



AKADÉMIAI KIADÓ

Journal of Behavioral Addictions

11 (2022) 3, 831–844

DOI:
10.1556/2006.2022.00059
© 2022 The Author(s)

FULL-LENGTH REPORT



The modulation of acute stress on model-free and model-based reinforcement learning in gambling disorder

FLORENT WYCKMANS¹ , NILOSMITA BANERJEE¹ ,
MÉLANIE SAEREMANS² , ROSS OTTO³ ,
CHARLES KORNREICH^{1,2} , LAETITIA VANDERIJST¹ ,
DAMIEN GRUSON⁴ , VINCENZO CARBONE⁴ ,
ANTOINE BECHARA⁵ , TONY BUCHANAN⁶ and
XAVIER NOËL^{1*}

¹ Psychological Medicine Laboratory, Université Libre de Bruxelles, Brussels, Belgium

² Psychiatric Institute, Universitary Hospital Brugmann, Brussels, Belgium

³ Department of Psychology, McGill University, Montréal, Canada

⁴ Department of Laboratory Medicine, Cliniques Universitaires St-Luc, Brussels, Belgium

⁵ Faculty of Psychology, University of Southern California, Los Angeles, USA

⁶ Department of Psychology, Saint Louis University, USA

Received: April 25, 2022 • Revised manuscript received: July 19, 2022 • Accepted: July 25, 2022
Published online: September 16, 2022

ABSTRACT

Background and aims: Experiencing acute stress is common in behavioral addictions such as gambling disorder. Additionally, like most substance-induced addictions, aberrant decision-making wherein a reactive habit-induced response (conceptualized as a Model-free [MF] in reinforcement learning) suppresses a flexible goal-directed response (conceptualized as a Model-based [MB]) is also common in gambling disorder. In the current study we investigated the influence of acute stress on the balance between habitual response and the goal-directed system. **Methods:** A sample of $N = 116$ problem gamblers (PG) and healthy controls (HC) performed an acute stress task – the Socially Evaluated Cold pressure task (SECPT) – or a control task. Self-reported stress and salivary cortisol were collected as measures of acute stress. Following the SECPT, participants performed the Two-Step Markov Task to account for the relative contribution of MB and MF strategies. Additionally, verbal working memory and IQ measures were collected to account for their mediating effects on the orchestration between MB/MF and the impact of stress. **Results:** Both groups had comparable baseline and stress-induced cortisol response to the SECPT. Non-stressed PG displayed lower MB learning than HC. MANOVA and regression analyses showed a deleterious effect of stress-induced cortisol response on the orchestration between MB and MF learning in HC but not in PG. These effects remained when controlling for working memory and IQ. **Discussion and Conclusions:** We found an abnormal pattern of modulation of stress on the orchestration between MB and MF learning among PG. Several interpretations and future research directions are discussed.

KEYWORDS

gambling disorder, reinforcement learning, model-based, model-free, Two-Step Markov Task, stress

*Corresponding author.
E-mail: xavier.noel@ulb.be

INTRODUCTION

Gambling is a ubiquitous recreational activity worldwide. Despite its recreational facet, for a small yet significant proportion of gamblers (estimated around 0.1–5.8% worldwide),

tenacious engagement in gambling results in the development of Gambling Disorder (GD), a persistent, excessive and uncontrollable urge towards gambling despite the looming negative consequences of this excessive engagement (APA, 2013). GD can be conceived as reflecting (1) an increased activity within driving paths involved in seeking pleasure, reward or reduction of stress and negative moods, as well as overreliance on habit-induced reactive responses and (2) a decreased self-control (Brand, 2022; Goldstein & Volkow, 2002; Noël, Brevers, & Bechara, 2013). Based on this theoretical consideration, two key aspects that characterize GD are: 1) gambling to escape stress (Buchanan, McMullin, Baxley, & Weinstock, 2020) and 2) making decisions that are informed by habit-induced, reactive responses instead of using a deliberative, goal-directed flexible approach (Everitt & Robbins, 2005). While these characteristic features of GD have been studied in isolation, to our knowledge a systematic investigation of a possible interaction between the experience of acute stress and its consequent impact on habit versus flexible decision-making has not been undertaken in GD. The aim of this study is to investigate the existence of this possible interaction. More specifically, the current study aims to examine stress-induced modulation of the arbitration between habit-induced, highly reactive decisions versus deliberative, goal-directed flexible decisions.

Stress represents a major accelerating factor in the development of addictive disorders (Enoch, 2011), including GD (Biback & Zack, 2015; Buchanan et al., 2020; Oakes, Pols, & Lawn, 2019). Stress has a paradoxical impact in GD. On the one hand recreational elements of gambling activities and the rewarding experience of gambling engagements are leveraged by gamblers to cope with stressful life events (Edgerton, Keough, & Roberts, 2018; Weinstock, Whelan, & Meyers, 2008), on the other hand consistent gambling engagement in itself can be a distressing experience (APA, 2013) making gambling a potential stressor (Russell et al., 2021). From a neurocognitive perspective, recurrent stress exposure can cause profound changes within neural systems involved in decision-making (Wirz, Bogdanov, & Schwabe, 2018). Particularly, recurrent exposure to acute stress may increase one's tendency to behave in an inflexible habitual way (Schwabe & Wolf, 2009, 2010, 2011; Seehagen, Schneider, Rudolph, Ernst, & Zmyj, 2015). Additionally, uncontrollable stress causes a rapid and dramatic loss of prefrontal cognitive abilities (e.g., response inhibition, mental shifting, working memory) (Arnsten, 2009), thus diminishing self-control (impaired 'stop now process', (Brand, 2022)) and favoring the maladaptive persistence of behaviors (e.g. gambling) despite being aware of its negative consequences (i.e. compulsion).

Clinically, inflexible and habitual responses (as opposed to deliberative and goal-directed responses) inform the decisions undertaken in addictive states like GD (c.f., Wyckmans et al., 2019), fueling the persistence of addictive behaviors like gambling despite the possibility of recurring negative consequences (e.g., financial losses; Everitt & Robbins, 2005). According to reinforcement learning (RL) theory, such habitual, inflexible responses and the contrary

goal-directed behaviors arise from two parallel learning systems, Model-Based (MB) and Model-Free (MF), which differ in their way of updating choice-value during instrumental learning (Daw, Niv, & Dayan, 2005, 2011). According to RL theory, when facing a decision, the MB-system relies on an internal model of the contingencies between actions and consequences, which informs the computations of the expected values for each candidate decision (Daw, 2018). Thus, MB strategies underlie goal-directed behavior, as it enables flexible adaptation to the environment and prospective update of action value in view of long-term goals (Dolan & Dayan, 2013). Conversely, the MF-system bypasses these laborious simulations and directly computes the action-value based on past reward history (Rummery & Niranjan, 1994; Sutton & Barto, 1998). Therefore, the MF-system is associated with habitual behavior as it depends solely on previously learned associations (Dolan & Dayan, 2013). Despite the advantage gained by the lowered cognitive burden in MF decision-making strategies, it comes at the cost of hampered flexibility: if the contingencies between action and consequences change, the stored action values become obsolete and lead to outdated choices (Daw, 2018). Thus, under RL theory, addictive behaviors (which occur from excessive reliance on habitual, reflexive responses coupled with diminished reliance on deliberative goal-directed systems), arise from the asymmetrical dynamics between the MF-MB systems wherein there is an exacerbated MF-system involvement coupled with a suppressed MB-system involvement underlying the persistence of the addictive behavior at hand (e.g., Groman, Massi, Mathias, Lee, & Taylor, 2019). In fact, recent research undertaken in the field of substance-induced addictions (e.g., binge drinking; Doñamayor, Strelchuk, Baek, Banca, & Voon, 2018) highlights that the presence of such an imbalance between the MF and MB systems acts as a key factor underlying the perpetuation of the addictive behaviors. Crucially, recent evidence also suggests that a preclinical existence of such imbalance could predispose an individual to develop addictive behaviors (e.g., Chen et al., 2021). Additionally, evidence from recent research has also replicated these findings with PG (i.e., in the behavioral addiction of GD; Wyckmans et al., 2019). Wyckmans et al. (2019) demonstrated a diminished MB-system involvement coupled with an heightened MF-system involvement in GD in the Two-step Markov task (Daw, Gershman, Seymour, Dayan, & Dolan, 2011) that systematically accounts for the MF/MB strategies in decision-making.

In terms of the impact of stress on MF/MB-systems, studies indicate that acute stress has a deleterious effect on MB learning (Otto, Raio, Chiang, Phelps, & Daw, 2013; Radenbach et al., 2015) which occurs due to the hampered functioning of the amygdala-dorsal striatum connectivity and prefrontal cortex, resulting in diminished executive processes (Wirz et al., 2018). Conversely, stress-induced cortisol release enhances MF learning (Park, Lee, & Chey, 2017) by increasing the firing rate of dopaminergic neurons in the dorsal striatum (Anstrom & Woodward, 2005), a region involved in habit formation (Everitt & Robbins, 2005, 2016)

such as those observed in addictive states. Considering the above evidence, it seems plausible to reason that repeated exposure to acute stress during gambling might present a risk factor for the development of GD by favoring MF learning, thus rendering the resulting behavior resistant to its negative consequences. Therefore, acute stress experienced from gambling (Russell et al., 2021), the increased reliance on MF-systems among PG (Wyckmans et al., 2019), and the amplification of MF-systems as a function of increased acute stress (Park et al., 2017), calls for a need to systematically investigate the effects of acute stress on the MF-system among PG.

To pursue this line of investigation and understand how acute stress modulates the arbitration between MF/MB-systems, computational modelling approaches that augment the previously used experimental inferential analyses approaches are needed (e.g., Wyckmans et al., 2019). Traditionally, MF/MB-systems are studied using the Two-step Markov task (RL-task; Daw et al., 2011). Experimental approaches which focus on inferential analyses of the choice behavior data such as reaction times and state-transition probabilities from the Two-Step Markov Task do not provide insights about the latent learning processes in decision-making that arise from the implementation of MF or MB strategies. Discerning these latent psychological variables is crucial to account for the impact of acute stress on the relative weighing between MF/MB systems. Thus, fitting behavioral choice data from the Two-Step Markov Task to a Hybrid-RL 7 parameter model (Daw et al., 2011), wherein each parameter will allow us to systematically estimate the learning processes involved during decision-making (Gillan, Otto, Phelps, & Daw, 2015; Mollick & Kober, 2020), would appropriately discern the impact of acute stress on the dynamics between MF and MB.

The aim of the current study is to systematically investigate the impact of acute stress on the exacerbation of the MF-system among PG. We aim to investigate this by fitting a Hybrid-RL 7-parameter computational model to the behavioral data obtained from Two-step Markov task, a novel approach which remains unexplored in this line of research. Acute stress will be induced with the Socially Evaluated Cold Pressor Task (SECPT) that requires participants to immerse one arm into cold water while being socially observed and videorecorded, a method that robustly increases laboratory stress reactivity (Schwabe, Haddad, &

Schachinger, 2008), including among PG (Grant & Chamberlain, 2019). Based on previous findings we hypothesized that PG will repeat previously rewarded choices without accounting for the task structure (Wyckmans et al., 2019). Thus, we expected to observe a general bias towards MF learning among PG as compared to healthy controls. Furthermore, we hypothesized that stress would further increase the activity of the MF-system and suppress the MB-system among stressed PG as opposed to stressed controls. Finally, we also accounted for verbal working memory and abstract reasoning on MB/MF to account for their mediating effects on the acute stress and MF/MB interaction.

METHODS

Participants

Seventy-three problem gamblers (PG) with a score of 6 or higher on the South Oaks Gambling Screen (Lesieur & Blume, 1987) and 75 healthy controls (HC) were recruited. Only male participants (to avoid the confounding effect of menstrual cycle on cortisol response to stress; Montero-López et al., 2018) over 18 and without a history of neurological or psychiatric diseases were included. We excluded five participants (two PG and three HC) due to corrupted cortisol samples, as well as 27 participants (13 PG and 14 HC) who did not meet the success criteria for the RL-Task (see below), yielding a final sample of 116 participants (58 PG and 58 HC).

Procedure

The entire procedure lasted two hours (Fig. 1). Remuneration was set at 30€, with up to 10€ extra depending on individual task performance. Participants first underwent a semi-structured interview assessing demographic characteristics and gambling habits (if applicable). Fluid intelligence was assessed with the nine-item forms of the Raven's Standard Progressive Matrices Test (see supplementary material, Bilker et al., 2012), followed by a working memory task, the Operation Span Task (OSPAN; see supplementary material, Unsworth, Heitz, Schrock, & Engle, 2005) and the RL-task instructions. Before completing the RL-task, half the participants underwent the SECPT (Schwabe et al., 2008), during which they immersed their forearm in a basin of cold

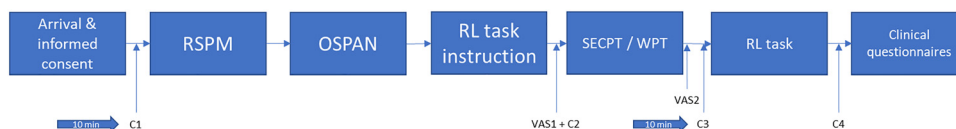


Fig. 1. Experimental protocol. After filling out the informed consent form, participants underwent the Raven's Standard Progressive Matrices Test (RSPM) and OSPAN task. The first cortisol measurement (C1) was collected 10 min after the arrival. The second (C2) was collected right after the dual-step Markov Task (RL task) instruction along with the first self-reported evaluation (VAS1). The second VAS was taken right after the stress induction (SECPT)/control (WPT) procedure, while the third cortisol measurement (C3) was collected 10 min after the procedure. The fourth measurement (C4) was collected right after the Two-Step Markov Task (RL task). Participants finished by filling out several questionnaires

water (3 °C) for three minutes while being closely observed by an experimenter and videorecorded. The other half underwent the control procedure (Warm Pressor Task; WPT), during which the water was at ambient temperature and no camera or observation was used. In total, 60 and 56 participants underwent the SECPT (30 HC and 30 PG) and the WPT (28 HC and 28 PG), respectively. Participants finished the session by filling out clinical questionnaires (see [supplementary material](#)).

Measures

Stress response assessment. Participants rated their desire to gamble, feeling of stress, and pain on visual analog scales (VAS; ranging from 0 to 10) before and after the stress induction procedure. Objective stress response measures were taken in the form of salivary cortisol levels, collected 10 min after arrival (C1), immediately following the RL-task instructions (C2), 10 min post-stress induction (C3), and after the RL-task (C4). C1 and C3 were delayed by 10 min as cortisol levels are expected to peak 10 min after stress onset (McRae et al., 2006). Salivary cortisol analysis procedure is described in the [supplementary material](#).

Two-Step Markov Task. All participants received extensive instructions for the Two-Step Markov Task (RL-task; Daw et al., 2011; Otto, Raio, et al., 2013; Wyckmans et al., 2019) and completed 10 practice trials at the end of which they answered three instruction-related questions. The instructions were reiterated in case of incorrect responses. Participants subsequently performed 200 trials of the RL-task (Fig. 2A). During the first step, they were required to choose between two fractal images displayed alongside each other on a black background. Each image commonly (70%) or rarely (30%) led to one of the two second-step stages (Fig. 2C). Participants were informed that these probabilities

would stay fixed. During the second step, they had to choose between two fractal images displayed alongside each other on a colored background (green or blue, depending on their first-step choice). Reward probabilities for each fractal image varied in function of background color. Participants were informed that these second-step probabilities would slowly fluctuate as the task progressed (following a Gaussian Random Walk with boundaries fixed at 0.25 and 0.75; SD = 0.025; Fig. 2B) to ensure continual exploration. After the second step, a 1s feedback slide was displayed to indicate whether the trial was rewarded (a 20c coin) or not (“0”). Participants had 3s to indicate their preference by pressing the letter “E” (left image) or “I” (right image) on an AZERTY keyboard. Inter-stage and inter-trial intervals lasted 1s each. Before data analyses, we automatically excluded participants who failed to answer within 3s more than 20 times, who picked the same first-step choice in 95% of the trials, or who repeated previously rewarded second-step responses at a rate lower than 50%. MB and MF algorithms predict different observable choice patterns during subsequent 2-step trials (Fig. 2D). Under MF strategy, decisions will be reinforced solely depending on the outcome of the previous trial. Conversely, choice-value will update depending on the interaction between previous outcomes and transitions under MB strategy.

Statistical analyses

Data cleaning and statistical analyses were performed using RStudio (v1.4.1103) and SPSS (v28). The sample size was estimated based on previous studies (Sebold et al., 2014; Voon et al., 2015). Accordingly, we aimed at finding a medium-sized difference between HC and PG with 80% power and 95% confidence, which required 64 participants per group.

Pre- and post-procedure cortisol concentrations were computed by averaging the first two and last two measures,

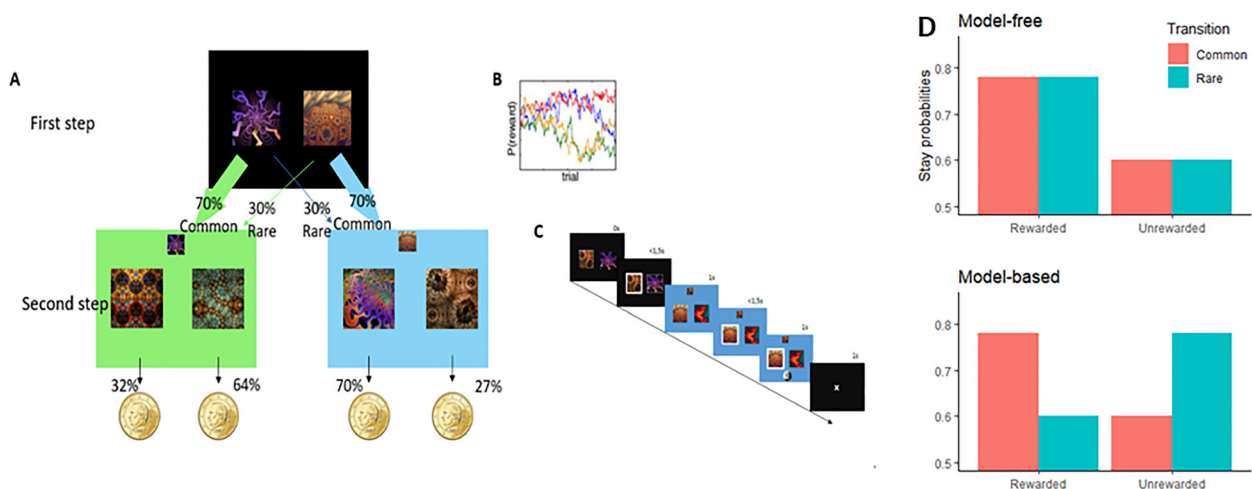


Fig. 2. A. Two-step decision task. (First step) Participants chose between the two images, leading preferentially to a green or a blue screen, according to fixed probabilities. (Second step) Subjects chose between the two images linked to probabilities to win money. Those probabilities slowly changed with time and differed according to the background color. B. Second step's changes in the probability of reward. C. Trial's design. D. Theoretical decision pattern according to a pure MF strategy and to a pure MB strategy

respectively. Results were log-transformed because of the skewed nature of their distribution (Petzold, Plessow, Goschke, & Kirschbaum, 2010). The influence of the SECPT/WPT procedures on cortisol concentrations and subjective ratings were assessed through repeated measures MANOVA. Our total sample was split into two groups depending on their cortisol response. Participants with a cortisol increase equal to or higher than 1.5 nmol L^{-1} ($N = 25$) were considered cortisol responders (Miller, Plessow, Kirschbaum, & Stalder, 2013).

We fitted choice behavior on the RL-task to a 7-parameter hybrid reinforcement learning algorithm (Daw et al., 2011) with the hBayesDM package (Ahn, Haines, & Zhang, 2017). Four MCMC chains of 6,000 samples each were run, with the first 3,000 samples as warm-up (see supplementary material). We mainly focused our analyses on the ω -parameter, which indicates the relative contribution of MB learning over MF learning during task completion. A between-subject MANOVA was used to assess the influence of stress group (Stressed vs. Non-stressed), diagnostic group (PG vs. HC), and their interaction on these 7 parameters. Multiple regression was performed to assess the continuous effect of cortisol increase, diagnostic group, and their interaction on the ω -parameter.

Normality was assessed by dividing the skewness and excess kurtosis by their respective standard deviation. Distributions with both scores $[-3.29; 3.29]$ were considered normal (Kim, 2013). The continuous variables were standardized before each regression. Demographic and clinical variables from both groups were compared with Mann-Whitney U or Welch t -tests. To avoid the confounding effect of outliers in regressions, scores that differed over 3 times the MAD from the median (Leys, Bernard, & Licata, 2013) were removed before each regression.

Ethics

Participants were recruited through advertisement and gave written informed consent to be part of the experiment. The experiment was approved by the C.H.U. Brugmann Ethics Committee (n° OM 026) and performed according to the Declaration of Helsinki.

RESULTS

Sample characteristics

Our final sample consisted of 116 participants, 58 HC, and 58 PG. Both groups were matched for age, gender (only males), and education. Table 1 depicts the demographic and clinical variables of PG and HC as well as between-group comparisons and Cronbach's alpha for each questionnaire. PG displayed significantly lower scores on working memory and fluid intelligence assessments than HC, as well as significantly more psychiatric comorbidities, state-anxiety, depressive symptoms, and reward sensitivity, which were included as covariates in separate analyses (see supplementary material).

A repeated measures MANOVA was performed to evaluate the effect of stress induction procedure (between-subject; SECPT vs. WPT), time (within-subject; before vs. after the procedure) and their interaction on salivary cortisol concentrations, as well as self-reported stress, desire to gamble (craving), and pain measures (Fig. 3). The interaction effect on the combined dependent variables was significant ($F(4, 111) = 7.49, P < 0.001$; Wilks' $\Lambda = 0.73, \eta^2_p = 0.28$). Univariate analyses showed a significant effect of the interaction for each dependent variable (Table 2), indicating a higher increase in cortisol concentration and subjective ratings among participants undergoing the SECPT than the WPT. No group effect (PG vs. HC) was found on pre- to post-procedure differences in cortisol concentration and subjective ratings ($F(4, 111) = 0.96, P = 0.43$; Wilks' $\Lambda = 0.97, \eta^2_p = 0.03$).

We divided our sample into two groups according to their cortisol response: participants with a cortisol elevation higher than 1.5 nmol L^{-1} were considered as demonstrating a cortisol response (Miller et al., 2013). The proportion of PG and HC did not significantly differ between the responders (10 PG and 15 HC) and the non-responders (48 PG and 43 HC) groups ($X^2(1) = 1.28, P = 0.26$). The responder group contained significantly more participants who underwent the SECPT than the WPT ($X^2(1) = 5.25, P = 0.02$).

Analyses of choice behavior

MB and MF algorithms predict differential probabilities of repeating the previous trial's first step's choice depending on its outcome and transition. These differential probabilities are illustrated in Fig. 4.

A between-subject MANOVA (Fig. 5) was performed to evaluate the effects of diagnostic group (PG vs. HC), stress response (responders vs. non-responders), and their interaction on each computational parameter. Results indicated a significant interaction effect on the combined dependent variables ($F(7, 106) = 4.23, P < 0.001$; Wilks' $\Lambda = 0.78, \eta^2_p = 0.22$). Univariate analyses (displayed in supplementary table 1) indicated a significant effect of the interaction on the ω -parameter ($F(1, 112) = 7.79, P = 0.006, \eta^2_p = 0.07$). Simple effect analyses with Bonferroni corrections showed that the ω -parameter was significantly lower among responders HC than non-responders HC ($F(1, 56) = 12.90, P < 0.01, \eta^2_p = 0.19$), but no significant difference was observed between responder and non-responder PG ($F(1, 56) = 0.22, P = 0.98, \eta^2_p = 0.004$). The ω -parameter was significantly lower among PG than HC in the non-responder group ($F(1, 89) = 9.32, P = 0.01, \eta^2_p = 0.10$), but not in the responder group ($F(1, 23) = 2.92, P = 0.34, \eta^2_p = 0.11$). In six additional ANCOVA, we added the OSPAN, Raven, SCL90R, BDI, STAI-YA, and RS scores as covariables. The interaction effect remained significant ($P < 0.05$, see supplementary table 2).

Interestingly, the interaction also showed a significant effect on the ρ -parameter, ($F(1,112) = 8.78, P = 0.004, \eta^2_p = 0.07$), an index quantifying perseveration. It was

Table 1. Descriptive scores of each clinical variable and their between-subject difference. Clinical variables assessed alcohol use disorder symptoms (AUDIT), gambling disorder symptoms (SOGS), number of DSM items (DSM), working memory performances (OSPAN), Raven score, psychiatric comorbidities (SCL90R), depressive symptoms (BDI), positive affects (PANAS pos), negative affects (PANAS neg), state-anxiety (STAI-YA), trait-anxiety (STAI-YB), adverse live events (SRRS), sensibility to punishment (PS), sensibility to reward (RS), routine tendencies (CoH R), automatism tendencies (CoH A), negative urgency (NU), positive urgency (PU), lack of premeditation (LoPr), lack of perseverance (LoPe), and sensation seeking (SS). Groups are compared with Welsh's *t*-tests or Mann-Whitney U according to the distribution of their scores. Cronbach's alphas are displayed when applicable. Significant group differences are displayed in bold

| Variable | PG (<i>n</i> = 58) | HC (<i>n</i> = 58) | Test | Alpha |
|----------------|---------------------------------|----------------------------------|--|-------|
| Age | 29.22 (8.39) Mdn = 28 | 31.16 (9.86) Mdn = 28.5 | U (116) = 1,498.5, <i>P</i> = 0.31 | N/A |
| Study Level | 13.88 (2.85) Mdn = 15 | 14.66 (2.4) Mdn = 15 | U (115) = 1,400, <i>P</i> = 0.14 | N/A |
| AUDIT | 8.87 (7.75) Mdn = 8 | 6.34 (4.49) Mdn = 5 | U (114) = 1748, <i>P</i> = 0.21 | 0.86 |
| SOGS | 9.05 (2.63) Mdn = 8.5 | 0.36 (1.1) Mdn = 0 | U(79) = 1,624, <i>P</i> < 0.001 | 0.86 |
| DSM | 3.97 (2.62) Mdn = 4 | 0.11 (0.32) Mdn = 0 | U(90) = 1881.5, <i>P</i> < 0.001 | 0.88 |
| Smoker (Y N) | 22 36 | 19 39 | X (1) = 0.34, <i>P</i> = 0.56 | N/A |
| OSPAN | 0.76 (0.13) Mdn = 0.77 | 0.8 (0.16) Mdn = 0.85 | U(116) = 1,207, <i>P</i> = 0.01 | N/A |
| Raven | 4.5 (1.88) Mdn = 5 | 5.33 (1.94) Mdn = 6 | U(116) = 1,214.5, <i>P</i> = 0.01 | N/A |
| SCL90R | 55.67 (49.01) Mdn = 44 | 35.1 (32.54) Mdn = 30.5 | U(116) = 2,119.5, <i>P</i> = 0.02 | 0.97 |
| BDI | 6.28 (5.94) Mdn = 5 | 3.71 (4.12) Mdn = 2 | U(116) = 2090.5, <i>P</i> = 0.02 | 0.88 |
| PANAS pos | 34.31 (7.8) Mdn = 34 | 36.21 (6.27) Mdn = 37 | <i>t</i> (108.98) = 1.44, <i>P</i> = 0.15 | 0.85 |
| PANAS neg | 20.6 (8.25) Mdn = 18 | 18.1 (6.85) Mdn = 16.5 | U (116) = 1970, <i>P</i> = 0.11 | 0.89 |
| STAI-YA | 36.54 (10.95) Mdn = 35 | 31.03 (8.83) Mdn = 29 | <i>t</i>(107.39) = 2.97, <i>P</i> = 0.004 | 0.90 |
| STAI-YB | 42.36 (12.3) Mdn = 41.5 | 39.72 (11.18) Mdn = 38.5 | <i>t</i> (110.12) = 1.2, <i>P</i> = 0.24 | 0.91 |
| SRRS | 244.96 (169.44) Mdn = 201 | 242.58 (206.81) Mdn = 184 | U (115) = 1,692.5, <i>P</i> = 0.58 | N/A |
| PS | 41.84 (11.54) Mdn = 41 | 39.26 (9.46) Mdn = 39 | <i>t</i> (108.08) = 1.31, <i>P</i> = 0.19 | 0.90 |
| RS | 43.98 (10.95) Mdn = 44 | 39 (7.89) Mdn = 40 | U(116) = 2,127, <i>P</i> = 0.01 | 0.88 |
| CoH R | 52.44 (12.69) Mdn = 53 | 45.77 (12.05) Mdn = 45 | <i>t</i> (31.46) = 1.63, <i>P</i> = 0.11 | 0.86 |
| CoH A | 33.94 (12.14) Mdn = 34.5 | 27 (9.42) Mdn = 22 | <i>t</i> (27.28) = 1.91, <i>P</i> = 0.07 | 0.88 |
| NU | 8.76 (3.35) Mdn = 9 | 7.67 (3.2) Mdn = 8 | <i>t</i> (113.76) = 1.78, <i>P</i> = 0.08 | 0.85 |
| PU | 10.17 (3.01) Mdn = 10.5 | 9.9 (2.59) Mdn = 10 | <i>t</i> (111.48) = 0.53, <i>P</i> = 0.60 | 0.75 |
| LoPr | 9.12 (3.5) Mdn = 8 | 8.47 (3.13) Mdn = 8 | U (116) = 1828, <i>P</i> = 0.42 | 0.88 |
| LoPe | 9.24 (3.63) Mdn = 9 | 9.07 (3.4) Mdn = 8 | <i>t</i> (113.51) = 0.26, <i>P</i> = 0.79 | 0.92 |
| SS | 10.38 (3.28) Mdn = 11 | 9.78 (3.5) Mdn = 10 | <i>t</i> (113.51) = 0.96, <i>P</i> = 0.34 | 0.86 |

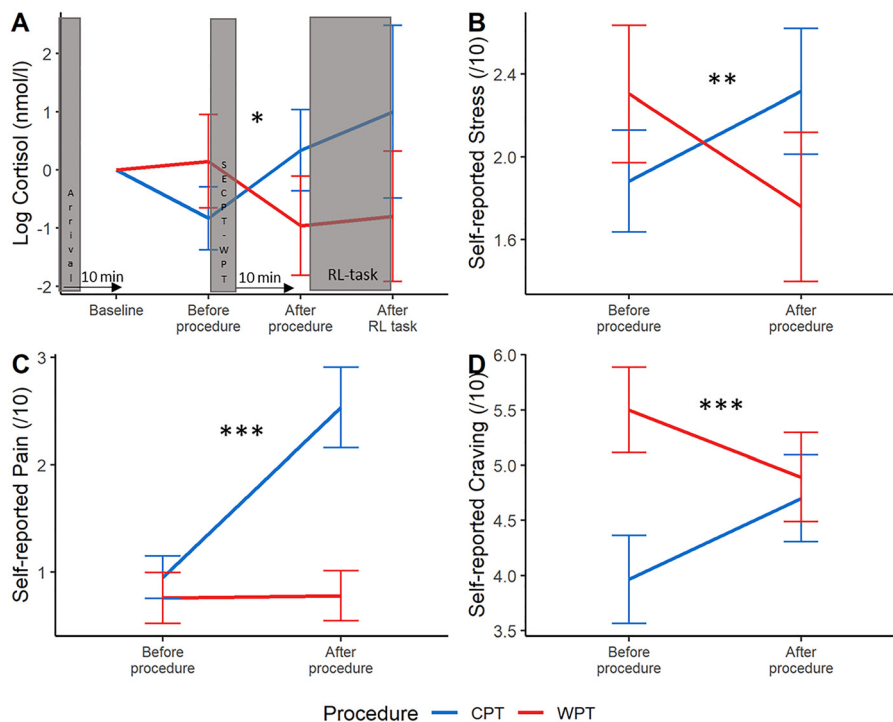


Fig. 3. Cortisol and self-reported measures for the group that underwent the cold-pressor task (SECPT) and the Warm-Pressor Task (WPT). Graphs show mean values \pm SE. *P*-values of the effect of interaction between time and procedure are reported * *P* < 0.05, ** *P* < 0.01, *** *P* < 0.001

Table 2. Mean (sd) cortisol concentrations (not log-transformed for interpretability) and self-reported measures, before and after the pressor task. Results of the univariate interactions between time and procedure type are displayed

| | Cold Pressor Task (<i>n</i> = 60) | | Warm Pressor Task (<i>n</i> = 58) | | Univariate results |
|--|------------------------------------|-----------------|------------------------------------|-----------------|---|
| | Before procedure | After procedure | Before procedure | After procedure | Time*procedure |
| Cortisol concentration (nmol l ⁻¹) | 5.94 (4.11) | 7.02 (6.70) | 7.63 (5.38) | 6.68 (6.53) | $F(1,114) = 9.41, P = 0.003, \eta_p^2 = 0.08$ |
| Stress (/10) | 1.88 (1.91) | 2.32 (2.35) | 2.30 (2.48) | 1.76 (2.70) | $F(1, 114) = 8.86, P = 0.004, \eta_p^2 = 0.07$ |
| Pain (/10) | 0.95 (1.53) | 2.53 (2.90) | 0.76 (1.78) | 0.78 (1.76) | $F(1, 114) = 17.69, P < 0.001, \eta_p^2 = 0.13$ |
| Craving (/10) | 3.32 (3.11) | 4.12 (3.20) | 4.07 (3.4) | 3.79 (3.25) | $F(1, 114) = 10.61, P = 0.001, \eta_p^2 = 0.09$ |

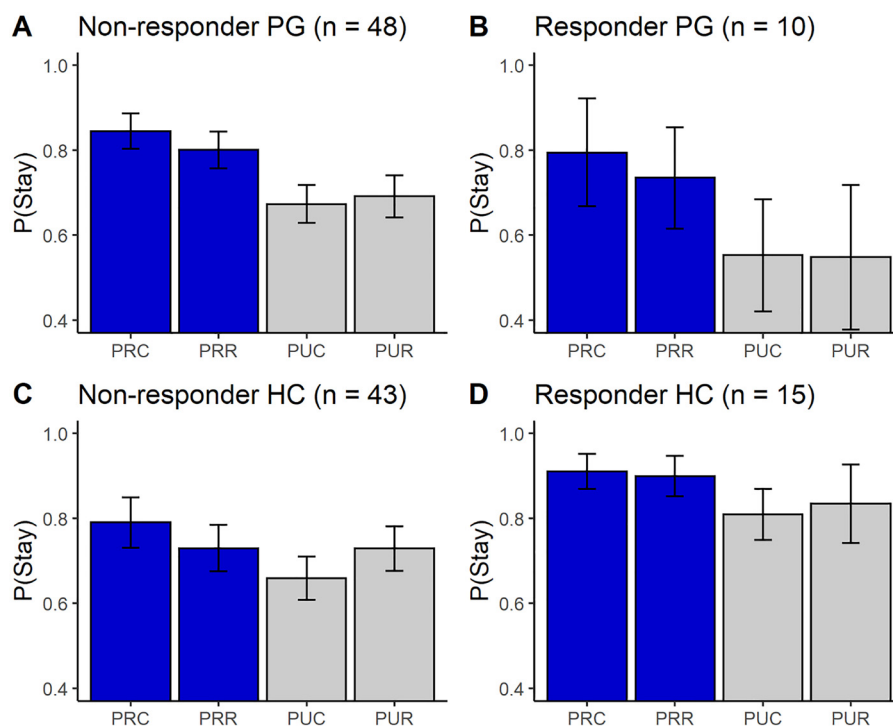


Fig. 4. Probability to maintain a first step choice in problem gamblers (PG) and healthy controls (HC) in trial following: a reward and a common transition (PRC), a reward and a rare transition (PRR), a loss and a common transition (PUC), as well as a loss and a rare transition (PUR). Graphs show mean values ± 2 SE

significantly lower among PG than HC in the stressed group ($F(1,23) = 7.47, P = 0.047, \eta_p^2 = 0.25$), but not in the non-responder group ($F(1,89) = 0.003, P > 0.99, \eta_p^2 < 0.001$). Additionally, the ρ -parameter was significantly higher among responder than non-responder HC ($F(1,56) = 8.51, P = 0.02, \eta_p^2 = 0.13$), but no significant difference was observed between responder and non-responder PG ($F(1,56) = 1.70, P = 0.59, \eta_p^2 = 0.03$).

The ω -parameter was regressed on the diagnostic group (DG; dummy coded: PG = -1 vs HC = 1), cortisol increase (dCort; continuous), and their interaction (Fig. 6). The model was significant ($F(3,93) = 4.37, P < 0.01, R^2 = 0.12$), with a significant dCort*DG interaction ($\beta = -0.56(0.19), P = 0.004$) indicating that stress was more deleterious on the ω -parameter among HC than PG. Bayesian Kendall's correlations indicated

strong evidence in favor of a negative correlation between the ω -parameter and dCort among HC ($\rho(58) = -0.35, BF = 15.1$), and moderate evidence against this correlation among PG ($\rho(58) = -0.05, BF = 0.32$). In six additional full factorial regressions, we added the OSPAN, Raven, SCL90R, BDI, STAI-YA, and RS scores as predictors. The interaction dCort*DG remained significant ($P < 0.05$) in each regression, and no additional predictor showed a significant main effect or interaction with DG and dCort ($P > 0.05$).

DISCUSSION AND CONCLUSIONS

The objective of the present study was to investigate how MB and MF were orchestrated in GD according to a hybrid-

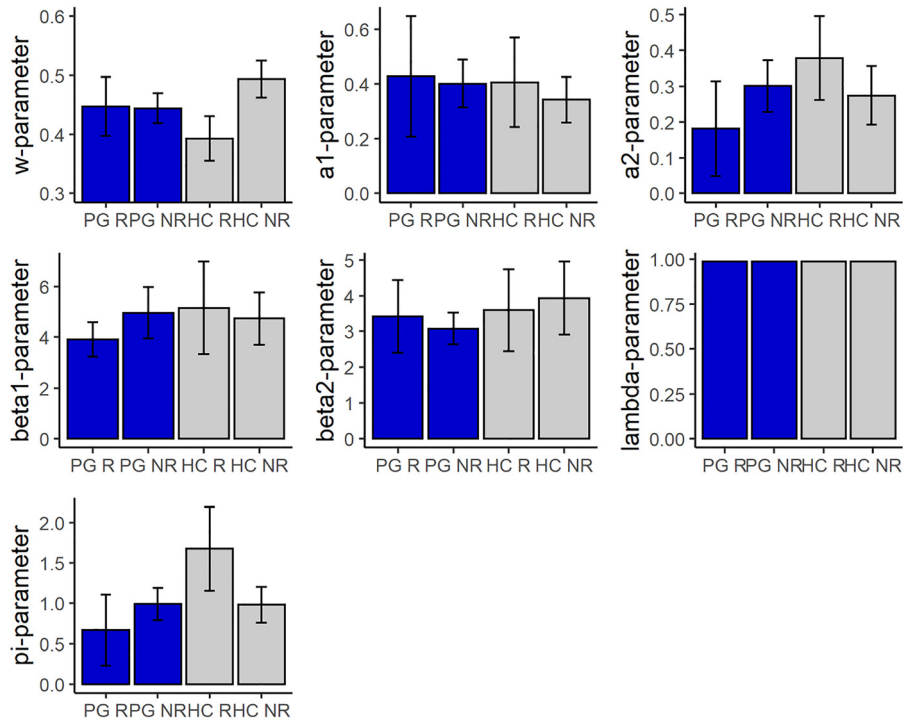


Fig. 5. Mean of the computational parameters among responder pathological gamblers (PG S), non-responder problem gamblers (PG NS), responder controls (HC S), and non-responder controls (HC NS). Graphs show mean values ± 2 SE

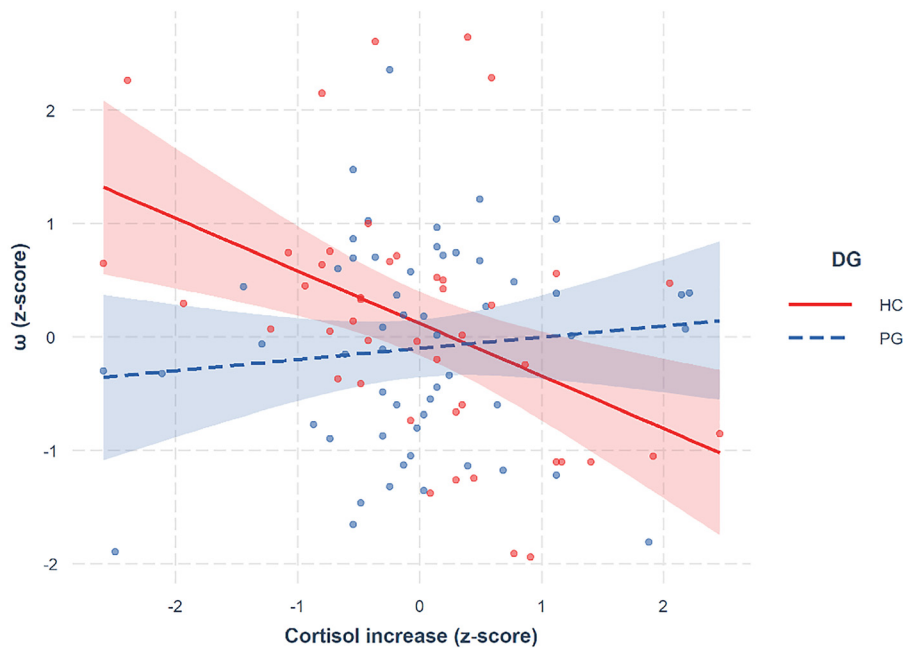


Fig. 6. Effect of the cortisol increase on ω -parameter in problem gamblers (PG) and healthy controls (HC), as determined by the computational model. Each continuous score was standardized, 95% IC is displayed around each regression line

RL 7-parameter computational model and whether PG were more sensitive than HC to the deleterious influence of stress on this dynamic. This is the first study to our knowledge that investigated the impact of stress on decision making in PG by employing a computational modelling approach that goes beyond the experimental approaches used in previous studies (e.g., Wyckmans et al., 2019).

Results indicated that the two groups had comparable baseline levels and stressor-induced elevation of salivary cortisol. A deleterious effect of enhanced cortisol level on MB/MF orchestration, characterized as a decrease of MB learning relative to MF learning, was found in HC but not in PG. Among non-stressed participants, results indicated that the balance between MB and MF contributions, formalized

by the ω -parameter, was biased towards MF learning in PG. This pattern held even while controlling for the lower verbal working memory performance found in PG. These results not only aligned with previous study with PG and MF/MB (Wyckmans et al., 2019) but also converged with similar studies conducted in subjects with substance-use disorders (Chen et al., 2021; Doñamayor et al., 2018; Sebold et al., 2017; Voon et al., 2015). Thus, these results support the idea that the inflexible repeated behaviors observed in PG might be underpinned by a greater inclination towards a habitual model of decision-making (Lucantonio, Caprioli, & Schoenbaum, 2014). This tendency indicates a higher reliance towards MF-strategies, wherein choices are based essentially upon past rewards, with little consideration for changes in contingency between response and consequence.

Despite similar objective (cortisol) and subjective (ratings) responses to stress between the PG and HC groups, cortisol response was only associated with changes on the ω -parameter among HC (Otto, Raio et al., 2013; Radenbach et al., 2015). Bayesian analyses further supported that the stress-induced cortisol increase was not associated with modulations in MB/MF orchestration in PG. A first explanation could be that the overall poorer working memory performance among PG and their general bias towards MF learning might explain the absence of a deleterious effect of the acute stressor on the MB/MF orchestration (i.e., floor effect). Indeed, evidence suggests that MB is partially supported by working memory (Culbreth, Westbrook, Daw, Botvinick, & Barch, 2016), as it involves the learning of the action-outcome-transition contingencies and the planning of the following choices (Otto, Gershman, Markman, & Daw, 2013). Accordingly, the shift from MB towards MF learning among HC following stress might be the consequence of a stress-induced decrease in working memory performance (Wirz et al., 2018). Hence, the cortisol increase might have limited influence on PG's already-impaired working memory, thereby restricting the modulation by stress on MB/MF orchestration in this population. However, group differences in the stress-induced modulation of learning strategies remained significant even when accounting for the effect of working memory performances, which suggests that this interpretation at best partially explain our results.

Another interpretation would be that the stress induced in the current study might have been insufficient among PG to elicit a behavioral adaptation during the Two-Step Markov Task. Despite similar cortisol increases across HC and PG, the latter are accustomed to functioning in highly stressful environments (Buchanan et al., 2020), which could be the cause of a persistent shift from goal-directed to habitual behavior (Dias-Ferreira et al., 2009). While we found no association between the ω -parameter and the social readjustment rating scale (SRRS; Holmes & Rahe, 1967), the latter does not account for stressful events specific to PG, such as financial preoccupation (Langham et al., 2015; Li, Browne, Rawat, Langham, & Rockloff, 2017), meaning that PG chronic stress might have been underestimated. Specific measures of chronic stress are therefore needed to further

validate the influence of chronic stress on the behavioral response to acute stress in GD.

Additionally, interoception, which critically contributes to subjective emotional experience by modulating the subjective strength of bodily responses (i.e., the ability to perceive internal state) might have been abnormally inaccurate in PG, as a recent study suggests (Moccia et al., 2021). Also, in line with impaired interoception, alexithymic tendencies (e.g., difficulty in identifying feelings) are greater in PG than in healthy participants (Noël et al., 2018). This lack of accurate perception of internal bodily states related to emotions, such as the somatic experience of stress, might contribute to downplaying the influence of stress-induction on diverse behavioral responses among PG, including MF/MB orchestration. Of note, the visual analog scales used to score perceived stress was not meant to adequately capture interoceptive processing. A multidomain, multidimensional approach to interoception including cardiac, gastric, and respiratory assessments should be preferred in future studies to examine the role of this function in the association between acute stress and reinforcement learning in PG (Murphy, Catmur, & Bird, 2018; Wang et al., 2019).

At the theoretical level, low working (Albein-Urios, Martinez-González, Lozano, & Clark, 2012; Brevers et al., 2012) and impaired model-based strategy (Wyckmans et al., 2019) found in PG may reflect the expression of two neurocognitive pathways driving maladaptive and persistent gambling behaviors, i.e. a “stop now” process mediated by the dorsolateral prefrontal cortex, and automatic responses to “must do” stimuli mediated by the dorsal striatum (Brand, 2022). An exciting line of future research would be to examine how a third system, the “feel better” pathway (i.e. positive and negative reinforcement) mediated by the ventral striatum, interacts with the “stop now” and “must do” processes to promote the escalation of gambling behaviors (Brand, 2022) see also (Noël et al., 2013). Thus, by considering two critical pathways of gambling disorder, the present study represents the first step toward a more nuanced multidimensional neurocognitive model of behavioral addiction.

Exploratory analyses showed that alongside the ω -parameter, the perseveration parameter (the maintenance of a choice irrespective of the previous outcomes and transitions) was higher among HC who showed a cortisol increase after the stress-inducing procedure. Stress-induced perseveration has never been reported and only one study found overall higher perseveration in binge-eating disorder and interpreted it as a lower involvement of cognitive control processes (Voon et al., 2015). While the ρ -parameter did not significantly differ between HC and PG, exacerbated perseveration among HC following acute stress might result from the deleterious effect of cortisol increase on the prefrontal cortex (Wirz et al., 2018).

As computational modelling allows us to quantify inter-individual differences in mechanisms underlying addiction development and maintenance, they hold great promises to promote targeted clinical interventions (Heinz et al., 2017). For instance, contingency management therapies, which are based on the principles of reinforcement that provide people



tangible rewards (e.g., money) for evidence of behavioral change (e.g., maintaining abstinence), show a high level of efficacy in substance use disorders (Davis et al., 2016; Kim & Hodgins, 2018; Mantzari et al., 2015), might specifically help PG displaying low MB capacities, as MB strategy use is increased when higher incentives are offered (Kool, Gershman, & Cushman, 2017; Patzelt, Kool, Millner, & Gershman, 2019). However, clinical trials in PG are underway, and the beneficial impact of contingency management on treatment attendance and gambling abstinence needs to be confirmed (Christensen, Witcher, Hudson-Breen, & Ofori-Dei, 2018). Episodic future thinking training represents another promising way to reduce addiction severity (Snider, LaConte, & Bickel, 2016). As outlined, PG encounter difficulties in mentally simulating future events in a vivid manner (Noël, Saeremans, Kornreich, Jaafari, & D'Argembeau, 2017). Episodic foresight might support model-based control in the sense that the agent prospectively evaluates actions based on their potential outcomes. However, little is known about the relationship between learning an internal model to prospectively make decisions and the capacity to mentally navigate the future. Finally, several brain stimulation techniques (tDCS, rTMS) might target critical neural networks involved in MB/MF orchestration, thus representing a promising way to treat gambling addiction (Smittenaar, FitzGerald, Romei, Wright, & Dolan, 2013; Weissengruber, Lee, O'Doherty, & Ruff, 2019).

The present study is not exempt from limitations. First, only a third of participants undergoing the SECPT showed a cortisol increase, lowering our power to observe small cortisol-induced behavioral changes. Two reasons for this low response rate compared to the initial validation study (Schwabe et al., 2008) can be put forward, our participants were older, and the experimenter in charge of the social assessment was not of the opposite sex. Alternative but longer stressors like the Trier Social Stress Test (Foley & Kirschbaum, 2010) or the Maastricht Acute Stress Test (Smeets, van Ruitenbeek, Hartogsveld, & Quaedflieg, 2019) generating a higher cortisol elevation may be recommended for further investigation, but only if the longer duration is not an issue. Nevertheless, both the HC and the PG groups displayed similar levels of the cortisol stress response, and the influence of acute stress on HC's decisional processes was consistent with the literature (Otto, Raio et al., 2013; Radenbach et al., 2015). Additionally, false beliefs about action and reward contingencies prevalent in GD, such as the hot-hand effect or gambler's fallacy (Joukhador, Blaszczynski, & Maccallum, 2004), have not been taken into account. As MB and MF are high-level processes, emerging from many sub-computations (Collins & Cockburn, 2020; Daw, 2018), these alternative decision strategies could have severely biased our strictly dichotomized behavioral assessment (Mohr et al., 2018). Further iterations of this paradigm in GD should therefore control for these false beliefs and assess how they might be related to weakened MB learning. Finally, we focused only on males in the present research, primarily because we aimed to discern the effects of experimentally induced stress on cortisol without other physiological confounds, since the cortisol response to stress among females remains susceptible to

variations across the phases of their menstrual (Kirschbaum, Wüst, & Hellhammer, 1992; Montero-López et al., 2018; Ozgocer, Ucar, & Yildiz, 2017). However, examination of reinforcement learning strategies in female PGs is warranted, particularly due to the finding that the phase of the menstrual cycle modulates reward-related neural function in females (Dreher et al., 2007). For instance, a relationship among ovarian hormones, negative mood, and gambling behavior (e.g., ovulation is accompanied by riskier gambling behaviors) has been recently reported (Joyce et al., 2019). Future research should investigate how physiological stress, for instance, in response to a social stressor across the female menstrual cycle, impacts model-based and model-free evaluations. This line of research is all the more urgent as the study of female-specific factors in behavioral addiction remains understudied (Joyce, Good, Tibbo, Brown, & Steward, 2021).

To summarize, the present study adds evidence towards the hypothesis that GD includes a compulsive component with a propensity towards inflexible habits formation when facing new decisions, arising from a bias towards MF control. PG's resistance to acute stress might be caused by several non-exclusive factors, namely low working-memory capacities, high chronic stress, and deleterious interoceptive accuracy. Despite a normal cortisol response to stress, we found an abnormal pattern of the modulation of stress on the relative weighing of goal-directed and habitual action strategies in GD.

Funding sources: This work was funded by the Belgian Fund for Scientific Research (PDR T10146.18).

Authors' contribution: Study concept and design: FW, XN. Writing original draft: FW. Review and editing: FW, MS, NB, OR, CK, LV, AB, TB, XN. Analysis and interpretation of data: FW, DG, VC. Statistical analysis: FW. Obtained funding: XN. Study supervision: XN. All authors had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflict of interest: The authors declare no competing interests.

Data availability: All the data and processing scripts are available on a GitHub folder: https://github.com/FWyckmans/MBMF_Study2.

SUPPLEMENTARY MATERIAL

Supplementary data to this article can be found online at <https://doi.org/10.1556/2006.2022.00059>.

REFERENCES

Ahn, W.-Y., Haines, N., & Zhang, L. (2017). Revealing neuro-computational mechanisms of reinforcement learning and



- decision-making with the hBayesDM package. *Computational Psychiatry*, 1(0), 24–57. https://doi.org/10.1162/CPSY_a_00002.
- Albein-Urios, N., Martinez-González, J. M., Lozano, O., & Clark, L. (2012). Comparison of impulsivity and working memory in cocaine addiction and pathological gambling: Implications for cocaine-induced neurotoxicity. *Drug and Alcohol Dependence*, 126(1–2), 1–6. <https://doi.org/10.1016/j.drugalcdep.2012.03.008>.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*, 5th ed.
- Anstrom, K. K., & Woodward, D. J. (2005). Restraint increases dopaminergic burst firing in awake rats. *Neuropsychopharmacology*, 30(10), 1832–1840. <https://doi.org/10.1038/sj.npp.1300730>.
- Arnsten, A. F. T. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience*, 10(6), 410–422. <https://doi.org/10.1038/nrn2648>.
- Biback, C., & Zack, M. (2015). The relationship between stress and motivation in pathological gambling: A focused review and analysis. *Current Addiction Reports*, 2(3), 230–239. <https://doi.org/10.1007/s40429-015-0064-9>.
- Bilker, W. B., Hansen, J. A., Bressinger, C. M., Richard, J., Gur, R. E., & Gur, R. C. (2012). Development of abbreviated nine-item forms of the Raven's standard progressive Matrices test. *Assessment*, 19(3), 354–369. <https://doi.org/10.1177/1073191112446655>.
- Brand, M. (2022). Can internet use become addictive? *Science*, 376(6595), 798–799. <https://doi.org/10.1126/science.abn4189>.
- Brevers, D., Cleeremans, A., Verbruggen, F., Bechara, A., Kornreich, C., Verbanck, P., & Noël, X. (2012). Impulsive action but not impulsive choice determines problem gambling severity. *PLoS One*, 7(11), e50647. <https://doi.org/10.1371/journal.pone.0050647>.
- Buchanan, T. W., McMullin, S. D., Baxley, C., & Weinstock, J. (2020). Stress and gambling. *Current Opinion in Behavioral Sciences*, 31, 8–12. <https://doi.org/10.1016/j.cobeha.2019.09.004>.
- Chen, H., Mojtahedzadeh, N., Belanger, M. J., Nebe, S., Kuitunen-Paul, S., Sebold, M., ... Smolka, M. N. (2021). Model-based and model-free control predicts alcohol consumption developmental trajectory in young adults – a three-year prospective study. *Biological Psychiatry*, S0006322321000767. <https://doi.org/10.1016/j.biopsych.2021.01.009>.
- Christensen, D. R., Witcher, C. S. G., Hudson-Breen, R., & Ofori-Dei, S. (2018). Piloting the addition of contingency management to best practice counselling as an adjunct treatment for rural and remote disordered gamblers: Study protocol. *BMJ Open*, 8(4), e018804. <https://doi.org/10.1136/bmjopen-2017-018804>.
- Collins, A. G. E., & Cockburn, J. (2020). Beyond dichotomies in reinforcement learning. *Nature Reviews Neuroscience*, 21(10), 576–586. <https://doi.org/10.1038/s41583-020-0355-6>.
- Culbreth, A. J., Westbrook, A., Daw, N. D., Botvinick, M., & Barch, D. M. (2016). Reduced model-based decision-making in Schizophrenia. *Journal of Abnormal Psychology*, 125(6), 777–787. <https://doi.org/10.1037/abn0000164>.
- Davis, D. R., Kurti, A. N., Skelly, J. M., Redner, R., White, T. J., & Higgins, S. T. (2016). A review of the literature on contingency management in the treatment of substance use disorders, 2009–2014. *Preventive Medicine*, 92, 36–46. <https://doi.org/10.1016/j.ypmed.2016.08.008>.
- Daw, N. D. (2018). Are we of two minds? *Nature Neuroscience*, 21(11), 1497–1499. <https://doi.org/10.1038/s41593-018-0258-2>.
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, 69(6), 1204–1215. <https://doi.org/10.1016/j.neuron.2011.02.027>.
- Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, 8(12), 1704–1711. <https://doi.org/10.1038/nn1560>.
- Dias-Ferreira, E., Sousa, J. C., Melo, I., Morgado, P., Mesquita, A. R., Cerqueira, J. J., ... Sousa, N. (2009). Chronic stress causes frontostriatal reorganization and affects decision-making. *Science*, 325(5940), 621–625. <https://doi.org/10.1126/science.1171203>.
- Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, 80(2), 312–325. <https://doi.org/10.1016/j.neuron.2013.09.007>.
- Doñamayor, N., Strelchuk, D., Baek, K., Banca, P., & Voon, V. (2018). The involuntary nature of binge drinking: Goal directedness and awareness of intention: Automaticity of binge drinking. *Addiction Biology*, 23(1), 515–526. <https://doi.org/10.1111/adb.12505>.
- Dreher, J. C., Schmidt, P. J., Kohn, P., Furman, D., Rubinow, D., & Berman, K. F. (2007). Menstrual cycle phase modulates reward-related neural function in women. *PNAS*, 104(7), 2465–2470. <https://doi.org/10.1073/pnas.0605569104>.
- Edgerton, J. D., Keough, M. T., & Roberts, L. W. (2018). Co-development of problem gambling and depression symptoms in emerging adults: A parallel-process latent class growth model. *Journal of Gambling Studies*, 34(3), 949–968. <https://doi.org/10.1007/s10899-018-9760-4>.
- Enoch, M.-A. (2011). The role of early life stress as a predictor for alcohol and drug dependence. *Psychopharmacology*, 214(1), 17–31. <https://doi.org/10.1007/s00213-010-1916-6>.
- Everitt, B. J., & Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: From actions to habits to compulsion. *Nature Neuroscience*, 8(11), 1481–1489. <https://doi.org/10.1038/nn1579>.
- Everitt, B. J., & Robbins, T. W. (2016). Drug addiction: Updating actions to habits to compulsions ten years on. *Annual Review of Psychology*, 67, 23–50. <https://doi.org/10.1146/annurev-psych-122414-033457>.
- Foley, P., & Kirschbaum, C. (2010). Human hypothalamus-pituitary-adrenal axis responses to acute psychosocial stress in laboratory settings. *Neuroscience and Biobehavioral Reviews*, 35(1), 91–96. <https://doi.org/10.1016/j.neubiorev.2010.01.010>.
- Gillan, C. M., Otto, A. R., Phelps, E. A., & Daw, N. D. (2015). Model-based learning protects against forming habits. *Cognitive, Affective & Behavioral Neuroscience*, 15(3), 523–536. <https://doi.org/10.3758/s13415-015-0347-6>.
- Goldstein, R. Z., & Volkow, N. D. (2002). Drug addiction and its underlying neurobiological basis: Neuroimaging evidence for the involvement of the frontal cortex. *The American Journal of Psychiatry*, 159(10), 1642–1652. <https://doi.org/10.1176/appi.ajp.159.10.1642>.



- Grant, J. E., & Chamberlain, S. R. (2019). Cold pressor pain and gambling disorder: Implications for the opioid system. *CNS Spectrums*, 25(3), 426–433. <https://doi.org/10.1017/S109285291900107X>.
- Groman, S. M., Massi, B., Mathias, S. R., Lee, D., & Taylor, J. R. (2019). Model-free and model-based influences in addiction-related behaviors. *Biological Psychiatry*, 85(11), 936–945. <https://doi.org/10.1016/j.biopsych.2018.12.017>.
- Heinz, A., Deserno, L., Zimmermann, U. S., Smolka, M. N., Beck, A., & Schlagenhauf, F. (2017). Targeted intervention: Computational approaches to elucidate and predict relapse in alcoholism. *NeuroImage*, 151, 33–44. <https://doi.org/10.1016/j.neuroimage.2016.07.055>.
- Holmes, T. H., & Rahe, R. H. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research*, 11(2), 213–218. [https://doi.org/10.1016/0022-3999\(67\)90010-4](https://doi.org/10.1016/0022-3999(67)90010-4).
- Joukhaider, J., Blaszczynski, A., & Maccallum, F. (2004). Superstitious beliefs in gambling among problem and non-problem gamblers: Preliminary data. *Journal of Gambling Studies*, 20(2), 171–180. <https://doi.org/10.1023/B:JOGS.0000022308.27774.2b>.
- Joyce, K. M., Good, K. P., Tibbo, P., Brown, J., & Stewart, S. H. (2021). Addictive behaviors across the menstrual cycle: A systematic review. *Archives of Women's Mental Health*, 24(4), 529–542. <https://doi.org/10.1007/s00737-020-01094-0>.
- Joyce, K. M., Hudson, A., O'Connor, R. M., Goldstein, A. L., Ellery, M., McGrath, D. S., Perrot, T. S., & Stewart, S. H. (2019). Retrospective and prospective assessments of gambling-related behaviors across the female menstrual cycle. *Journal of Behavioral Addictions*, 8(1), 135–145. <https://doi.org/10.1556/2006.7.2018.133>.
- Kim, H. S., & Hodgins, D. C. (2018). Component model of addiction treatment: A pragmatic transdiagnostic treatment model of behavioral and substance addictions. *Frontiers in Psychiatry*, 9. <https://www.frontiersin.org/articles/10.3389/fpsy.2018.00406>.
- Kim, H.-Y. (2013). Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics*, 38(1), 52–54. <https://doi.org/10.5395/rde.2013.38.1.52>.
- Kirschbaum, C., Wüst, S., & Hellhammer, D. (1992). Consistent sex differences in cortisol responses to psychological stress. *Psychosomatic Medicine*, 54(6), 648–657. <https://doi.org/10.1097/00006842-199211000-00004>.
- Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost-benefit arbitration between multiple reinforcement-learning systems. *Psychological Science*, 28(9), 1321–1333. <https://doi.org/10.1177/0956797617708288>.
- Langham, E., Thorne, H., Browne, M., Donaldson, P., Rose, J., & Rockloff, M. (2015). Understanding gambling related harm: A proposed definition, conceptual framework, and taxonomy of harms. *BMC Public Health*, 16(1). <https://doi.org/10.1186/s12889-016-2747-0>.
- Lesieur, H. R., & Blume, S. B. (1987). The South Oaks gambling screen (SOGS): A new instrument for the identification of pathological gamblers. *American Journal of Psychiatry*, 144(9). <https://doi.org/10.1176/ajp.144.9.1184>.
- Leys, C., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764–766. <https://doi.org/10.1016/j.jesp.2013.03.013>.
- Li, E., Browne, M., Rawat, V., Langham, E., & Rockloff, M. (2017). Breaking bad: Comparing gambling harms among gamblers and affected others. *Journal of Gambling Studies*, 33(1), 223–248. <https://doi.org/10.1007/s10899-016-9632-8>.
- Lucantonio, F., Caprioli, D., & Schoenbaum, G. (2014). Transition from “model-based” to “model-free” behavioral control in addiction: Involvement of the orbitofrontal cortex and dorso-lateral striatum. *Neuropharmacology*, 76(Pt B), 407–415. <https://doi.org/10.1016/j.neuropharm.2013.05.033>.
- Mantzari, E., Vogt, F., Shemilt, I., Wei, Y., Higgins, J. P. T., & Marteau, T. M. (2015). Personal financial incentives for changing habitual health-related behaviors: A systematic review and meta-analysis. *Preventive Medicine*, 75, 75–85. <https://doi.org/10.1016/j.ypmed.2015.03.001>.
- McRae, A. L., Saladin, M. E., Brady, K. T., Upadhyaya, H., Back, S. E., & Timmerman, M. A. (2006). Stress reactivity: Biological and subjective responses to the cold pressor and Trier Social stressors. *Human Psychopharmacology: Clinical and Experimental*, 21(6), 377–385. <https://doi.org/10.1002/hup.778>.
- Miller, R., Plessow, F., Kirschbaum, C., & Stalder, T. (2013). Classification criteria for distinguishing cortisol responders from nonresponders to psychosocial stress: Evaluation of salivary cortisol pulse detection in panel designs. *Psychosomatic Medicine*, 75(9), 832–840. <https://doi.org/10.1097/PSY.000000000000002>.
- Moccia, L., Quintigliano, M., Janiri, D., Martin, V. D., Rogier, G., Sani, G., ... Nicola, M. D. (2021). Heart rate variability and interoceptive accuracy predict impaired decision-making in Gambling Disorder. *Journal of Behavioral Addictions*, 10(3), 701–710. <https://doi.org/10.1556/2006.2021.00067>.
- Mohr, H., Zwosta, K., Markovic, D., Bitzer, S., Wolfensteller, U., & Ruge, H. (2018). Deterministic response strategies in a trial-and-error learning task. *PLOS Computational Biology*, 14(11), e1006621. <https://doi.org/10.1371/journal.pcbi.1006621>.
- Mollick, J. A., & Kober, H. (2020). Computational models of drug use and addiction: A review. *Journal of Abnormal Psychology*, 129(6), 544–555. <https://doi.org/10.1037/abn0000503>.
- Montero-López, E., Santos-Ruiz, A., García-Ríos, M. C., Rodríguez-Blázquez, M., Rogers, H. L., & Peralta-Ramírez, M. I. (2018). The relationship between the menstrual cycle and cortisol secretion: Daily and stress-invoked cortisol patterns. *International Journal of Psychophysiology*, 131, 67–72. <https://doi.org/10.1016/j.ijpsycho.2018.03.021>.
- Murphy, J., Catmur, C., & Bird, G. (2018). Alexithymia is associated with a multidomain, multidimensional failure of interoception: Evidence from novel tests. *Journal of Experimental Psychology. General*, 147(3), 398–408. <https://doi.org/10.1037/xge0000366>.
- Noël, X., Brevers, D., & Bechara, A. (2013). A neurocognitive approach to understanding the neurobiology of addiction. *Current Opinion in Neurobiology*, 23(4), 632–638. <https://doi.org/10.1016/j.conb.2013.01.018>.
- Noël, X., Saeremans, M., Kornreich, C., Bechara, A., Jaafari, N., & Fantini-Hauwel, C. (2018). On the processes underlying the



- relationship between alexithymia and gambling severity. *Journal of Gambling Studies*, 34(4), 1049–1066. <https://doi.org/10.1007/s10899-017-9715-1>.
- Noël, X., Saeremans, M., Kornreich, C., Jaafari, N., & D'Argembeau, A. (2017). Future-oriented mental time travel in individuals with disordered gambling. *Consciousness and Cognition*, 49, 227–236. <https://doi.org/10.1016/j.concog.2017.02.006>.
- Oakes, J., Pols, R., & Lawn, S. (2019). The 'merry-go-round' of habitual relapse: A qualitative study of relapse in electronic gaming machine problem gambling. *International Journal of Environmental Research and Public Health*, 16(16), 2858. <https://doi.org/10.3390/ijerph16162858>.
- Otto, A. R., Gershman, S. J., Markman, A. B., & Daw, N. D. (2013). The curse of planning: Dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychological Science*, 24(5), 751–761. <https://doi.org/10.1177/0956797612463080>.
- Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences*, 110(52), 20941–20946. <https://doi.org/10.1073/pnas.1312011110>.
- Ozgoer, T., Ucar, C., & Yildiz, S. (2017). Cortisol awakening response is blunted and pain perception is increased during menses in cyclic women. *Psychoneuroendocrinology*, 77, 158–164. <https://doi.org/10.1016/j.psyneuen.2016.12.011>.
- Park, H., Lee, D., & Chey, J. (2017). Stress enhances model-free reinforcement learning only after negative outcome. *Plos One*, 12(7), e0180588. <https://doi.org/10.1371/journal.pone.0180588>.
- Patzelt, E. H., Kool, W., Millner, A. J., & Gershman, S. J. (2019). Incentives boost model-based control across a range of severity on several psychiatric constructs. *Biological Psychiatry*, 85(5), 425–433. <https://doi.org/10.1016/j.biopsych.2018.06.018>.
- Petzold, A., Plessow, F., Goschke, T., & Kirschbaum, C. (2010). Stress reduces use of negative feedback in a feedback-based learning task. *Behavioral Neuroscience*, 124, 248–255. <https://doi.org/10.1037/a0018930>.
- Radenbach, C., Reiter, A. M. F., Engert, V., Sjoerds, Z., Villringer, A., Heinze, H.-J., ... Schlagenhaut, F. (2015). The interaction of acute and chronic stress impairs model-based behavioral control. *Psychoneuroendocrinology*, 53, 268–280. <https://doi.org/10.1016/j.psyneuen.2014.12.017>.
- Rummery, G. A., & Niranjana, M. (1994). On-line Q-learning using connectionist systems (Vol. 37). Cambridge, UK: University of Cambridge, Department of Engineering.
- Russell, A. M. T., Browne, M., Hing, N., Visintin, T., Begg, S., Rawat, V., & Rockloff, M. (2021). Stressful life events precede gambling problems, and continued gambling problems exacerbate stressful life events; A life course calendar study. *Journal of Gambling Studies*. <https://doi.org/10.1007/s10899-021-10090-7>.
- Schwabe, L., Haddad, L., & Schachinger, H. (2008). HPA axis activation by a socially evaluated cold-pressor test. *Psychoneuroendocrinology*, 33(6), 890–895. <https://doi.org/10.1016/j.psyneuen.2008.03.001>.
- Schwabe, L., & Wolf, O. T. (2009). Stress prompts habit behavior in humans. *Journal of Neuroscience*, 29(22), 7191–7198. <https://doi.org/10.1523/JNEUROSCI.0979-09.2009>.
- Schwabe, L., & Wolf, O. T. (2010). Socially evaluated cold pressor stress after instrumental learning favors habits over goal-directed action. *Psychoneuroendocrinology*, 35(7), 977–986. <https://doi.org/10.1016/j.psyneuen.2009.12.010>.
- Schwabe, L., & Wolf, O. T. (2011). Stress-induced modulation of instrumental behavior: From goal-directed to habitual control of action. *Behavioural Brain Research*, 219(2), 321–328. <https://doi.org/10.1016/j.bbr.2010.12.038>.
- Sebold, M., Deserno, L., Nebe, S., Schach, D. J., Garbusow, M., Hägele, C., ... Huys, Q. J. M. (2014). Model-based and model-free decisions in alcohol dependence. *Neuropsychobiology*, 70(2), 122–131. <https://doi.org/10.1159/000362840>.
- Sebold, M., Nebe, S., Garbusow, M., Guggenmos, M., Schach, D. J., Beck, A., ... Heinz, A. (2017). When habits are dangerous: Alcohol expectancies and habitual decision making predict relapse in alcohol dependence. *Biological Psychiatry*, 82(11), 847–856. <https://doi.org/10.1016/j.biopsych.2017.04.019>.
- Seehagen, S., Schneider, S., Rudolph, J., Ernst, S., & Zmyj, N. (2015). Stress impairs cognitive flexibility in infants. *Proceedings of the National Academy of Sciences*, 112(41), 12882–12886. <https://doi.org/10.1073/pnas.1508345112>.
- Smeets, T., van Ruitenbeek, P., Hartogsveld, B., & Quaedflieg, C. W. E. M. (2019). Stress-induced reliance on habitual behavior is moderated by cortisol reactivity. *Brain and Cognition*, 133, 60–71. <https://doi.org/10.1016/j.bandc.2018.05.005>.
- Smittenaar, P., FitzGerald, T. H. B., Romei, V., Wright, N. D., & Dolan, R. J. (2013). Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. *Neuron*, 80(4), 914–919. <https://doi.org/10.1016/j.neuron.2013.08.009>.
- Snider, S. E., LaConte, S. M., & Bickel, W. K. (2016). Episodic future thinking: Expansion of the temporal window in individuals with alcohol dependence. *Alcoholism: Clinical and Experimental Research*, 40(7), 1558–1566. <https://doi.org/10.1111/acer.13112>.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1). Cambridge: MIT Press.
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, 37(3), 498–505. <https://doi.org/10.3758/BF03192720>.
- Voon, V., Derbyshire, K., Rück, C., Irvine, M. A., Worbe, Y., Enander, J., ... Bullmore, E. T. (2015). Disorders of compulsivity: A common bias towards learning habits. *Molecular Psychiatry*, 20(3), 345–352. <https://doi.org/10.1038/mp.2014.44>.
- Wang, X., Wu, Q., Egan, L., Gu, X., Liu, P., Gu, H., ... Fan, J. (2019). Anterior insular cortex plays a critical role in interoceptive attention. *eLife*, 8, e42265. <https://doi.org/10.7554/eLife.42265>.
- Weinstock, J., Whelan, J., & Meyers, A. (2008). College students' gambling behavior: When does it become harmful? *Journal of American College Health: J of ACH*, 56, 513–521. <https://doi.org/10.3200/JACH.56.5.513-522>.
- Weissengruber, S., Lee, S. W., O'Doherty, J. P., & Ruff, C. C. (2019). Neurostimulation reveals context-dependent arbitration between model-based and model-free reinforcement learning. *Cerebral Cortex*, 29(11), 4850–4862. <https://doi.org/10.1093/cercor/bhz019>.



- Wirz, L., Bogdanov, M., & Schwabe, L. (2018). Habits under stress: Mechanistic insights across different types of learning. *Current Opinion in Behavioral Sciences*, 20, 9–16. <https://doi.org/10.1016/j.cobeha.2017.08.009>.
- Wyckmans, F., Otto, A. R., Sebold, M., Daw, N., Bechara, A., Saeremans, M., ... Noël, X. (2019). Reduced model-based decision-making in gambling disorder. *Scientific Reports*, 9(1). <https://doi.org/10.1038/s41598-019-56161-z>.

