between the two groups to investigate whether personal traits could explain the differences observed throughout our analyses. We focus on intolerance of uncertainty (EII [25]), impulsivity (UPPS-P [26]), sensation-seeking (SSS [23]), and sensitivity to

punishment and reward (SPSRQ [24]). Comparisons between healthy controls and problem gamblers revealed no differences in the scores obtained from EII (p = .785, BF_{01} = 3.61), UPPS-P (p = .217, BF_{01} = 1.89), SSS (p = .483, BF_{01} = 3.02), and SPSRQ (sensitivity to reward p = .399, BF_{01} = 2.81; sensitivity to punishment p = .266, BF_{01} = 2.4), suggesting that the behavioral alterations observed in problem gamblers are unlikely to be explained as differences in terms of personality traits (or in some cases there was not substantial evidence in favor of the alternative hypothesis). These results appear to suggest that information-seeking impairments in addictive disorders might be related to a process or mechanism that is independent from individual subjective preferences toward uncertainty, novelty, and reward.

The 'big win' hypothesis

The results reported in this study showed that problem gamblers' decision-making alterations when solving our sequential decision-making were driven by alterations in information behavior as a consequence of a failure to represent or incorporate novelty. However, these parametric alterations might have been confounded by the inability of problem gamblers of moving away from an option after experiencing fairly positive outcomes in the past, i.e., the 'big win' hypothesis. To better investigate this point, we computed the empirical probability of choosing an option associated with an unusually high score ("big win" options) when first selected in the forced-choice task. A two-sample t test showed no differences in the probability of choosing the "big win" option in problem gamblers (M = 0.607 SD = 0.187) compared to controls (M = 0.596 SD = 0.144), p = .798 suggesting that alterations in decision-making observed in problem gamblers were not driven by their particular persistence in choosing options associated with unusually good outcomes in the past.

Beyond gambling addiction

We investigate whether the decision-making task we adopted (and related computational model) might be able to explain other psychopathologies. We observed a positive correlation between Beck Depression Inventory (BDI) and undirected exploration (r = 0.3, p = 0.048, n =62) suggesting that the more participants were scoring high on the depression scale the more they were increasing undirected exploration as already observed in [16]. We additionally observed a negative correlation between the learning rate computed estimated nkRL model and BDI (r = - 0.3, p = 0.045, n =62) and a positive correlation between the softmax parameter β and BDI (r = 0.3, p = 0.02, n =62).

SUPPLEMENTARY REFERENCES

1 Ferris, J. & Wynne, H. The Canadian problem gambling index: Final report. (Canadian Centre on Substance Abuse., Ottawa (ON), 2001).

2 APA. *Diagnostic and Statistical Manual of Mental Disoders*. (American Psychiatry Publishing, 2013).

3 Skinner, H. A. The Drug Abuse Screening Test. *Addict Behav* **7**, 363-371 (1982).

4 Babor, T. F., de la Fuente, J. R., Saunders, J. & Grant, M. (World Health Organization).

5 Fagerström, K. O. Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment. *Addictive Behaviors*, 235–241 (1978).

6 Lorains, F. K. *et al.* Strategic and non-strategic problem gamblers differ on decision-making under risk and ambiguity. *Addiction* **109**, 1128-1137, doi:10.1111/add.12494 (2014).

7 Grant, J. E., Odlaug, B. L., Chamberlain, S. R. & Schreiber, L. R. Neurocognitive dysfunction in strategic and non-strategic gamblers. *Prog Neuropsychopharmacol Biol Psychiatry* **38**, 336-340, doi:10.1016/j.pnpbp.2012.05.006 (2012).

8 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. *Journal of experimental psychology. General* **143**, 2074-2081, doi:10.1037/a0038199 (2014).

9 Cogliati Dezza, I., Yu, A. J., Cleeremans, A. & Alexander, W. Learning the value of information and reward over time when solving exploration-exploitation problems. *Sci Rep* **7**, 16919, doi:10.1038/s41598-017-17237-w (2017).

10 Cogliati Dezza, I., Cleeremans, A. & Alexander, W. Should we control? The interplay between cognitive control and information integration in the resolution of the exploration-exploitation dilemma. *Journal of experimental psychology. General*, doi:10.1037/xge0000546 (2019).

11 Kassinove, J. I. & Schare, M. L. Effects of the "near miss" and the "big win" on persistence at slot machine gambling. *Psychol Addict Behav* **15**, 155-158 (2001).

12 Rescorla, R. A. & Wagner, A. R. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning: Current research and theory*, 64-99 (1972).

13 Bishop, C. M. *Pattern recognition and machine learning*. (2006).

14 Schwarz, G. Estimating the dimension of a model. *Ann. Stat* **6**, 461-464 (1978).

15 Daw, N. D. in *Affect, Learning and Decision Making, Attention and Performance XXIII* (eds E.A. Phelps, T.W. Robbins, & M. Delgado) (Oxford University Press, 2009).

16 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, doi:10.1016/j.cognition.2013.08.018 (2013).

873

874

SUPPLEMENTARY FIGURES

Figure Captions

Figure S1. ***Reward Context*****.** Probability to perform *reward-driven*, *information-driven* and *other* strategy under High (**left panel**) and Low (**right panel**) Reward Context in both problem gambling and healthy group. Each group reduced information-driven choices and increased reward-driven choices in Low Reward Context compared to High Reward Context. And, compared to controls, problem gamblers (PG) frequently engaged in reward-driven decisions at the expense of information-driven in both reward contexts. Error bars are represented as s.e.m.

Figure S2. ***Familiarity vs. Ignorance*****.** Participant's tendency to choose Most-Sampled and Least-Sampled alternatives (**a**,**c**), Least-Sampled and Mid-Sampled alternatives (**b**,**d**) with the respect of a baseline (i.e., mean tendency to choose Least-Sampled and Most-Sampled options or Least-Sampled and Mid-Sampled across all subjects).

Figure S3. ***gkRL's estimated parameters*****.** Model fit on all free-choices revealed a decrease in information integration ω (**a**), learning rate α (**b**) and an increase in the information magnitude γ (**c**) in problem gamblers compared to controls. The softmax parameter β did not differ between the two groups (**d**).

Figure

Figure S1

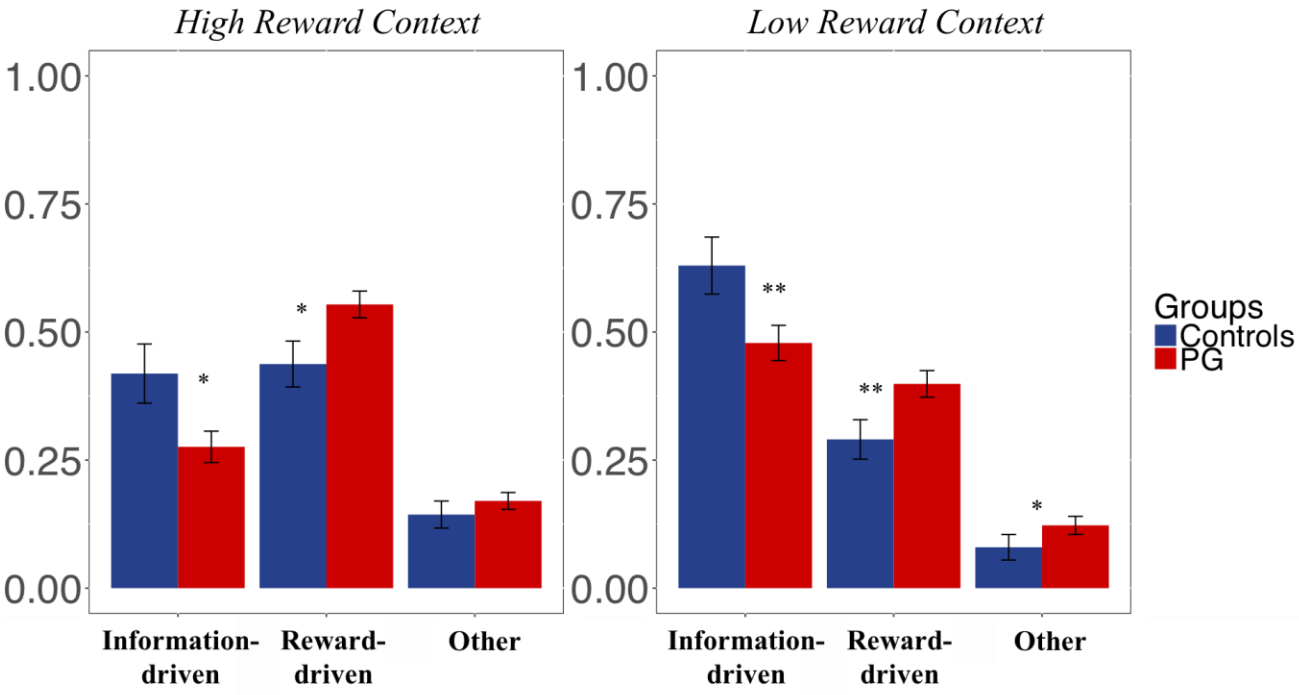


Figure S2

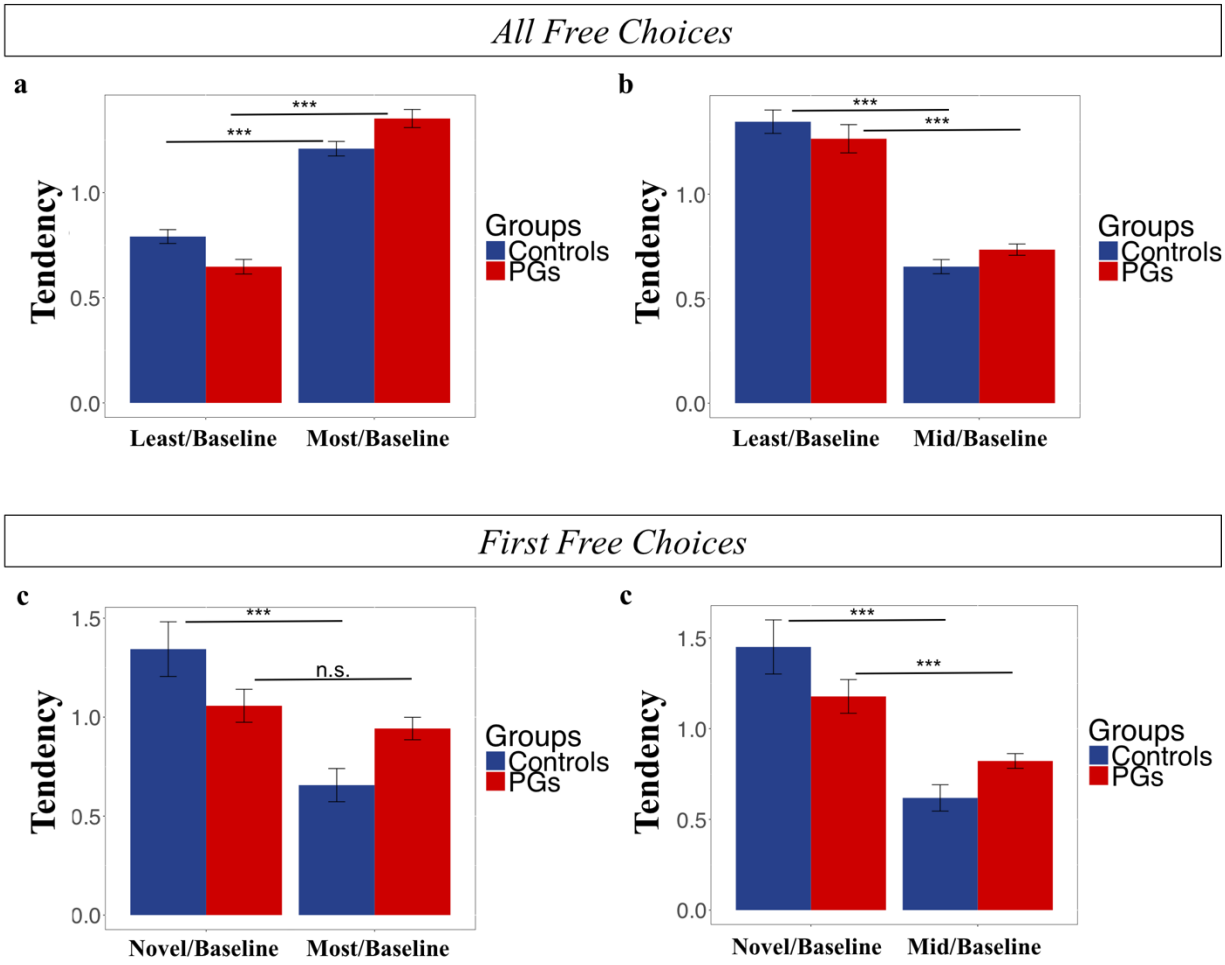
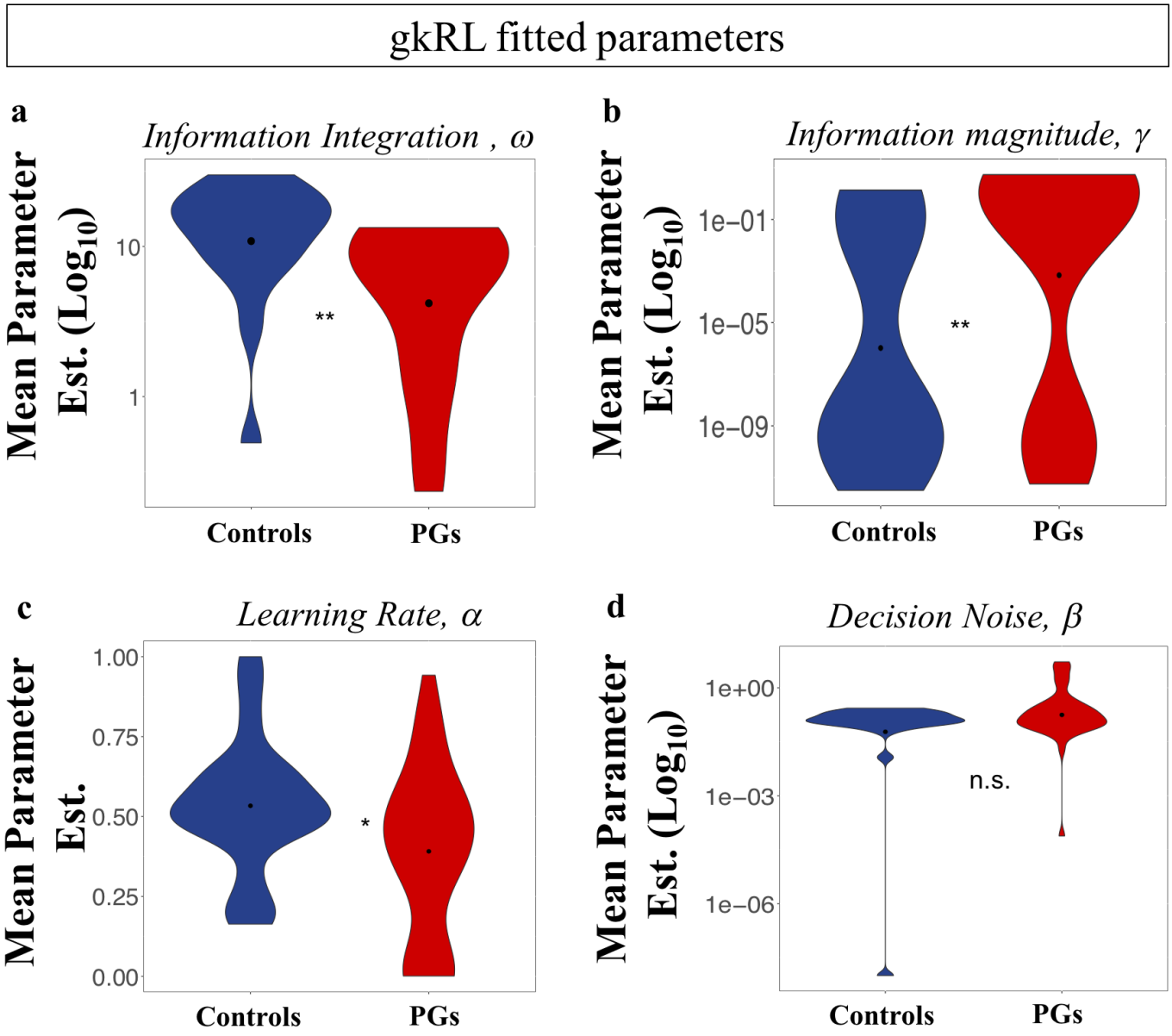


Figure S3



SUPPLEMENTARY TABLEs

Table Captions

Table S1 *gkRL individual fitted parameters in PGs*

Table S2 *gkRL individual fitted parameters in Controls*

Table S3 *nkRL individual fitted parameters in PGs*

Table S4 *nkRL individual fitted parameters in Controls*

Table

Table S1

PGs	α	β	ω	γ
Subject01	0.545	0.055	-0.6	1.528
Subject02	0.53	0.098	-0.005	4.187
Subject03	0.491	0.07	-7.753	0.546
Subject04	0.942	0.019	1.201	1.79
Subject05	0.396	0.095	8.124	0.202
Subject06	0.257	0.356	-13.703	0.387
Subject07	0.438	0.378	8.559	5.66E-10
Subject08	0.565	0.11	6.752	1.63E-11
Subject09	0.014	1.687	1.64	0.038
Subject10	0.285	0.17	10.901	1.39E-10
Subject11	0.471	0.22	-0.205	2.015
Subject12	0.601	0.095	-6.066	0.283
Subject13	0.529	0.075	-0.07	3.251
Subject14	0.362	0.186	11.213	7.52E-10
Subject15	0.007	5.283	0.234	4.73E-10
Subject16	0.007	2.14	-0.242	0.535
Subject17	0.588	0.098	9.784	2.18E-09
Subject18	0.007	1.156	0.404	0.132
Subject19	0.017	0.103	-0.212	1.971
Subject20	0.555	0.197	8.278	7.24E-10
Subject21	0.077	0.307	4.265	2.18E-11
Subject22	0.363	0.09	-0.024	3.52
Subject23	0.018	0.531	1.605	8.47E-11
Subject24	0.741	0.062	-23.731	5.62E-12
Subject25	0.405	0.166	-0.085	2.767
Subject26	0.512	0.202	13.36	1.52E-10
Subject27	0.001	0.0001	-601.036	2.247
Subject28	0.398	0.252	-12.281	0.63
Subject29	0.49	0.13	-1.424	1.087
Subject30	0.751	0.104	10.206	1.57E-10
Subject31	0.33	0.175	-0.288	1.916
Subject32	0.007	3.948	-0.038	1.231
Subject33	0.451	0.243	0.0001	5.459
Subject34	0.001	4.096	0.0001	3.602
Subject35	0.359	0.426	-0.209	1.571
Subject36	0.472	0.126	-0.789	1.561
Subject37	0.818	0.045	-7.032	0.215
Subject38	0.691	0.114	-0.152	1.719
Subject39	0.813	0.087	-0.921	1.482
Subject40	0.336	0.264	11.056	4.25E-10

Table S2

Controls	α	β	ω	γ
Subject01	0.791	0.135	21.873	2.32E-10
Subject02	0.485	0.174	8.479	0.053
Subject03	0.517	0.192	29.985	3.75E-10
Subject04	0.527	0.12	13.341	0.164
Subject05	0.644	0.103	-7.579	4.30E-09
Subject06	0.546	0.096	-2.605	1.013
Subject07	0.211	0.211	14.686	0.009
Subject08	0.609	0.11	-10.394	0.279
Subject09	0.204	0.235	2.889	1.49E-09
Subject10	0.163	0.272	8.177	4.21E-12
Subject11	0.466	0.139	11.404	3.17E-12
Subject12	0.468	0.094	15.366	9.65E-10
Subject13	0.58	0.012	-152.108	3.07E-10
Subject14	0.657	0.122	19.181	0.086
Subject15	0.561	0.268	5.456	1.51E-09
Subject16	0.653	0.129	22.011	0.182
Subject17	0.908	0.062	-9.774	7.10E-10
Subject18	1	0.083	19.082	7.24E-10
Subject19	0.473	0.123	29.375	0.039
Subject20	0.317	0.202	8.53	1.80E-10
Subject21	0.486	0.185	16.532	3.87E-10
Subject22	0.474	1.00E-08	0.491	1.375

Table S3

PGs	α	β	κ	ν
Subject01	0.471	0.06	0.464	-3.412
Subject02	0.327	0.137	-0.216	1.768
Subject03	0.32	0.093	-0.023	-11.466
Subject04	0.953	0.019	9.721	36.714
Subject05	0.411	0.092	1.09	14.321
Subject06	0.23	0.339	-0.525	-24.123
Subject07	0.429	0.397	-0.772	5.453
Subject08	0.567	0.109	-0.46	5.053
Subject09	0.015	1.576	-0.032	1.669
Subject10	0.341	0.153	-2.008	4.61
Subject11	0.374	0.259	-0.351	-3.353
Subject12	0.578	0.099	0.082	-8.211
Subject13	0.057	0.447	-0.251	0.93
Subject14	0.363	0.184	-0.511	9.34
Subject15	0.259	0.241	-0.967	2.08
Subject16	0.007	2.434	-0.006	-0.448
Subject17	0.587	0.098	0.166	10.411
Subject18	0.157	0.067	-0.145	7.452
Subject19	0.019	0.108	-1.325	-3.154
Subject20	0.545	0.198	-0.501	6.368
Subject21	0.081	0.29	-0.331	3.275
Subject22	0.219	0.138	-1.07	0.148
Subject23	0.018	0.532	-0.081	1.317
Subject24	0.74	0.062	-0.166	-24.325
Subject25	0.249	0.258	-0.911	1.684
Subject26	0.509	0.201	-0.892	9.942
Subject27	1.01E-08	0.019	-10	-75.45
Subject28	0.258	0.244	0.532	-21.555
Subject29	0.38	0.152	-0.752	-6.141
Subject30	0.76	0.103	-0.595	7.975
Subject31	0.215	0.253	-0.639	-4.756
Subject32	0.007	4.317	-0.027	-0.207
Subject33	0.39	0.27	-0.262	-0.992
Subject34	0.001	4.207	-0.026	-0.135
Subject35	0.313	0.491	-0.373	-0.416
Subject36	0.278	0.175	-0.84	-5.974
Subject37	0.831	0.047	2.615	0.848
Subject38	0.666	0.119	-0.183	0.508
Subject39	0.729	0.096	-0.74	-9.481
Subject40	0.323	0.267	-0.859	7.627

Table S4

Controls	α	β	κ	ν
Subject01	0.792	0.135	-0.092	21.556
Subject02	0.49	0.172	-0.107	8.647
Subject03	0.554	0.182	-0.967	27.239
Subject04	0.566	0.115	1.206	21.26
Subject05	0.652	0.103	0.534	-5.61
Subject06	0.395	0.117	0.105	-6.518
Subject07	0.214	0.209	0.068	15.278
Subject08	0.542	0.119	-0.204	-14.775
Subject09	0.209	0.227	-0.61	0.751
Subject10	0.172	0.26	-0.27	7.534
Subject11	0.485	0.133	-1.299	6.819
Subject12	0.473	0.094	0.461	17.079
Subject13	0.585	0.012	11.524	-105.482
Subject14	0.666	0.121	0.897	24.784
Subject15	0.561	0.267	0.055	5.67
Subject16	0.69	0.124	1.622	33.115
Subject17	0.911	0.062	-0.68	-12.356
Subject18	1	0.083	0.235	19.923
Subject19	0.473	0.123	0.513	32.645
Subject20	0.306	0.206	-0.759	5.464
Subject21	0.485	0.181	-1.683	10.729
Subject22	1.00E-08	0.978	0.044	-0.465