Decision-making (in)flexibility in gambling disorder

Ana Perandrés-Gómez[1], Juan F. Navas[2, 3], Tim van Timmeren[4,5], & José C Perales[1]

1. Department of Experimental Psychology; Mind, Brain and Research Center (CIMCYC); Universidad de Granada, Spain
2. Department of Personality, Assessment and Clinical Psychology, Universidad Complutense de Madrid, Spain
3. Universitat Oberta de Catalunya, Spain
4. Department of Clinical Psychology, University of Amsterdam, Amsterdam, The Netherlands
5. Amsterdam Brain and Cognition, University of Amsterdam, Amsterdam, The Netherlands

Corresponding author:

José C. Perales
Department of Experimental Psychology
University of Granada
Campus de Cartuja s/n
18071 Granada, Spain
Spain
jcesar@ugr.es

Abstract

Background: Behavioral flexibility –the ability to dynamically readjust our behavior in response to reward contingency changes– is often investigated using probabilistic reversal learning tasks (PRLT). Poor PRLT performance has been proposed as a proxy for compulsivity, and theorized to be related to perseverative gambling. Previous attempts to measure inflexibility with the PRLT in patients with gambling disorder have, however, used a variety of indices that may conflate inflexibility with more general aspects of performance in the task. *Methods*: Trial-by-trial PRLT acquisition and reacquisition curves in 84 treatment-seeking patients with gambling disorder and 64 controls (non-gamblers and non-problem recreational gamblers) were analyzed to distinguish between (a) variability in acquisition learning, and (b) reacquisition learning in reversed contingency phases. Complementarily, stay/switch responses throughout the task were analyzed to identify (c) premature switching, and (d) sensitivity to accumulated negative feedback. *Results and interpretation*: Even after controlling for differences in acquisition learning, patients were slower to readjust their behavior in reversed contingency phases, and were more prone to maintain their decisions despite accumulated negative feedback. Inflexibility in patients with gambling disorder is thus a robust phenomenon that could predate gambling escalation, or result from massive exposure to gambling activities.

Keywords: Gambling disorder, Reversal learning, Reward-based learning, Instrumental learning, Decision-making under ambiguity, Compulsivity

1. Decision-making inflexibility in addictive processes

Behavioral flexibility is defined as the ability to adjust decisions to a dynamically changing environment (Bond, Kamil, & Balda, 2007; Cools, Clark, Owen, & Robbins, 2002). More precisely, flexibility implies detecting changes in reward and punishment contingencies and changing choice preferences to track such changes in an adaptive manner (Izquierdo & Jentsch, 2012). Alterations in this ability have been shown to be a transdiagnostic marker of psychopathology, including addictive disorders (Verdejo-García et al., 2015; Yücel et al., 2019).

Behavioral flexibility has been studied in human and non-human animals using decision-making-under-ambiguity tasks. In these tasks, the probabilities of reward and punishment for each of the response options are initially unknown to the learner, and preferences are gradually established by feedback (e.g. Bechara, Damasio, Tranel, & Damasio, 1997). Flexibility is required as long as reward contingencies vary throughout the task (Brand, Recknor, Grabenhorst, & Bechara, 2007).

The most commonly used protocol to assess behavioral flexibility is the *affective probabilistic reversal learning task* (*PRLT*, Swainson, 2000). In an initial acquisition phase, the learner is presented with a series of trials. On each trial, one of two response options (e.g. clicking on one of two buttons, or touching one of two objects on the screen) is rewarded with a high probability and punished with a low one, whereas the other available option is punished with a high probability and rewarded with a low one. After preference for the high-reward option is established, contingencies are reversed, so that the advantageous option becomes disadvantageous and *vice versa*. This reversed-contingency phase continues until the most-frequently rewarded response becomes predominant and stable again, and the task can consist of several phases with alternating contingencies. In most versions of the PRLT, these contingency changes are abrupt, so the task consists of phases in which reward contingencies are in the acquisition direction, and phases in which the sign of the contingency is reversed, relative to the acquisition phase. Thus, flexibility in this paradigm is defined as the capability to explore away from the initially established responses when these become disadvantageous.

The relevance of reversal learning in relation to addiction stems from the hypothesis that inflexibility can accelerate the loss of control over addictive behaviors. According to a first family of theories (e.g. Berridge & Robinson, 2016; Everitt & Robbins, 2016; Lüscher, Robbins, & Everitt, 2020) certain activities become addictive as they transition from being motivated by the hedonic utility of their consequences to being triggered by conditioned cues (i.e. they become compulsive). In alternative theories (e.g. Lamb & Ginsburg, 2018), addictive behaviors remain goal-driven, yet aberrant learning processes make certain rewards acquire a disproportionate motivational value. Regardless of what the mechanism driving the addictive process is,

inflexibility could render some individuals more vulnerable to addiction (Izquierdo & Jenstch, 2012). Alternatively, the progression of the addictive disorder can interfere with behavioral flexibility (Lucantonio, Caprioli, & Schoenbaum, 2014). For one reason or another, people suffering from different addictive disorders seem to show poorer performance, relative to healthy controls, in PRLT protocols (Camchong et al., 2011; de Ruiter et al., 2009).

2. The role of reversal learning abnormalities in gambling disorder

Gambling disorder presents behavioral and neurobiological similarities with substance use disorders (Fauth-Bühler, Mann, & Potenza, 2017; Petry, Blanco, Jin, & Grant, 2014; Potenza, 2006; Shaffer et al., 2004), which has led to its widespread recognition as a behavioral addiction. Still, the specific mechanisms by means of which gambling, in the absence of an external chemical agent, can become addictive remains debated (Perales et al., 2020).

A number of researchers have proposed that behavioral inflexibility, and specifically abnormalities in reversal learning, could be a marker of gambling disorder vulnerability (Fernández-Serrano, Pérez-García, Perales, & Verdejo-García, 2010; Leeman & Potenza, 2012; Verdejo-García & Manning, 2015). More precisely, not being able to fully reverse preferences, once established, will subsequently facilitate persevering in gambling despite its long-term negative utility. Indeed, patients with gambling disorder (henceforth, PGD) frequently relate their excessive gambling to their inability to overcome the allure of big wins in their early gambling history (Turner, Zangeneh, & Littman-Sharp, 2006). The importance of early reward density has also been corroborated with quantitative analyses of gaming machine data (Haw, 2008), and could be related to the well-known fact that the 'remembered high' plays an important role in motivating compulsive gambling (Volkow, Wise, & Baler, 2017; see also Lister, Nower, & Wohl, 2016).

Some studies seem to support this relationship between reversal learning abnormalities and perseverative gambling. For instance, Boog, et al., (2014) found PGD to display inflexibility in a reversal learning task, but not in the Wisconsin Sorting Card Test (WSCT; a task requiring non-reward based cognitive flexibility). Inflexibility also correlated with psychological distress. With different degrees of evidential strength, a similar argument has been made in a number of previous studies. For instance, de Ruiter et al. (2009) found signs of perseveration in PGD, but not in smokers. PGD's poorer performance in the PRLT occurred in parallel with a lower responsiveness to both losses and wins in the right ventrolateral prefrontal cortex. Navas et al. (2015), Perales et al. (2019), and Torres et al. (2013), in turn, have reported different indices of reversal learning inflexibility to be associated with gambling symptoms severity in both clinical and subclinical ranges (although the latter did not find any case-control differences in inflexibility-related indices).

Converging evidence also arises from related paradigms, such as the Iowa Gambling Task (IGT), in which performance has been hypothesized to partially depend on more basic reversal learning processes (see Fellows & Farah, 2005; Haines, Vassileva, & Ahn, 2018). For instance, Nigro, Ciccarelli, and Cosenza (2018) showed that performance in the IGT was significantly poorer in gamblers with an increased propensity to chase losses. Relatedly, Linnet, Møller, Peterson, Gjedde, and Doudet (2010) found increased levels of dopamine release in the ventral striatum in response to losses during the IGT, restricted to PGDs who lost money in the task (allegedly, loss chasers).

3. Present study: dissociating inflexibility from other behavioral components of PRLT in gambling disorder

The evidence briefly reviewed above arises from studies in which inflexibility has been measured in a very variable manner (see van Timmeren, Daams, van Holst, & Goudriaan, 2018, for a similar argument). Most studies use the number of perseverative errors –the count of consecutive choices of the previously (but no longer) advantageous option after a reversal– as the main dependent variable (Clarke, Dalley, Crofts, Robbins, & Robert, 2004; Clarke, Walker, Dalley, Robbins, & Roberts, 2007; D'Cruz et al., 2013; Ersche et al., 2011). This measure is, however, influenced by how well responses were established *before* reversal (namely, acquisition differences). This problem can be palliated by setting an acquisition criterion prior to reversal. Still, the fact that two participants have reached the same criterion does not ensure that their acquisition asymptotes are comparable. Moreover, this measure partially conflates perseveration with decision stability. In probabilistic tasks, maintaining your response after one or two (or even more) consecutive negative feedback instances can be an appropriate way to avoid premature switching after random punishment. In principle, there is no easy way to determine whether a given decision to maintain the same decision as in the previous trial (e.g. after two consecutive punishments) reflects noise-proof stability or perseveration. This problem calls for the necessity to model choice dynamics as negative feedback accrues in a detailed manner.

Other studies have used global performance indices (number of correct choices) in reversal phases versus phases with the original contingency as the main dependent variable (de Ruiter et al., 2009; Torres et al., 2013). This variable can be interpreted as a measure of the difficulty to adapt preferences to the new circumstances, or *reversal cost*. However, this is not free of interpretation problems either. The most important one stems from the fact that learners differ in their asymptotic learning levels. A proportion of people, for instance, never reach the economically optimal strategy of selecting the high reward option in every trial (which would maximize the amount of reward accrued). More often, learners behave as 'probability matchers'

(Da Silva, Victorino, Caticha, & Baldo, 2017), in such a way that the probability with which they select the high-reward option equates at the asymptote with the probability of receiving a reward for choosing that option. Learners with higher acquisition asymptotes will find it more difficult to change their preferences after reversal (so they will temporarily make more 'incorrect' choices), but this apparent disadvantage should not be defined as learning inflexibility.

In spite of their interpretational problems, these two measures are representative of the two ways in which inflexibility can be operationalized: as perseveration (tendency to maintain decisions in spite of mounting negative/corrective feedback) and as reversal cost. In the present work we present novel ways of analyzing both reversal cost and perseveration, while trying to overcome these limitations, and controlling for potential confounders, with the largest sample of PGD and controls used in a study of this type so far (*N* = 148).

Our version of the PRLT consists of 40 acquisition trials (*acquisition phase*), followed by 40 trials with reversed contingency (*first reversal phase*), 40 trials with a reward contingency of the same sign as the original one (*reacquisition phase*), and 40 trials with a reversed contingency (*second reversal phase*). In a first analytic approach (*reversal cost* modelling), each response will be coded as correct (clicking on the advantageous option) or incorrect (clicking on the disadvantageous one). Correctness or incorrectness of each response will be predicted on the basis of *group* (PGD vs low or non-gambling controls), *task phase* (1-4, namely acquisition, reversal, reacquisition, second reversal), and the *trial* within each phase (1-40). Our main aim is to dissociate inflexibility, conceptualized as slower learning functions in phases 2 and 4 (reversed) versus 1 and 3 (non-reversed), from potentially contaminating factors that are by definition present in raw decisions, including general acquisition differences. Once inflexibility is isolated, we will evaluate its sensitivity to group.

In the second approach (*perseveration* modelling) responses will be coded as 'same as previous' (*stay*), or 'different to previous' (*switch*). This response will be predicted on the basis of group and *insensitivity to cumulative negative feedback* (which is defined as the number of consecutive *stay* responses preceding the present trial, in the presence of negative feedback). On the one hand, switching after positive feedback or a single punishment would indicate that responses are more instable or stochastic; on the other hand, staying despite several consecutive punishments would be a sign of perseveration. As in the first approach, the effect of sensitivity to cumulative negative feedback will be analyzed as a function of group, but, in contrast to raw counts of perseverative errors, this approach allows for dissociating the effects of perseveration in the face of error from (in)stability of decisions.

The main aim of the present study is thus to establish the existence of increased reversal learning inflexibility in patients with gambling disorder. In order to do so, we describe a method to dissociate inflexibility from general differences in acquisition learning[1], as well as from decision-making (in)stability. If this dissociation can be corroborated, computational models will need to account for it, in such a way that the behavioral pattern typically observed in patients can be reproduced by tuning the model parameters in specific ways. Computational approaches to these learning mechanisms will be briefly reviewed in the discussion section, in light of the observed results.

4. Methods

4.1. Participants

Eighty-four patients with gambling disorder (PGD) and 64 low or non-gambling controls took part in this study. Groups were similar in gender proportion, age and number of education years (Table 1). Inclusion criteria for the PGD group were: (1) having a diagnosis of gambling disorder with DSM5 criteria, as established by their therapist, and (2) abstaining from gambling for at least 15 days. At the moment of participation, severity of gambling symptoms was further assessed using the SOGS questionnaire. One participant in the PGD group scored below the problem gambling SOGS threshold at the moment of doing the PRLT task, but was not excluded from any analyses. Inclusion criteria for controls were identical, except for GD diagnosis and being in treatment. Additionally, all participants in the control group presented SOGS severity scores below the problem gambling threshold, and gambled less than once a week. The two participants with highest SOGS values in the control group scored 3 (see Table 1 for information on the distribution of SOGS scores in both groups; scores for all participants are included in the open data file accompanying this article).

Exclusion criteria for both groups were (1) lifetime diagnosis or treatment for any psychopathology (apart from gambling disorder in the case of PGD), and (2) any history of neurological disease or brain trauma causing unconsciousness for 10 minutes or longer (for the two criteria, as informed by the participant in semi-structured interviews).

In the case of PGD, criteria were initially checked by their therapist by means of a semi-structured interview carried out upon admission to treatment. This interview belongs of the

[1] This is by no means the first attempt to properly dissociate inflexibility from more general alterations of performance in reversal learning tasks and, indeed, the implications of these other aspects have been also scrutinized. As noted by a reviewer, Walton, Behrens, Buckley, Rudebeck, and Rushworth (2010) used a 3-arm bandit task with macaques to discriminate between altered reversal learning and an excess of "inappropriate switches" in response to feedback. In human studies, depressed patients have been found to be hyper-responsive to negative feedback, which misleads them to inappropriate rule switching, while acquisition and reversal learning were preserved (e.g. Murphy, Michael, Robbins, & Sahakian, 2003).

treatment facilities' shared protocol, and its aim is to examine symptoms of psychological and personality disorders. Therapists received instructions from our research team to filter participants based on the abovementioned criteria, so we do not have any records of potential participants who were excluded. Subsequently, and right before the assessment for the present study, all participants completed a brief semi-structured interview, exploring (1) whether they had ever been diagnosed of psychopathological disorders or serious physical illnesses, (2) the need to have received psychological or psychiatric treatment throughout life, (3) traumatic accidents with loss of consciousness, (4) whether they thought they had ever lost control over their use of alcohol or other substances. This brief interview was carried out personally by the second author, or under his supervision.

4.2. Procedure

PGD were recruited from different association of rehabilitated disordered gamblers (*AGRAJER*, *APLIJER*, and *ALUJER*), all based in Andalucía, Spain. Controls were recruited by advertisements posted in the University of Granada premises and social networks, and among researchers' and patients' acquaintances. All participants were informed about the aims of the research and they gave their informed consent before taking part in the study. The assessment took place in a single session with a duration of approximately two hours. The protocol consisted of different questionnaires and neuropsychological computer tasks administered in a quasi-randomized order. The present sample was recruited across three different phases of a research project with slightly different aims and assessment protocols (GBrain, see details in the funding section). Overlap between the current sample and the ones in other studies by our team was 33.78% (Megías et al., 2018), 58.11% (Navas et al., 2016; Navas et al., 2017, study 1), and 44.59% (Perales, Navas, Ruiz de Lara, Maldonado, & Catena, 2017; each of these articles describes the tasks that were relevant of their respective aims[2]). Participants were compensated with €10/hour. In the case of patients, this payment was made through a relative. The present study was approved by the Ethic Committee of the University of Granada, according to the Helsinki Declaration.

4.3. Measures and tasks

The affective Probabilistic Reversal Learning Task (*PRLT*, Moreno-López et al., 2015). Different versions of the PRLT differ in the number of reversals and the specific probabilities programmed (in the extreme case, contingencies can be deterministic; e.g. Ghahremani, Monterosso, Jentsch,

[2] Phase 1 of the GBrain Project also included tests for numeracy, impulsivity, sensitivity to reward and punishment, gambling-related cognitions, and associative learning. Phase 2 included tests in these domains plus depression, emotion regulation, risk taking, and nicotine dependence. Phase 3 did not include numeracy, but assessed gambling motives. Phases 2 and 3 also included a separate session (in a different day) with psychophysiological or neuroimaging recordings.

Bilder, & Poldrack, 2009). In probabilistic reinforcement schedules, contingencies are slower to acquire, and elicit slower preference adjustment rates after reversal; namely, they elicit more perseverative behavior (e.g. Cools et al., 2002). Additionally, probabilistic designs are more realistic and representative of decision-making in daily life (Amitai et al., 2014).

The current version is a computerized task divided into four phases, each of which consists of 40 trials. In each trial, participants are presented with two differently colored squares. The two squares were equal in size (approximately 5 x 5 cm), were simultaneously displayed at the center of each half of the screen, and could be yellow, green, red, and blue. Color blindness was not explicitly tested, but participants selected the two colors to be used themselves, and none of them informed of difficulties to discriminate between the two colors. The position of the colors (left/right) varied randomly from trial to trial (so that position was uninformative).

In each trial, participants were asked to choose a color by mouse-clicking on it. The aim of the task is to accumulate points. In each phase, one of the response options is pre-programmed as correct and probabilistically rewarded with virtual points. When the incorrect option is selected, points are probabilistically subtracted from the account. In Phase 1, the correct color is rewarded in 80% of the trials, and punished in 20%, whereas the incorrect one is rewarded in 20% of the trials, and punished in 80%. In Phase 2, the correct and incorrect colors are simply reversed. Phases 3 and 4 are identical to phases 1 and 2, respectively, with the only difference that the 80/20 contingency is degraded to 70/30 to avoid perfect performance and ceiling effects.

The South Oaks Gambling Screen (SOGS; Spanish version, Echeburúa, Báez, Fernández-Montalvo, & Paéz, 1994) was used to assess gambling severity. This is the most widely used tool to assess GD severity in Spain, and has good psychometric properties. Cronbach's α for the SOGS scale in the present study was 0.93.

The MultiCAGE CAD-4 (Pedrero-Pérez et al., 2007) is a quick screening questionnaire used to assess the risk of suffering from clinically significant levels of alcohol misuse, illegal drug misuse, excessive gambling, and other potentially problematic behaviors such as hyper-sexuality, excessive Internet surfing, excessive buying, dysregulated eating, and excessive video-gaming. The alcohol misuse-related items in MultiCAGE are identical to the CAGE screening tool (Ewing, 1984), with the items for the other problematic behaviors being strongly inspired by the alcohol-related ones. Cronbach's α for the MultiCAGE scale in the present study was 0.75 for the full scale (0.63 for alcohol, and 0.69 for illegal drugs subscales).

Intelligence was measured through the vocabulary and abstract reasoning subtests of the *Wechsler Adult Intelligence Scale (WAIS-IV*; Wechsler, 2012). A general IQ score was estimated from these tests following the standard instructions.

5. Statistical analysis and results

5.1. Sample characterization

First, PGD and control groups were compared on gender, age, intelligence, education years, and risk of substance abuse (alcohol and illegal drugs) using Bayes Factors (BF) with JASP software (JASP Team, 2019; default settings). Please note that small BFs (e.g. BF<1/3) are customarily interpreted as favoring the null hypothesis (the absence of differences), large BF (e.g. BF>3) are customarily interpreted as supporting the alternative hypothesis, and BFs around 1 (e.g. 1/3<BF<3) are considered uninformative, or just anecdotally supporting either the alternative or the null hypothesis. In the present study, variables in which the BF favored the alternative hypothesis (of a difference between groups), whatever the strength of this support is (i.e. BF>1) were included as fixed-effects variables in further analyses. Groups' descriptives and BFs are reported in Table 1.

5.2. Reversal cost modelling

This analysis aimed to investigate group differences on trial-by-trial accuracy within and across phases (observed proportion of correct responses across phases and trials are reported in the Appendix). Model fitting was carried out using the *glmer* function in the *lme4* R package (Bates, Mächler, Bolker, & Walker, 2015).

The output variable was trial-by-trial accuracy (modelled as a binomial variable, correct/incorrect, using a logit link function). This is predicted by a best fitting model (resulting from a hierarchical procedure) including a mixture of the following fixed-effects factors: Phase (1-4), Trial within Phase (1-40), Group (PGD vs. controls), and IQ (previously identified as a potential confounder)[3]. As responses were collected within-participant across phases and trials, Participant was included as a random intercept, and Trial as a random slope. To use Trial as a quantitative variable and to model learning as a linear function, Trial was translated into a natural logarithmic scale (which will henceforth be referred to simply as Trial). The PRLT file corresponding to one participant from the PGD group was corrupted and not considered for any further analyses[4].

[3] All results reported here are qualitatively identical without IQ as a covariate. For the sake of brevity and readability, those analyses are not reported here.

[4] Three alternative modellings of the effect of trial were considered when building the baseline model: a simple linear effect of trial, the linear effect of log-transformed trial (the one finally selected), and a polynomial decomposition (linear + quadratic components). The logarithmic model yielded the best fitting

The effect of Phase was decomposed into an orthogonal set of contrasts that modelled different patterns of change in accuracy rates across phases. The main contrast of interest compared Phases 1 and 3 (acquisition-sign contingencies) versus Phases 2 and 4 (reversed contingencies), i.e. contrast C1 (1, -1, 1, -1). Contrast C2 (1, 1, -1, -1) modelled the effect of contingency degradation in the second half of the task, i.e. how much performance differed between the easy (80/20%) and difficult phases (70/30%) of the task. Contrast C3 (1, -1, -1, 1) compares phases 2 and 3 against 1 and 4 (which is theoretically irrelevant, but necessary for the three contrasts together to conform an exhaustive orthogonal set).

A first (baseline) model was built with Phase, Trial, IQ scores, and all possible interactions between these variables as fixed-effects factors. IQ was zero-centered and scaled to facilitate model convergence[5]. A saturated model was subsequently built by adding Group, Group x Phase, Group x Trial, and Group x Phase x Trial upon the baseline model. Finally, the saturated model was tested against a simpler one with the three-way interaction removed (Model 1). Removing the three-way interaction resulted in a substantial loss of model fit (AIC = 29465, and AIC = 29467, for the model with and without the interaction, respectively; $\chi^2 = 8.063$, $p = 0.045$). In other words, the Group x Phase x Trial interaction significantly contributed to account for variance in participants' choices. (Details of the three models, and how they fare against each other are reported in Table A1 in the Appendix, Supplementary Materials). Table 2 shows the estimates and significance levels for all effects in the saturated model (for the sake of simplicity, direct and interactive effects of IQ are not presented).

As shown in Figure 1 and Table 2, Group was involved in several effects. Globally, the proportion of correct responses was lower among PGD than among controls. Group also interacted with trial, in such a way that, for controls relative to PGD, learning functions were steeper, and reached higher asymptotes in all phases. Controls were also more clearly subject to transitory reversal costs in the first trials after reversal points. However, from a theoretical point of view, the most relevant effect is the C1 x Trial x Group interaction. This interaction reveals that the slowing of learning in phases 2 and 4 (reversed) relative to 1 and 3 (non-reversed), was more marked in PGD. (Learning curves in reversed contingency phases became flatter and asymptotes lower, in comparison to controls). A comparison of Figure 1 and observed

indices and was selected for further analyses. These comparisons are included in the R code accompanying this article.

[5] Running this model also allows to estimate the potential confounding effect of IQ differences before considering the effect of Group. The interactions of IQ with the C1 component of phase (the one containing information about inflexibility) and with C1 and Trial were both non-significant (p=0.36 for C1 x IQ, and p=0.85 for C1 x IQ x Trial). The interactions of C1 with group in subsequent models are thus not likely to be attributable to IQ differences across groups.

responses (Figure A1, Appendix, Supplementary Materials) shows that predicted responses closely reproduced the pattern of observed ones. The difference between groups across phases and trials in the task, revealing the origin of significant contrasts, is depicted in Figure A2 in the Appendix.

5.3. Reversal cost modelling in efficient learners

Results so far suggest that performance on the PRLT is hampered in PGD. However, the group interaction with phase and trial (suggesting increased inflexibility in PGD) remains somewhat overshadowed by more global differences in task performance.

A complementary analysis was thus run with the subset of participants in the two groups who made 4 consecutive correct choices in the last four trials of Phase 1 (henceforth efficient learners, 36 controls, and 34 PGD). The 4-trial criterion was chosen to maintain the largest possible number of participants in the sample while ensuring the two groups showed no differences in acquisition in the first phase. As depicted in the first panel of Figure 2, this yields virtually identical learning curves in Phase 1 for PGD and controls. The analysis rationale was similar to the analysis used for the whole sample (details of the three models: Baseline, Saturated, and Model 2 are reported in Table A2 in the Appendix, Supplementary Materials). Removing the three-way interaction resulted again in a substantial loss of model fit (AIC = 12509, and AIC = 12516, for the model with and without the interaction, respectively; $\chi^2 = 12.872$, p = 0.005). In other words, the Group x Phase x Trial interaction significantly contributed to account for variance in observed responses, even more clearly than in the analysis with the whole sample.

This restricted analysis allows a neater comparison of individual effects after contingencies are reversed (Table 3). PGD made less correct choices than controls, but the effect was restricted to Phases 2 and 4, that is, to phases with reversed contingency, as shown by the largely significant C1 x Trial x Group effect (see Figure 2). The matching of learning curves in Phase 1 was externally imposed by our subsample selection criterion. Most importantly, however, the two groups departed from each other in Phase 2, with PGD performing increasingly worse than controls as the phase progressed, converged in Phase 3, and once again diverged in Phase 4.

5.4. Perseveration modelling

Following the same hierarchical strategy as in the previous analysis, we started by building a baseline model with cumulative negative feedback, standardized IQ, and their interaction in the fixed part, and the stay/switch response as dependent variable. Participant was treated as a random intercept, and Cumulative negative feedback also as a random slope. Cumulative negative feedback was expressed on a 0-3 scale, where 0 stands for a positive feedback on the

last trial; 1 stands for a single negative feedback in the last trial; 2 stands for two consecutive negative feedbacks in the last two trials with one *stay* response in the previous trial; and 3 stands for three consecutive negative feedbacks in the last three trials with two consecutive *stay* responses in the previous two trials.

In a second step, the Group factor and the Group x Cumulative negative feedback interaction were included in the model (saturated model). Most importantly, removing the Cumulative negative feedback x Group interaction from the saturated model (Model 3 in Table A3 in the Appendix, Supplementary Materials) hampered model fit (AIC = 27966, and AIC = 27975, for the model with and without the interaction, respectively; $\chi^2 = 11.688$, $p < 0.001$).

Figure 3 shows the shape of the interaction between Cumulative negative feedback and Group (see Table 4 for estimates). In the two groups, as expected, participants were more likely to switch with the accumulation of negative feedback. However, the different slopes for the two groups suggest a mixture of premature switching (instability) and perseveration in PGD. This is reflected in an elevated proneness to switch with no or little negative feedback, and a slightly elevated proneness to stay when accumulation of negative feedback would make switching a more optimal option.

5.5. Perseveration in efficient learners

We applied the same analysis strategy to the subset of efficient learners. Results from model fitting were very similar to the ones from the full sample, with the model with the interaction outperforming the one without the interaction (AIC = 11886, and AIC = 11889, for the model with and without the interaction, respectively; $\chi^2 = 4.639$, $p = 0.031$; see Table A4 in the Appendix, Supplementary Materials for more details).

Table 5 and Figure 4 display results from the analysis. The slope for the effect of Cumulative negative feedback was again steeper for controls. However, in this subsample, PGD differed less from controls in response instability (both were almost equally unlikely to switch after a positive feedback), but differed more strongly from controls in their tendency to perseveration (staying after 2-3 negative feedbacks for the same choice).

6. Discussion

Results can be summarized as follows. Patients with gambling disorder (PGD) showed worse performance on the PRLT, relative to controls. PGD's behavior adjusted more slowly to the initial contingencies, and they reached a lower number of correct responses by the end of the first phase of the task. When contingency was reversed, so that the correct option turned incorrect and *vice versa*, controls' performance seemed to (initially) deteriorate more, but this difference was restricted to the first trials of the second phase, and further learning seemed to proceed

more slowly for PGD. The pattern observed in phases 1 and 2 was qualitatively reproduced in phases 3 and 4.

In contrast with previous studies in gambling disorder, however, we were able to dissociate general acquisition learning and inflexibility effects by including specific trial and phase effects (and their interaction) in the model used for analysis. Once the marginal and interactive effects of trial and phase, and the interaction between group and trial, were discounted, a remaining interaction between group, phase and trial showed that learning in reversed contingency phases (2 and 4) was significantly slower in PGD than in controls. This result was reinforced by a supplementary analysis restricted to participants showing high learning levels in phase 1. Among these, PGD kept on showing a substantial decrease in re-acquisition learning in reversed contingency phases (2 and 4), relative to controls, i.e. learning inflexibility.

Although their results are consistent, the joint interpretation of these two analyses should be made with some caution. On the one hand, running analyses only with "efficient" learners is likely to have removed not only "less-efficient" learners, but also probability matchers: participants whose decisions match programmed reward probabilities for each option, and should not be regarded as worse than the ones of maximizers. Nevertheless, this subgrouping made the detection of inflexibility easier. Equating acquisition curves resulted in a similar matching of the curves for phase 3 (when contingency returns to its initial sign; see Figure 3), but differences in reversed phases remained. On the other hand, the fact that the same pattern of inflexibility is statistically detectable in the whole sample, in which the signal/noise ratio for that effect is obviously lower, reinforces the capacity of our approach to operationally dissociate inflexibility from other behavioral components in the PRLT.

A congruent pattern emerged from the analysis of stay/switch responses. In global analyses, PGD seemed to behave more stochastically than controls, with a stronger tendency to change their decisions after being rewarded or punished just once, and also to maintain the same decision after 2 or 3 consecutive punishments. As in the previous analysis, even when less efficient learners were removed, PGD remained more perseverative than controls. In other words, regardless of whether inflexibility was measured as slowed learning in reversed contingency phases or insensitivity to cumulative negative feedback, PGD showed stronger signs of inflexibility than controls.

Isolating inflexibility from other aspects of reward learning is necessary to establish an operational link between PRLT performance and different aspects gambling disorder and other addictive processes (see, for example, Figee et al., 2016). Problem gamblers' behavior can be undeniably considered inflexible in gambling contexts, as they keep on participating in gambling

activities despite their negative utility. Additionally, compulsivity, understood as a type of behavior that is not driven by the utility of its consequences, has been proposed as a transdiagnostic dimension of addictive disorders in the Research Domain Criteria (RDoC) framework, and has been hypothesized to play a central role in their etiology (Brooks, Lochner, Shoptaw, & Stein, 2017; Ersche, Roiser, Robbins, & Sahakian, 2008; Yücel et al., 2019). Some cross-sectional studies have indeed shown that individuals suffering from substance use disorders behave more inflexibly than controls in domain-general reward learning tasks (Banca, Harrison, & Voon, 2016; Lee, Hoppenbrouwers, & Franken, 2019). However, it remains unclear whether this type of inflexibility is due to neuroadaptations caused by chronic drug exposure, or predate the escalation of drug use, making individuals more vulnerable to develop an addictive disorder.

In the specific area of gambling disorder, research on behavioral inflexibility is scarcer. In the only systematic review and meta-analysis published to date, van Timmeren et al., (2018) identified four studies using the PRLT in PGD samples. The ones by Boog et al. (2014), and de Ruiter et al. (2009) reported significant differences between groups, whereas the one by Torres et al. (2013), and Verdejo-García et al. (2015) did not (although in the former an index of inflexibility was found to be linked to severity of gambling disorder). In this review, considerable heterogeneity between studies was found, which may reflect differences in the way PRLT performance was assessed. As noted in the introduction, these different scoring methods could well also conflate inflexibility and acquisition deficits.

The distinction between inflexibility and acquisition deficits is relevant beyond methodological discussion. Although this is a matter of intense theoretical debate, there is some agreement that acquisition learning relies on simpler mechanisms than behavioral adjustment to unsignaled contingency changes. Extinction, for instance, is not just unlearning of previously learnt associations, but also requires a type of context-dependent learning about the absence of the reinforcer (Bouton, 2019). In a similar vein, reversal learning has been hypothesized to require higher-order mechanisms to restructure the set of learned associations (for an updated view, see Izquierdo, Brigman, Radke, Rudebeck, & Holmes 2017). Gambling disorder is likely to involve anomalies of associative learning. Yet, the understanding of these anomalies has so far been hindered by the difficulty to establish the connection between gambling and inflexibility, unconfounded with other aspects of performance on the PRLT. This corroboration should warrant further research on the potential association between inflexibility (as detected in laboratory tasks), and inflexibility-like manifestations of gambling disorder, as well as on the identification of inflexibility as a potential therapeutic target (in line with a recent proposal by Goudriaan, 2020).

Importantly, we also found that, on average, learning was slower in PGD than in controls even in phases with non-reversed contingency. In a recent theoretical paper (Perales et al., 2020), we posit that more basic and general neurocognitive alterations (e.g. inattention, reduced control resources, hyper- or hyposensitivity to reward and punishment) play ancillary roles in the etiology of behavioral addictions, but are neither necessary nor sufficient for the addictive process to develop. Whatever the disorder of interest is, case-control studies tend to find between-group differences in these processes. However, given their lack of specificity, such differences provide little information about the etiological mechanisms that are specific to each disorder. Reversal learning, like virtually all neuropsychological tasks, is multicomponential (Yaple & Yu, 2019), and different components are likely to be differentially involved in different disorders, or aspects of the same disorder. Our proposal is that inflexibility, once isolated, is a good candidate as a feature specific of disordered gambling (see also Wiehler, Chakroun, & Peters, 2019).

Our (data-driven) approach is complementary to (theory-driven) computational modeling approaches. Computational models can describe individual choice behavior as theoretically meaningful parameters based on potential underlying processes, such as learning rates and choice randomness. As noted in the introduction, behavioral findings constrain computational models in such a way that such behavioral patterns can be reproduced by tuning the model parameters in specific ways. If a model is able to capture (in)flexibility as described here, and its parameters can be tuned to account for differences between PGD and non-patients, those parameters can be used to confer cognitive content to such differences. Once a model is found to be a viable account of a family of behavioral findings, it can describe individual choice behavior as theoretically meaningful parameters based on potential underlying processes, and thus capture choice tendencies independently of task-specifics. (For an example of how that approach can be applied to the relationship between IGT performance and gambling, see Kildahl, Hansen, Brevers, & Skewes, 2020, in this special volume).

Although the precise computational modelling of inflexibility is beyond the aims of the present work, our results seem compatible with current models. For example, in the experience-weighted attraction (EWA) model (Camerer & Ho, 1999; den Ouden et al., 2013), the experience weight parameter (ρ) captures the intuition that updating becomes slower with the accumulation of experience. Thus, it becomes more and more difficult to adjust behavior during the acquisition phase, which leads to increased perseveration at the moment of reversal. Our finding that PGD show increased perseveration is in line with an increased experience decay factor in the EWA model, as opposed to a more general problem in reward/punishment learning (i.e. learning rate, α). Interestingly, increased perseveration (as captured by experience weight)

has been linked specifically to genetic variations in dopamine transporter genes (den Ouden et al, 2013). Distorted dopamine function has previously been implicated in gambling disorder (e.g. van Holst et al., 2018) and our results seem to point in a similar direction, albeit indirectly. Alternatively, our results may be explained by differences in more complex (sequential) exploratory behavior as captured by another recent model (VSE; Ligneul, 2019). Future work may directly compare these models to see which captures participants' choice behavior better.

7. Final remarks

The present study presents a number of strengths. First, it presents two methods to operationalize domain-general inflexibility, and clearly dissociates it both from general acquisition learning, and (in)stability of decisions. Second, the sample size is by far the largest used to date in studies with PRLT in PGD, which allowed to perform restricted analyses with efficient learners. Finally, the fact that patients suffer from a behavioral addiction, and current substance abuse was controlled to some degree, allows us to lessen the possibility that differences are attributable to exposure to neurotoxic substances.

Still, the present study is also affected by several limitations. First, using a sample partially coincident with other studies requires some precaution regarding generalization, as the different studies describe different phenomena in non-independent samples. Second, it suffers from the same weaknesses attributable to all case-control studies, in relation to the establishment of directional causal links. The observed differences can be attributed either to (1) chronic gambling-driven neuroadaptations, i.e. transference of compulsivity from gambling to gambling-unrelated tasks, as seen in animal and human studies on substance-use disorders (Lucantonio, et al., 2014), or (2) individual differences that make some people more prone to compulsivity than others, and thus more vulnerable to developing a gambling habit (Groman, Massi, Mathias, Lee, & Taylor, 2019). Third, our sample consisted almost exclusively of males (which limits the generalization to the general population). And fourth, although patients were not clinically depressed at the moment of assessment, symptoms of mood disorder were not systematically controlled. Mood alterations are of special relevance in this context, as some studies have found a link between depression and PRLT performance (see, for example, Murphy et al., 2003). As noted earlier, however, depression seems to be associated with hyper-sensitivity to negative feedback and premature switching, but not with increased perseveration or inflexibility. The effect of depression is thus not expected to explain the observed case-control inflexibility effects away.

These results can have important clinical implications. First, they strongly suggest that inflexibility is not restricted to gambling scenarios (at least in some PGD), but that more general acquisition anomalies are more frequent among them than among controls. Second, these two

types of learning problems are not necessarily linked, as inflexible behavior is clearly identifiable in the subgroup of gamblers who reached a level of performance comparable to the one of controls during initial acquisition. Third, although that possibility has not been explored here, differences in inflexibility within the PGD population are presumably linked to individual differences in the manifestations of gambling disorder. For instance, inflexibility as described in the present work could arise from a tendency to prematurely form an internal model of the task contingencies that becomes resistant to later reinforcement history. A similar difficulty to utilize structure when it is available in the reinforcement history of games has been observed to be associated with gambling-related cognitive distortions (Lim, Jocham, Hunt, Behrens, & Rogers, 2015). And fourth, although, on average, performance in several aspects of the task was hampered in the group of patients, there is also a substantial overlap between groups (there are PGD showing neither slowed acquisition nor inflexibility, and controls with substandard performance). Individuals suffering from an addictive disorder who show specific reward learning deficits, but not marked compulsivity, may benefit of the inclusion of contingent management techniques involving tangible, swift reinforcers contingent to well-defined behaviours (McDonell et al., 2013), within broad cognitive-behavioural therapy programs. Compulsivity/inflexibility, on the contrary, may require the inclusion of extinction or response prevention techniques (Verdejo-García, Alcázar-Córcoles, & Albein-Urios, 2019).

Funding

This work was supported by grants from the Spanish Government (PSI2017-85488-P: Ministerio de Economía y Competitividad, Secretaría de Estado de Investigación, Desarrollo e Innovación; Convocatoria 2017 de Proyectos I+D de Excelencia, Spain; co-funded by the Fondo Europeo de Desarrollo Regional, FEDER, European Union; and PSI2013-45055: Ministerio de Economía y Competitividad, Secretaría de Estado de Invetigación, Desarrollo e Innovación; Convocatoria 2013 de Proyectos I+ D de Excelencia). Additionally, JFN was supported by a grant from the Spanish Government (PSI2017-85159-P. Ministerio de Ciencia, Innovación y Universidades). Funding agencies had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript, or the decision to submit the paper for publication.

Competing interests

The authors have declared that no competing interests exist.

Authors contributions

JCP and JFN designed the study. JCP performed the statistical analysis. JCP and AP wrote the original draft. All authors contributed to the interpretation of findings and to write and edit further versions of the manuscript. All authors have approved the final version of this manuscript.

Data Availability

The open database and code files for these analyses are available at the Open Science Framework repository: https://osf.io/9tnwr/

References

Amitai, N., Young, J. W., Higa, K., Sharp, R. F., Geyer, M. A., & Powell, S. B. (2014). Isolation rearing effects on probabilistic learning and cognitive flexibility in rats. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(1), 388-406. https://doi.org/10.3758/s13415-013-0204-4.

Banca, P., Harrison, N. A., & Voon, V. (2016). Compulsivity across the pathological misuse of drug and non-drug rewards. *Frontiers in Behavioral Neuroscience*, *10*, 154. https://doi.org/10.3389/fnbeh.2016.00154.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). "Fitting linear mixed-effects models using lme4." *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01.

Bechara, A., Damasio, H., Tranel, D., & Damasio A.R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275* (5304), 1293-1294. https://doi.org/10.1126/science.275.5304.1293.

Berridge, K.C., Robinson, T.E., (2016). Liking, wanting, and the incentive-sensitization theory of addiction. *American Psychologist, 71*(8), 670-679. https://doi.org/10.1037/amp0000059

Bond, A. B., Kamil, A. C., & Balda, R. P. (2007). Serial reversal learning and the evolution of behavioral flexibility in three species of North American corvids (Gymnorhinus cyanocephalus, Nucifraga columbiana, Aphelocoma californica). *Journal of Comparative Psychology*, *121*(4), 372-379. https://doi.org/10.1037/0735-7036.121.4.372.

Boog, M., Höppener, P., Goudriaan, A. E., Wetering, B. J. M., Boog, M. C., & Franken, I. H. (2014). Cognitive inflexibility in gamblers is primarily present in reward-related decision making. *Frontiers in Human Neuroscience*, *8*, 569. https://doi.org/10.3389/fnhum.2014.00569.

Bouton, M. E. (2019). Extinction of instrumental (operant) learning: Interference, varieties of context, and mechanisms of contextual control. *Psychopharmacology, 236*(1), 7-19. https://doi.org/10.1007/s00213-018-5076-4

Brand, M., Recknor, E. C., Grabenhorst, F., & Bechara, A. (2007). Decisions under ambiguity and decisions under risk: correlations with executive functions and comparisons of two different gambling tasks with implicit and explicit rules. *Journal of Clinical and Experimental Neuropsychology*, *29*(1), 86-99. https://doi.org/10.1080/13803390500507196.

Brooks, S. J., Lochner, C., Shoptaw, S., & Stein, D. J. (2017). Using the research domain criteria (RDoC) to conceptualize impulsivity and compulsivity in relation to addiction. *Progress in Brain Research*, 235, 177-218. https://doi.org/10.1016/bs.pbr.2017.08.002.

Camchong, J., MacDonald III, A. W., Nelson, B., Bell, C., Mueller, B. A., Specker, S., & Lim, K. O. (2011). Frontal hyperconnectivity related to discounting and reversal learning in cocaine

subjects. *Biological Psychiatry, 69*(11), 1117-1123. https://doi.org/10.1016/j.biopsych.2011.01.008

Camerer, C., and Ho, T. (1999). Experience-weighted attraction learning in normal form games. *Econometrica, 67*(4)*,* 827-874. https://doi.org/10.1111/1468-0262.00054

Clarke, H. F., Dalley, J. W., Crofts, H. S., Robbins, T. W., & Roberts, A. C. (2004). Cognitive inflexibility after prefrontal serotonin depletion. *Science*, *304*(5672), 878-880. https://doi.org/10.1007/s10899-004-4578-7.

Clarke, H. F., Walker, S. C., Dalley, J. W., Robbins, T. W., & Roberts, A. C. (2007). Cognitive inflexibility after prefrontal serotonin depletion is behaviorally and neurochemically specific. *Cerebral Cortex*, *17*(1), 18-27. https://doi.org/10.1093/cercor/bhj120.

Cools, R., Clark, L., Owen, A. M., & Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *Journal of Neuroscience*, *22*(11), 4563-4567. https://doi.org/10.1523/JNEUROSCI.22-11-04563.2002.

Da Silva, C. F., Victorino, C. G., Caticha, N., & Baldo, M. V. C. (2017). Exploration and recency as the main proximate causes of probability matching: a reinforcement learning analysis. *Scientific Reports, 7*(1), 15326. https://doi.org/10.1038/s41598-017-15587-z.

D'Cruz, A.M., Ragozzino, M. E., Mosconi, M. W., Shrestha, S., Cook, E. H., & Sweeney, J. A. (2013). Reduced behavioral flexibility in autism spectrum disorders. *Neuropsychology, 27*(2), 152-160. https://doi.org/10.1037/a0031721

den Ouden, H. E., Daw, N. D., Fernandez, G., Elshout, J. A., Rijpkema, M., Hoogman, M., ... & Cools, R. (2013). Dissociable effects of dopamine and serotonin on reversal learning. *Neuron*, *80*(4), 1090-1100. https://doi.org/10.1016/j.neuron.2013.08.030

de Ruiter, M.B., Veltman, D.J., Goudriaan, A.E., Oosterlaan, J., Sjoerds, Z. & van den Brink, W. (2009). Response perseveration and ventral prefrontal sensitivity to reward and punishment in male problem gamblers and smokers. *Neuropsychopharmacology, 34*, 1027-1038. https://doi.org/10.1038/npp.2008.175.

Echeburúa, C., Báez, E., Fernández-Montalvo, J., & Paéz, D. (1994). Cuestionario de juego patológico de South Oaks (SOGS): Validación española [The South Oaks Gambling Screen (SOGS): Spanish validation]. *Análisis y Modificación de Conducta, 20,* 769-791.

Ersche, K. D., Roiser, J. P., Abbott, S., Craig, K. J., Müller, U., Suckling, J., ... & Bullmore, E. T. (2011). Response perseveration in stimulant dependence is associated with striatal dysfunction and can be ameliorated by a D2/3 receptor agonist. *Biological Psychiatry, 70*(8), 754-762. https://doi.org/10.1016/j.biopsych.2011.06.033

Ersche, K. D., Roiser, J. P., Robbins, T. W., & Sahakian, B. J. (2008). Chronic cocaine but not chronic amphetamine use is associated with perseverative responding in humans. *Psychopharmacology*, 197, 421-431. https://doi.org/10.1007/s00213-007-1051-1.

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.

Ewing, J. A. (1984). Detecting alcoholism: The CAGE questionnaire. *JAMA, 252*(14), 1905-1907. https://doi.org/jama.1984.03350140051025

Fauth-Bühler, M., Mann, K., & Potenza, M. N. (2017). Pathological gambling: A review of the neurobiological evidence relevant for its classification as an addictive disorder. *Addiction Biology*, *22*(4), 885-897. https://doi.org/10.1111/adb.12378.

Fellows, L. K., & Farah, M. J. (2005). Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cerebral Cortex, 15*(1), 58-63. https://doi.org/10.1093/cercor/bhh108.

Fernández-Serrano, M. J., Pérez-García, M., Perales, J. C., & Verdejo-García, A. (2010). Prevalence of executive dysfunction in cocaine, heroin and alcohol users enrolled in therapeutic communities. *European Journal of Pharmacology*, *626*(1), 104-112. https://doi.org/10.1016/j.ejphar.2009.10.019.

Figee, M., Pattij, T., Willuhn, I., Luigjes, J., van den Brink, W., Goudriaan, A., Potenza, M.N., Robbins, T.W., & Denys, D. (2016). Compulsivity in obsessive–compulsive disorder and addictions. *European Neuropsychopharmacology, 26*(5), 856-868. https://doi.org/10.1016/j.euroneuro.2015.12.003.

Ghahremani, D. G., Monterosso, J., Jentsch, J. D., Bilder, R. M., & Poldrack, R. A. (2009). Neural components underlying behavioral flexibility in human reversal learning. *Cerebral Cortex*, *20*(8), 1843-1852. https://doi.org/10.1093/cercor/bhp247.

Goudriaan, A. E. (2020). Integrating neurocognition from bench to bedside in gambling disorder: from neurocognitive to translational studies. *Current Opinion in Behavioral Sciences, 31*, 83-88. https://doi.org/10.1016/j.cobeha.2019.12.012

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/j.biopsych.2018.12.017.

Haw, J. (2008). Random-ratio schedules of reinforcement: The role of early wins and unreinforced trials. *Journal of Gambling Issues*, (21), 56-67. https://doi.org/10.4309/jgi.2008.21.6

Haines, N., Vassileva, J., & Ahn, W. Y. (2018). The Outcome-Representation Learning Model: A novel reinforcement learning model of the Iowa Gambling Task. *Cognitive Science, 42*(8), 2534-2561. https://doi.org/10.1111/cogs.12688

Izquierdo, A., & Jentsch, J. D. (2012). Reversal learning as a measure of impulsive and compulsive behavior in addictions. *Psychopharmacology,219*(2), 607-620. https://doi.org/10.1007/s00213-011-2579-7.

Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017). The neural basis of reversal learning: An updated perspective. *Neuroscience, 345*, 12-26. https://doi.org/10.1016/j.neuroscience.2016.03.021.

JASP Team (2019). JASP (Version 0.11.1) [Computer software].

Kildahl, N., Hansen, S., Brevers, D., & Skewes, J. (2020). Individual differences in learning during decision-making may predict specific harms associated with gambling. *Addictive Behaviors, 110,* 106496. https://doi.org/10.1016/j.addbeh.2020.106496

Lamb, R. J., & Ginsburg, B. C. (2018). Addiction as a BAD, a behavioral allocation disorder. *Pharmacology Biochemistry and Behavior, 164,* 62-70. https://doi.org/10.1016/j.pbb.2017.05.002

Lee, R. S., Hoppenbrouwers, S., & Franken, I. (2019). A systematic meta-review of impulsivity and compulsivity in addictive behaviors. *Neuropsychology Review, 29*(1), 14-26. https://doi.org/10.1007/s11065-019-09402-x.

Leeman, R. F., & Potenza, M. N. (2012). Similarities and differences between pathological gambling and substance use disorders: A focus on impulsivity and compulsivity. *Psychopharmacology, 219(*2), 469-490. https://doi.org/10.1007/s00213-011-2550-7.

Lim, M. S., Jocham, G., Hunt, L. T., Behrens, T. E., & Rogers, R. D. (2015). Impulsivity and predictive control are associated with suboptimal action-selection and action-value learning in regular gamblers. *International Gambling Studies, 15*(3), 489-505. https://doi.org/10.1080/14459795.2015.1078835

Linnet, J., Møller, A., Peterson, E., Gjedde, A., & Doudet, D. (2010). Dopamine release in ventral striatum of pathological gamblers losing money. *Acta Psychiatrica Scandinavica, 122*(4), 326-333. https://doi.org/10.1111/j.1600-0447.2010.01591.x.

Ligneul, R. (2019). Sequential exploration in the Iowa gambling task: Validation of a new computational model in a large dataset of young and old healthy participants. *PLoS Computational Biology, 15*(6), e1006989. https://doi.org/10.1371/journal.pcbi.1006989.

Lister, J. J., Nower, L., & Wohl, M. J. (2016). Gambling goals predict chasing behavior during slot machine play. *Addictive Behaviors, 62*, 129-134. https://doi.org/10.1016/j.addbeh.2016.06.018

Lucantonio, F., Caprioli, D., & Schoenbaum, G. (2014). Transition from 'model-based' to 'model-free' behavioral control in addiction: involvement of the orbitofrontal cortex and dorsolateral striatum. *Neuropharmacology,* 76, 407-415. https://doi.org/10.1016/j.neuropharm.2013.05.033.

Lüscher, C., Robbins, T. W., & Everitt, B. J. (2020). The transition to compulsion in addiction. *Nature Reviews Neuroscience, 21,* 247-263. https://doi.org/10.1038/s41583-020-0289- z

McDonell, M. G., Srebnik, D., Angelo, F., McPherson, S., Lowe, J. M., Sugar, A., ... & Ries, R. K. (2013). Randomized controlled trial of contingency management for stimulant use in community mental health patients with serious mental illness. *American Journal of Psychiatry*, *170*(1), 94-101. https://doi.org/10.1176/appi.ajp.2012.11121831.

Megías, A., Navas, J. F., Perandrés-Gómez, A., Maldonado, A., Catena, A., & Perales, J. C. (2018). Electroencephalographic evidence of abnormal anticipatory uncertainty processing in gambling disorder patients. *Journal of Gambling Studies, 34*(2), 321-338. https://doi.org/10.1007/s10899-017-9693-3.

Moreno-López, L., Perales, J. C., Son, D. V., Albein-Urios, N., Soriano-Mas, C., Martinez-Gonzalez, J. M., … & Verdejo-García, A. (2015). Cocaine use severity and cerebellar gray matter are associated with reversal learning deficits in cocaine-dependent individuals. *Addiction Biology, 20*(3), 546-556. https://doi.org/10.1111/adb.12143.

Murphy, F. C., Michael, A., Robbins, T. W., & Sahakian, B. J. (2003). Neuropsychological impairment in patients with major depressive disorder: The effects of feedback on task performance. *Psychological Medicine, 33*(3), 455-467. https://doi.org/10.1017/S0033291702007018

Navas, J. F., Contreras-Rodríguez, O., Verdejo-Román, J., Perandrés-Gómez, A., Albein-Urios, N., Verdejo-García, A., & Perales, J. C. (2017). Trait and neurobiological underpinnings of negative emotion regulation in gambling disorder. *Addiction*, *112*(6), 1086-1094. https://doi.org/10.1111/add.13751.

Navas, J. F., Torres, A., Vilar, R., Verdejo-García, A., Catena, A., & Perales, J. C. (2015). Nonmonetary decision-making indices discriminate between different behavioral components of gambling. *Journal of Gambling Studies, 31*, 1545-1560. https://doi.org/10.1007/s10899-014-9482-1.

Navas, J. F., Verdejo-Garcia, A., Lopez-Gomez, M., Maldonado, A., & Perales, J. C. (2016). Gambling with rose-tinted glasses on: Use of emotion-regulation strategies correlates with dysfunctional cognitions in gambling disorder patients. *Journal of Behavioral Addictions, 5*(2), 271-281. https://doi.org/10.1556/2006.5.2016.040

Nigro, G., Ciccarelli, M., & Cosenza, M. (2018). The illusion of handy wins: Problem gambling, chasing, and affective decision-making. *Journal of Affective Disorders, 225*, 256-259. https://doi.org/10.1016/j.jad.2017.08.010

Pedrero-Pérez, E. J., Rodríguez-Monje, M. T., Gallardo-Alonso, F., Fernández-Girón, M., Pérez-López, M., & Chicharro-Romero, J. (2007). Validación de un instrumento para la detección de trastornos de control de impulsos y adicciones: el MULTICAGE CAD-4 [Validation of a tool for screening of impulse control disorders and addiction: MULTICAGE CAD-4]. *Trastornos Adictivos, 9,* 269-278.

Perales, J. C., King, D. L., Navas, J. F., Schimmenti, A., Sescousse, G., Starcevic, V., ... & Billieux, J. (2020). Learning to lose control: A process-based account of behavioral addiction. *Neuroscience & Biobehavioral Reviews, 108,* 771-780. https://doi.org/10.1016/j.neubiorev.2019.12.025

Perales, J. C., Navas, J. F., Ruiz de Lara, C. M., Maldonado, A., & Catena, A. (2017). Causal learning in gambling disorder: Beyond the illusion of control. *Journal of Gambling Studies, 33*(2), 705-717. https://doi.org/10.1007/s10899-016-9634-6

Petry, N. M., Blanco, C., Jin, C., & Grant, B. F. (2014). Concordance between gambling disorder diagnoses in the DSM-IV and DSM-5: Results from the national epidemiological survey of alcohol and related disorders. *Psychology of Addictive Behaviors, 28*(2), 586-591. https://doi.org/10.1037/a0034661

Potenza, M. N. (2006). Should addictive disorders include non-substance-related conditions? *Addiction, 101*(s1), 142-151. https://doi.org/10.1111/j.1360-0443.2006.01591.x.

Shaffer, H. J., Laplante, D. A., Labrie, R. A., Kidman, R. C., Donato, A. N., & Stanton, M. V. (2004). Toward a syndrome model of addiction: Multiple expressions, common etiology. *Harvard Review of Psychiatry, 12*(6), 367-374. https://doi.org/.1080/10673220490905705

Swainson, R., Rogers, R. D., Sahakian, B. J., Summers, B. A., & Polkey, C. E. (2000). Probabilistic learning and reversal deficits in patients with Parkinson's disease or frontal or temporal lobe lesions: Possible adverse effects of dopaminergic medication. *Neuropsychologia, 38*(5),596-612. https://doi.org/10.1016/S0028-3932(99)00103-7

Torres, A., Catena, A., Cándido, A., Maldonado, A., Megías, A., & Perales, J. C. (2013). Cocaine dependent individuals and gamblers present different associative learning anomalies in feedback-driven decision making: A behavioral and ERP study. *Frontiers in Psychology, 4*,122. https://doi.org/10.3389/fpsyg.2013.00122.

Turner, N. E., Zangeneh, M., & Littman-Sharp, N. (2006). The experience of gambling and its role in problem gambling. *International Gambling Studies, 6*(2), 237-266. https://doi.org/10.1080/14459790600928793

van Holst, R. J., Sescousse, G., Janssen, L. K., Janssen, M., Berry, A. S., Jagust, W. J., & Cools, R. (2018). Increased striatal dopamine synthesis capacity in gambling addiction. *Biological Psychiatry*, *83*(12), 1036-1043. https://doi.org/10.1016/j.biopsych.2017.06.010

van Timmeren, T., Daams, J. G., Van Holst, R. J., & Goudriaan, A. E. (2018). Compulsivity-related neurocognitive performance deficits in gambling disorder: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews, 84*, 204-217. https://doi.org/10.1016/j.neubiorev.2017.11.022.

Verdejo-García, A., Alcázar-Córcoles, M. A., & Albein-Urios, N. (2019). Neuropsychological interventions for decision-making in addiction: A systematic review. *Neuropsychology Review, 29*(1), 79-92. https://doi.org/10.1007/s11065-018-9384-6.

Verdejo-García, A., Clark, L., Verdejo-Roman, J., Albein-Urios, N., Martinez-Gonzalez, J. M., Gutierrez, B., & Soriano-Mas, C. (2015). Neural substrates of cognitive flexibility in cocaine and gambling addictions. *The British Journal of Psychiatry*, *207*(2), 158-164. https://doi.org/10.1192/bjp.bp.114.152223.

Verdejo-Garcia, A., & Manning, V. (2015). Executive functioning in gambling disorder: Cognitive profiles and associations with clinical outcomes. *Current Addiction Reports, 2*(3), 214-219.

Volkow, N. D., Wise, R. A., & Baler, R. (2017). The dopamine motive system: Implications for drug and food addiction. *Nature Reviews Neuroscience*, *18*(12), 741. https://doi.org/10.1038/nrn.2017.130.

Walton, M. E., Behrens, T. E., Buckley, M. J., Rudebeck, P. H., & Rushworth, M. F. (2010). Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron, 65*(6), 927-939. https://doi.org/10.1016/j.neuron.2010.02.027

Wechsler, D. (2012). WAIS-IV. *Escala de inteligencia de Wechsler para adultos-IV. Manual técnico y de interpretación*. Madrid: NCS Pearson, Inc. Edición original, 2008.

Wiehler, A., Chakroun, K., & Peters, J. (2019). Attenuated directed exploration during reinforcement learning in gambling disorder. *BioRxiv,* 823583. Unpublished results. https://doi.org/10.1101/823583

Yaple, Z. A., & Yu, R. (2019). Fractionating adaptive learning: A meta-analysis of the reversal learning paradigm. *Neuroscience & Biobehavioral Reviews, 102,* 85-94. https://doi.org/10.1016/j.neubiorev.2019.04.006

Yücel, M., Oldenhof, E., Ahmed, S. H., Belin, D., Billieux, J., Bowden-Jones, H., ... & Verdejo-García, A. (2019). A transdiagnostic dimensional approach towards a neuropsychological assessment for addiction: An international Delphi consensus study. *Addiction*, *114*(6), 1095-1109. https://doi.org/10.1111/add.14424

Table 1. Descriptive statistics and matching tests for the two groups in the study (PGD: Patients with gambling disorder).

		All participants			***Efficient learners***		
	Group	***Mean***	***SD***	***BF10***	***Mean***	***SD***	***BF10***
Age	Control	33.78	8.74	0.38 (anecdotal, null)	33.72	8.51	0.48 (anecdotal, null)
	PGD	36.26	11.32		36.59	10.75	
Education years	Control	13.36	3.29	0.36 (anecdotal, null)	13.39	3.27	0.27 (substantial, null)
	PGD	12.63	4.02		13.19	4.77	
IQ	Control	104.18	10.94	86.37 (strong, alternative)	104.30	11.22	1.81 (anecdotal, alternative)
	PGD	97.29	12.40		98.09	12.91	
MC alcohol	Control	1.08	1.15	0.23 (substantial, null)	1.31	1.22	0.57 (anecdotal, null)
	PGD	0.99	1.23		0.97	1.27	
MC drugs	Control	0.59	0.96	0.44 (anecdotal, null)	0.86	1.10	13.17 (strong, alternative)
	PGD	0.44	0.94		0.21	0.48	
SOGS	Control	0.58	0.91	> 100 (very strong, alternative)	0.58	0.84	>100 (very strong, alternative)
	PGD	10.42	3.04		10.03	2.90	
Gender	Control	2F/63M		0.074 (substantial, null)	0.00	0F/36M	0.113 (substantial, null)
	PGD	2F/82M			0.03	1F/34M	

Note. Risk of suffering from a substance use disorder was assessed with the MultiCAGE CAD-4 test (MC, Pedrero-Pérez et al., 2007). Two scales from the MC questionnaire were used for alcohol and illegal drug abuse. Scores represent the number of risk criteria (out of 4) participants report to meet. Bayes factors (BF10) for quantitative variables were computed using Bayesian Mann-Whitney U-tests as implemented with default settings in JASP software. The Bayes factor for reported gender was calculated using the joint multinomial test as implemented in JASP. The label between brackets indicates the degree of support favoring the null (BF<1) or the alternative (BF>1) hypothesis, according to the customary standard.

Abbreviations. IQ: Intelligence quotient. SOGS: South Oaks Gambling Screen (Echeburúa et al., 1994).

Table 2. Full sample: Effect estimates for the best-fitting model of correct choices as a function of Phase, Trial and Group (Intelligence quotient is included in the model, but not shown; for Phase effect contrasts significant $p = 0.05/3 = 0.017$).

Predictors	***Odds Ratios***	***CI***	***p***
Intercept	1.95	1.65 – 2.31	**<0.001**
Phase (C1)	0.72	0.69 – 0.75	**<0.001**
Phase (C2)	0.89	0.85 – 0.93	**<0.001**
Phase (C3)	0.89	0.85 – 0.93	**<0.001**
Trial	1.63	1.47 – 1.81	**<0.001**
Group	0.66	0.53 – 0.83	**<0.001**
Phase (C1) x Trial	1.11	1.06 – 1.16	**<0.001**
Phase (C2) x Trial	0.92	0.88 – 0.96	**<0.001**
Phase (C3) x Trial	1.10	1.05 – 1.15	**<0.001**
Phase (C1) x Group	1.00	0.94 – 1.06	0.954
Phase (C2) x Group	1.03	0.97 – 1.09	0.373
Phase (C3) x Group	0.96	0.91 – 1.02	0.217
Trial x Group	0.76	0.66 – 0.87	**<0.001**
Phase (C1) x Trial x Group	0.93	0.87 – 0.98	**0.010**
Phase (C2) x Trial x Group	1.04	0.98 – 1.10	0.208
Phase (C3) x Trial x Group	1.00	0.94 – 1.06	0.979
Random part			
σ^2	3.29		
$\tau_{\text{Participant}}$	0.41		
$\tau_{\text{Participant x Trial}}$	0.14		
$\rho_{\text{Participant}}$	0.87		
ICC	0.14		

Abbreviations. CI: Confidence interval. ICC: Intraclass correlation coefficient.

Table 3. Efficient learners: Effect estimates for the best-fitting model of correct choices as a function of Phase, Trial and Group (Intelligence quotient is included in the model, but not shown).

Predictors	***Odds Ratios***	***CI***	***p***
Intercept	2.66	2.05 – 3.46	**<0.001**
Phase (C1)	0.69	0.64 – 0.73	**<0.001**
Phase (C2)	0.82	0.77 – 0.87	**<0.001**
Phase (C3)	0.82	0.77 – 0.87	**<0.001**
Trial	2.14	1.79 – 2.55	**<0.001**
Group	0.66	0.45 – 0.95	**0.026**
Phase (C1) x Trial	1.15	1.08 – 1.22	**<0.001**
Phase (C2) x Trial	0.90	0.84 – 0.96	**0.001**
Phase (C3) x Trial	1.15	1.07 – 1.22	**<0.001**
Phase (C1) x Group	0.74	0.68 – 0.81	**<0.001**
Phase (C2) x Group	0.97	0.88 – 1.06	0.458
Phase (C3) x Group	0.87	0.80 – 0.95	**0.003**
Trial x Group	0.73	0.57 – 0.94	**0.013**
Phase (C1) x Trial x Group	0.86	0.79 – 0.94	**0.001**
Phase (C2) x Trial x Group	0.98	0.89 – 1.07	0.648
Phase (C3) x Trial x Group	0.92	0.84 – 1.01	0.073
Random Effects			
σ^2	3.29		
$\tau_{\text{Participant}}$	0.54		
$\tau_{\text{Participant x Trial}}$	0.23		
$\rho_{\text{Participant}}$	0.92		
ICC	0.19		

Abbreviations. CI: Confidence interval. ICC: Intraclass correlation coefficient.

Table 4. Full sample: Effect estimates for the best-fitting model of stay/switch responses as a function of (cumulative) Negative feedback and Group (Intelligence quotient is included in the model, but not shown).

Predictors	***Odds Ratios***	***CI***	***p***
Intercept	0.30	0.24 – 0.38	**<0.001**
Negative feedback	1.76	1.55 – 2.00	**<0.001**
Group	1.60	1.16 – 2.21	**0.004**
Negative feedback x Group	0.75	0.63 – 0.88	**0.001**
Random Effects			
σ^2	3.29		
$\tau_{\text{Participant}}$	0.83		
$\tau_{\text{Participant x Negative feedback}}$	0.18		
$\rho_{\text{Participant}}$	-0.55		
ICC	0.19		

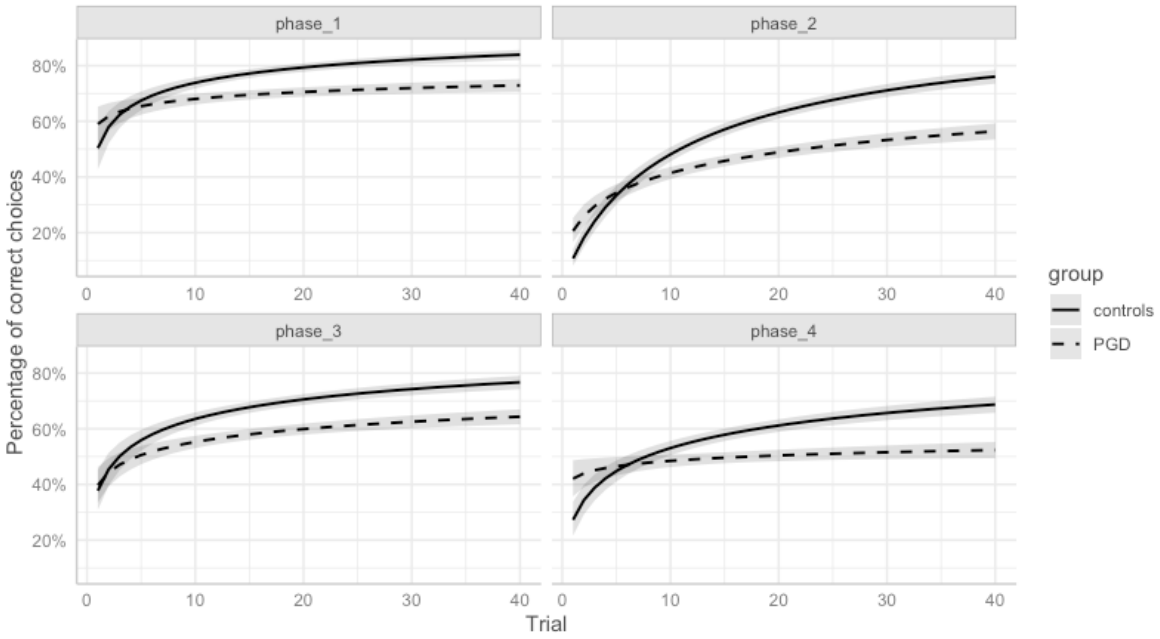
Abbreviations. CI: Confidence interval. ICC: Intraclass correlation coefficient.

Table 5. Efficient learners: Effect estimates for the best-fitting model of stay/switch responses as a function of (cumulative) Negative feedback and Group (IQ is included in the model, but not shown).

Predictors	***Odds Ratios***	***CI***	***p***
Intercept	0.20	0.14 – 0.29	**<0.001**
Negative feedback	1.97	1.66 – 2.33	**<0.001**
Group	1.32	0.78 – 2.25	0.306
Negative feedback x Group	0.77	0.61 – 0.98	**0.031**
Random Effects			
σ^2	3.29		
$\tau_{\text{Participant}}$	1.14		
$\tau_{\text{Participant x Negative feedback}}$	0.17		
$\rho_{\text{Participant}}$	-0.52		
ICC	0.24		

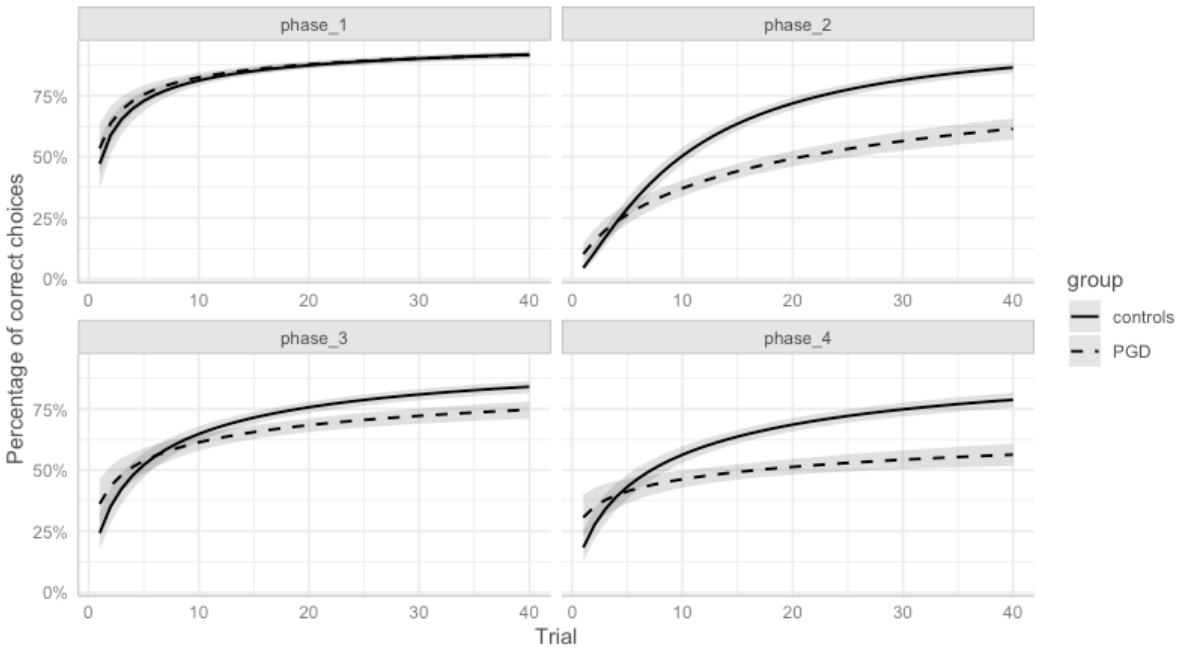
Abbreviations. CI: Confidence interval. ICC: Intraclass correlation coefficient.

Figure 1. Predicted percentages of correct responses from the saturated model for the full sample across Phase, Trial, and Group (the Trial scale has been reconstituted in the 1-40 range to facilitate interpretation).



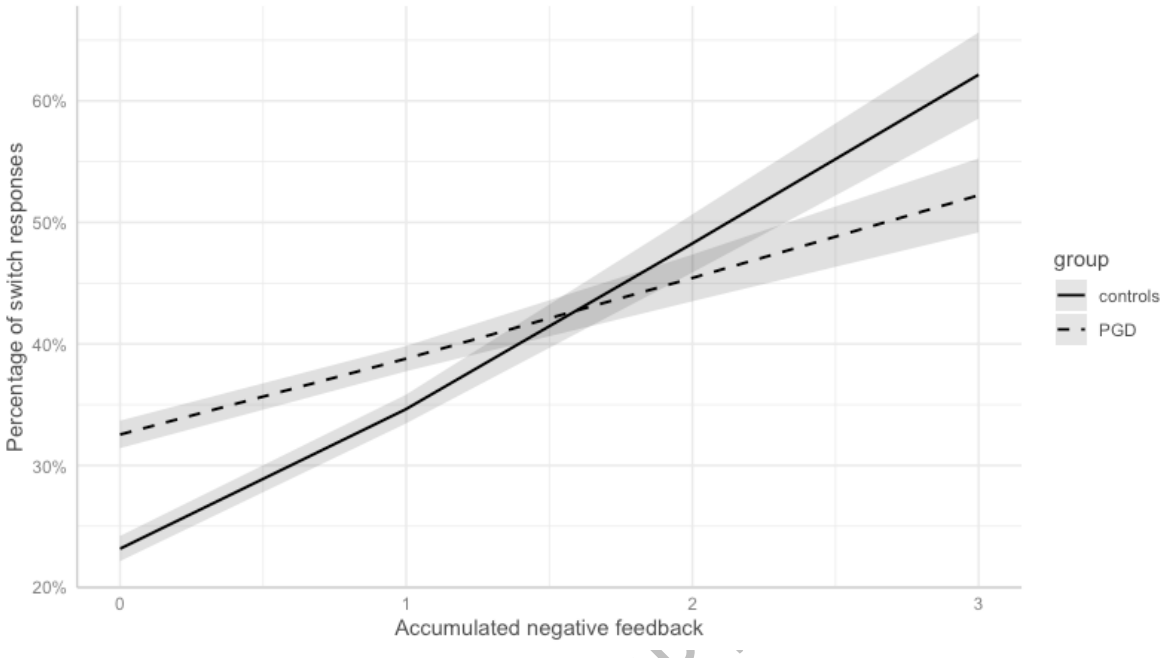
Abbreviations. PGD: Patients with Gambling Disorder.

Figure 2. Predicted percentages of correct responses from the saturated model for efficient learners across Phase, Trial, and Group (the Trial scale has been reconstituted to facilitate interpretation).



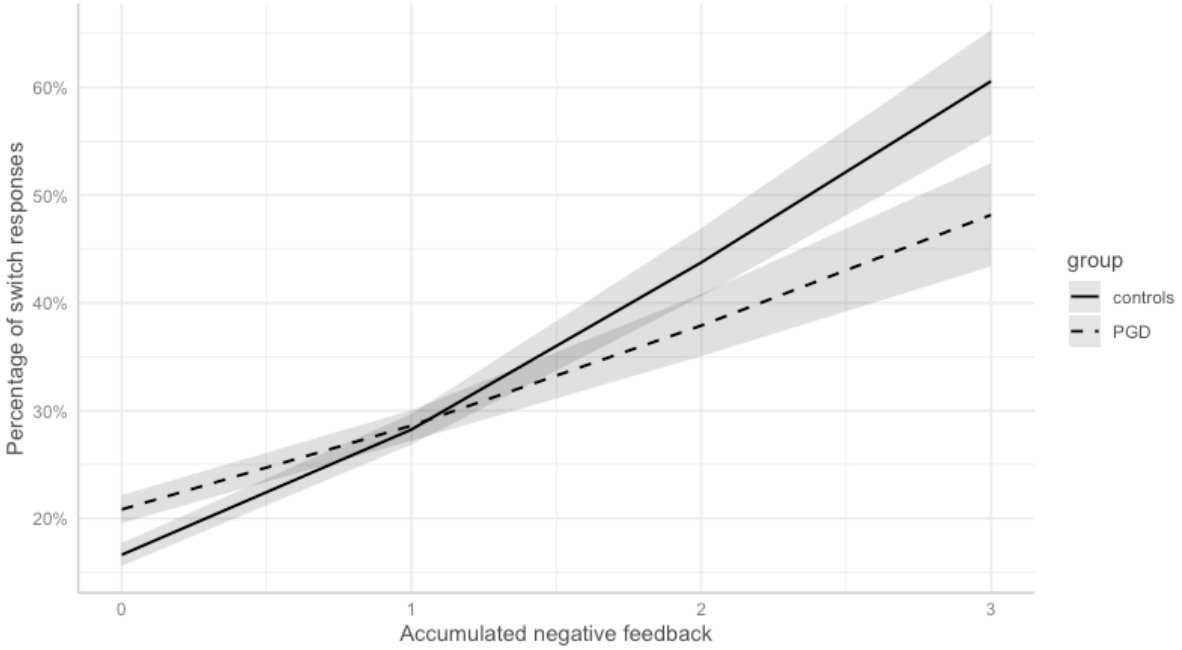
Abbreviations. PGD: Patients with Gambling Disorder.

Figure 3. Predicted percentage of switch responses from the saturated model of switch/stay responses for the full sample across Group and Cumulative negative feedback.



Abbreviations. PGD: Patients with Gambling Disorder.

Figure 4. Predicted percentage of switch responses from the saturated model of switch/stay responses for efficient learners across Group and Cumulative negative feedback.



Abbreviations. PGD: Patients with Gambling Disorder.

Appendix

Table A1. Full sample: Fitting indices for mixed-effects models of correct responses.

Model	***Fixed factors***	***df***	***AIC***	χ^2	***p***
a. Baseline	Phase, Trial, IQ	19	29475		
b. Saturated	Baseline *plus* Group + (Group x Phase) + (Group x Trial) + (Group x Phase x Trial)	27	29465	26.052	0.001 (b>a)
c. Model 1	Saturated *minus* (Group x Phase x Trial)	24	29467	8.063	**0.045** (b>c)

Note. The baseline model also includes all the possible interactions between the three predictors. Participant is included in all models as a random intercept, and Trial as a random slope. All comparisons are made by pitching the more complex model against the simpler one in each pair.

Abbreviations. df: degrees of freedom; AIC: Akaike Information Criterion; IQ: Intelligence Quotient.

Table A2. Efficient learners: Fitting indices for mixed-effects models of correct responses.

Model	***Fixed factors***	***df***	***AIC***	χ^2	***p***
a. Baseline	Phase, Trial, IQ	19	12557		
b. Saturated	Baseline + Group + Group x Phase + Group x Trial + Group x Phase x Trial	27	12509	63.598	<0.001 (b>a)
c. Model 2	Saturated *minus* Group x Phase x Trial	24	12516	12.872	0.005 (b>c)

Note. The baseline model also includes all the possible interactions between the three predictors. Participant is included in all models as a random intercept, and Trial as a random slope. All comparisons are made by pitching the more complex model against the simpler one in each pair.

Abbreviations. df: degrees of freedom; AIC: Akaike Information Criterion; IQ: Intelligence Quotient.

Table A3. Full sample: Fitting indices for mixed-effects models of stay/switch responses.

Model	***Fixed factors***	***df***	***AIC***	χ^2	***p***
a. Baseline	Negative feedback, IQ	7	27975		
b. Saturated	Baseline + Group + (Group x Negative feedback)	9	27966	12.987	**0.002** (b>a)
c. Model 3	Saturated *minus* (Group x Negative feedback)	8	27975	11.688	**<0.001** (b>c)

Note. The baseline model includes the interaction between (cumulative) Negative feedback and IQ. Participant is included in all models as a random intercept, and (cumulative) Negative feedback as a random slope. All comparisons are made by pitching the more complex model against the simpler one in each pair.

Abbreviations. df: degrees of freedom; AIC: Akaike Information Criterion; IQ: Intelligence Quotient.

Table A4. Efficient learners: Fitting indices for mixed-effects models of stay/switch responses.

Model	***Fixed factors***	***df***	***AIC***	χ^2	***p***
a. Baseline	Negative feedback, IQ	7	11887		
b. Saturated	Baseline + Group + (Group x Negative feedback)	9	11886	4.650	0.098
c. Model 4	Baseline *minus* (Group x Negative feedback)	8	11889	4.639	**0.031** (b>c)

The baseline model includes the interaction between Cumulative negative feedback and IQ. Participant is included in all models as a random intercept, and Cumulative negative feedback as a random slope. All comparisons are made by pitching the more complex model against the simpler one in each pair.

Abbreviations. df: degrees of freedom; AIC: Akaike Information Criterion; IQ: Intelligence Quotient.

Figure A1. Proportion of observed correct responses across trials and phases for the Control (CNTRL) and the Patients with Gambling Disorder (PGD) group.

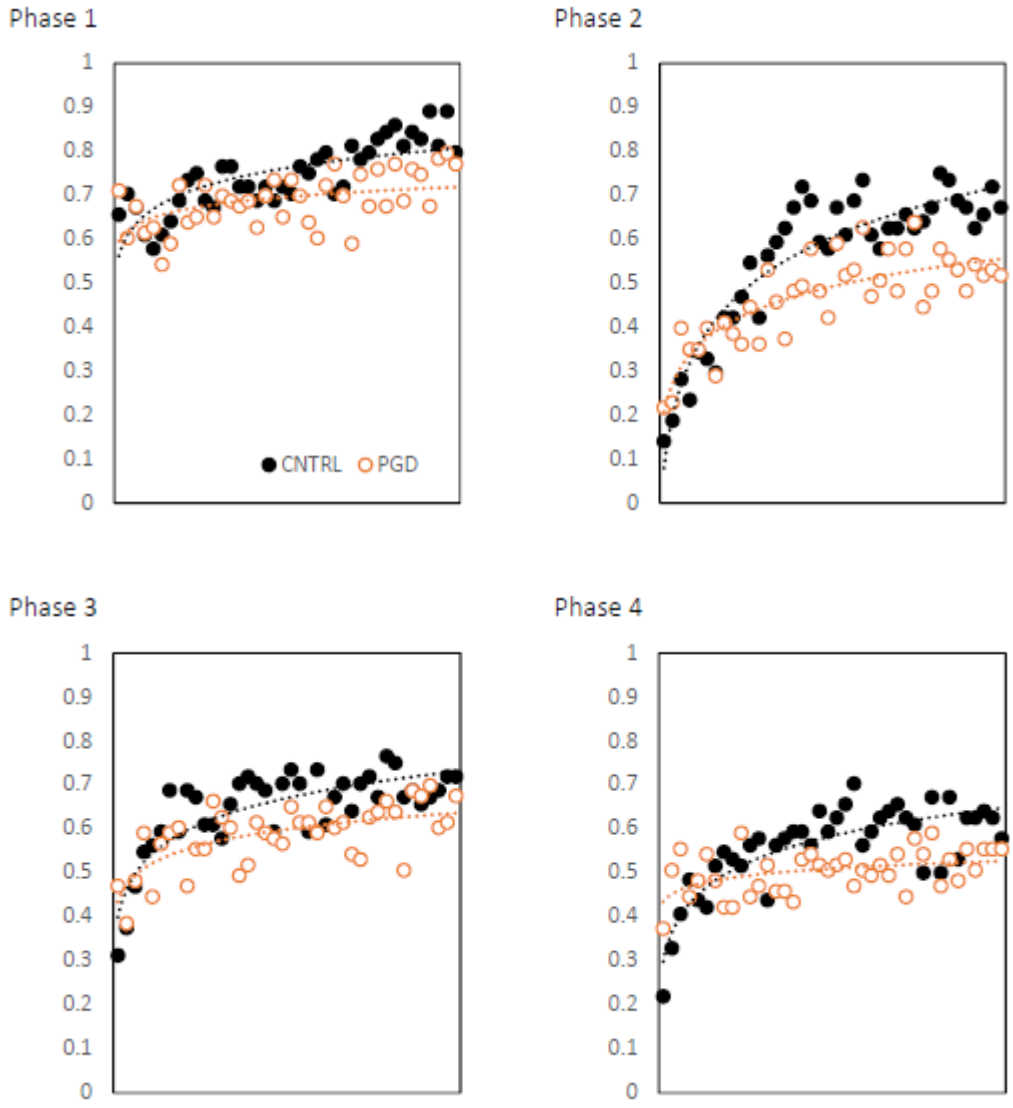


Figure A2. Evolution of differences in proportion of correct responses between groups, across phases and trials. Dots represent the difference of proportions in each single trial, and lines fitted logarithmic functions.