Neural correlates of cue-induced changes in decision-making distinguish subjects with gambling disorder from healthy controls

Running title: Neural correlates of pavlovian-to-instrumental transfer in gambling disorder

Authors:

Alexander Genauck[1)2)], Caroline Matthis[4)2)], Milan Andrejevic[7)], Lukas Ballon[1)], Francesca Chiarello[5)], Katharina Duecker[6)], Andreas Heinz[1)], Norbert Kathmann[3)], Nina Romanczuk-Seiferth[1)]

Corresponding Author:

Alexander Genauck

Department of Psychiatry and Psychotherapy

Charité - Universitätsmedizin Berlin

Charitéplatz 1, 10117 Berlin

Germany

+49 30 450 517-307

alexander.genauck@charite.de

Affiliations:

1) Charité – Universitätsmedizin Berlin, Department of Psychiatry and Psychotherapy
2) Bernstein Center for Computational Neuroscience Berlin
3) Humboldt-Universität zu Berlin, Department of Psychology
4) Technische Universität Berlin, Institute of Software Engineering and Theoretical Computer Science, Neural Information Processing

5) Psychiatry Unit, Department of Health Sciences, University of Florence, Italy
6) Carl-von-Ossietzky Universität Oldenburg, Department of Psychology
7) Melbourne School of Psychological Sciences, The University of Melbourne

Funding Sources:

This study was funded by a research grant by the Senatsverwaltung für Gesundheit, Pflege und Gleichstellung, Berlin. A.G. was funded by Deutsche Forschungsgemeinschaft (DFG) HE2597/15-1, HE2597/15-2, and DFG Graduiertenkolleg 1519 "Sensory Computation in Neural Systems".

Authors' contribution:

AG designed the experiment, collected the data, analyzed the data, wrote the manuscript. CM reviewed the machine-learning algorithm and revised the manuscript. MA collected data, revised manuscript. AH revised the manuscript, oversaw manuscript drafting. NK revised the manuscript, advised first author. LB collected and analyzed data. FC analyzed data, revised manuscript. KD collected data, analyzed data, revised manuscript. NRS designed and supervised study and experiment, oversaw manuscript drafting and data analysis.

Conflict of interest:

The authors declare no conflict of interest.

Key words: (should accompany the abstract)

gambling disorder, decision-making, pavlovian-to-instrumental transfer, fMRI

Words (including title page, abstract, captions, tables):

12504

Words abstract:

346

Date when submitted (first revision):

09.04.2019

Co-Authors' email addresses:

Caroline Matthis: matthis@ni.tu-berlin.de

Andreas Heinz: andreas.heinz@charite.de

Norbert Kathmann: kathmann@hu-berlin.de

Nina Romanczuk-Seiferth: nina.seiferth@charite.de

Francesca Chiarello: francescachiarello@gmail.com

Katharina Duecker katharina.duecker@uni-oldenburg.de

Milan Andrejevic: milan.andrejevic@unimelb.edu.au

Lukas Ballon: lballon@uos.de

Remarks

To ensure a more convenient reviewing process, we positioned figures and tables at their destined position.

ABSTRACT

Background:

Just as substance use disorders (SUDs), gambling disorder (GD) is characterized by an increase in cue-dependent decision-making (similar to Pavlovian-to-instrumental transfer, PIT). PIT, as studied in SUDs and healthy subjects, is associated with altered communication between Nucleus Accumbens (NAcc), amygdala, and orbitofrontal cortex (OFC). These neural differences are, however, poorly understood. For example, it is unclear whether they are due to the physiological effects of substance abuse, or rather related to learning processes and/or other etiological factors like innate traits associated with addiction. We have thus investigated whether network activation patterns during a PIT task are also altered in GD, an addictive disorder not involving substance abuse. We have specifically studied which *neural* PIT patterns were best at distinguishing GD from HC subjects, all to improve our understanding of the neural signatures of GD and of addiction-related PIT in general.

Methods:

30 GD and 30 HC subjects completed an affective decision-making task in a functional magnetic resonance imaging (fMRI) scanner. Gambling associated and other emotional cues were shown in the background during the task, allowing us to record multivariate neural PIT signatures focusing on a network of NAcc, amygdala and OFC. We built and tested a classifier based on these multivariate neural PIT signatures using cross-validated elastic net regression.

Results and Discussion:

As expected, GD subjects showed stronger PIT than HC subjects because they showed stronger increase in gamble acceptance when gambling cues were presented in the background. Classification based on neural PIT signatures yielded a significant AUC-ROC (0.70, p = 0.013).

When inspecting the features of the classifier, we observed that GD showed stronger PIT-related functional connectivity between NAcc and amygdala elicited by gambling background cues, as well as between amygdala and OFC elicited by negative and positive cues.

Conclusion:

We propose that HC and GD subjects are distinguishable by PIT-related neural signatures including amygdala-NAcc-OFC functional connectivity. Our findings suggest that neural PIT alterations in addictive disorders might not depend on the physiological effect of a substance of abuse, but on related learning processes or even innate neural traits, also found in behavioral addictions.

INTRODUCTION

Gambling disorder (GD) has been classified as an addiction alongside substance-use disorders (SUDs), such as alcohol or cocaine dependence (American Psychiatric Association et al., 2013). This new classification was indicated because GD and SUDs share the same core symptoms (including craving, withdrawal, tolerance) and both GD and SUDs show similar neuro-behavioral signatures (Clark, 2014; Clark et al., 2013; Leeman and Potenza, 2012; Petry et al., 2014; Romanczuk-Seiferth et al., 2014).

For instance, just like patients suffering from SUDs, GD subjects show increased neural activity elicited by addiction-related stimuli (i.e. "cues") and a reduced neural response towards stimuli signaling natural rewards (Crockford et al., 2005; Goudriaan et al., 2010; Potenza et al., 2003; Rømer Thomsen et al., 2014). In addiction, a cue can be any formerly neutral stimulus that has been repeatedly paired with the effects of the addictive behavior (Mucha et al., 2000; Potenza et al., 2003). The effect of increased responsivity towards addiction-related cues is termed cue reactivity and is pivotal in explaining a range of behaviors related to addictive disorders, such as arousal, attentional bias, craving, and relapse (Beck et al., 2012; Carter and Tiffany, 1999; Field et al., 2009; Goudriaan et al., 2010; Heinz et al., 2003; Leyton and Vezina, 2012, 2013; Schacht et al., 2013; Vezina and Leyton, 2009; Wölfling et al., 2011).

Besides cue reactivity, and just like in SUDs, GD subjects display impaired value-based decision-making. For example, GD subjects show increased risk taking, higher discounting of delayed rewards (delay discounting) and reduced loss aversion (Clark et al., 2013; Dixon et al., 2003; Genauck et al., 2017; Glimcher and Rustichini, 2004; Lorains et al., 2014; MacKillop et al., 2011; Madden et al., 2009; Petry, 2012; Platt and Huettel, 2008; Romanczuk-Seiferth et al., 2014; Wiehler and Peters, 2015). Impaired value-based decision-making in addiction may partly be explained, or even further exacerbated, by cues that modulate decision-making processes.

The modulating influence of conditioned cues on instrumental behavior (i.e. increasing the vigor with which a behavior is displayed or increasing the likelihood of choosing a certain option) has been termed Pavlovian-to-instrumental transfer (PIT) (Cartoni et al., 2016; De Tommaso et al., 2018; Schulreich et al., 2016; Talmi et al., 2008). PIT is one of the key effects deepening our understanding of cue-controlled behaviors (Dickinson Anthony and Balleine Bernard, 2002; Dickinson and Balleine, 1994; Holmes et al., 2010; Niv et al., 2007). Interestingly, PIT effects can persist even when the outcome of the instrumental behavior has been devalued (De Tommaso et al., 2018; Steenbergen et al., 2017), and a stronger PIT has been associated with heightened impulsivity (Garofalo and Robbins, 2017) and with reduced model-based behavior (Sebold et al., 2016). This is why PIT has gained considerable attention in the field of addiction research. Increased PIT has been associated with SUDs in animal studies (Corbit et al., 2007; Corbit and Janak, 2007, 2016; Krank et al., 2008; Saddoris et al., 2011) and in human studies (Garbusow et al., 2016; Schad et al., 2018).

Investigating PIT in GD is of particular importance to understanding addictive disorders in general, because GD is an addictive disorder independent of any neurotropic substance of abuse. The study of PIT in GD therefore helps us distinguish whether PIT effects seen in SUDs are a physiological result caused by the abused substance, or rather by addiction-related learning (Heinz, 2017, 113ff.), or even by innate traits putatively associated with developing and maintaining an addiction (Barker et al., 2012).

So far only a small number of studies have investigated cue-induced effects on decision-making or PIT in GD subjects. It has been observed in GD subjects that delay discounting is increased under the influence of high-craving gambling cues vs. low-craving gambling cues (Dixon et al., 2006; Miedl et al., 2014). GD subjects also have shown to be more strongly influenced by gambling cues in a response inhibition task than HC subjects (van Holst et al., 2012). To investigate PIT in GD, Genauck et al. (2019) used a mixed-gambles task, i.e. a task where participants have to decide whether they want to accept gambles that entail both possible gains and losses. They coupled the task with emotional and gambling-related cues (affective mixed-gambles task) to estimate subject-specific behavioral PIT parameters with regards to loss aversion. The authors found that behavioral PIT parameters lend themselves to classify subjects into HC vs. GD subjects. The most successful model to separate GD subjects from HC subjects was the one explaining the shift in general gamble acceptance by the influence of different cue categories, while loss aversion and loss-aversion specific PIT did not improve the distinction between GD from HC. In the present study, subjects performed a very similar affective mixed-gambles task in a functional magnetic resonance imaging (fMRI) scanner. Genauck et al. (2019) successfully used the behavioral data of the present study as an independent sample to validate their classifier.

We mentioned studies that suggest that GD is associated with increased PIT, despite the disorder being independent of any substance of abuse. However, it is unclear if there are also *neural* PIT signatures associated with cue-induced decision-making which distinguish GD from HC subjects, just like there are between SUD and HC subjects. If there are neural PIT signatures associated with GD then this would be additional evidence for functional brain changes related to addictive disorders independent of a substance of abuse (Goudriaan et al., 2010; Koehler et al., 2013; Romanczuk-Seiferth et al., 2015; Sescousse et al., 2013; van Holst et al., 2012). Our study is the first to investigate functional brain changes in GD compared to HC related to cue-induced changes in value-based decision making. We expected that neural PIT signatures derived from SUD studies should underlie behavioral PIT increase also in GD, and thus lend themselves to distinguish GD from HC subjects.

At the neural level, PIT depends on the functions of amygdala and the ventral striatum (VS) (Corbit et al., 2001; Corbit and Balleine, 2005; de Borchgrave et al., 2002; Hall et al., 2001; Prévost et al., 2012; Talmi et al., 2008). The VS denotes the ventral parts of caudate and putamen in humans and it is often used interchangeably with the nucleus accumbens (NAcc) region. Garbusow et al. (2016) distinguished alcohol dependent relapsers from abstainers using a nucleus accumbens (NAcc) PIT signal, reaching an accuracy of 71% in leave-one-out cross-validation. Note that cue reactivity, which PIT arguably is based upon, is also associated with altered activity of amygdala and NAcc in addictive disorders (Kühn and Gallinat, 2011; Schacht et al., 2013).

In addition to possible activity differences in limbic regions being associated with PIT, recent literature suggests that functional NAcc-amygdala connectivity plays a role in decision-making changes due to emotional cues (Charpentier et al., 2015). Other authors have argued that Pavlovian influence on instrumental behavior require the modulation of ongoing processes in the striatum by the amygdala (Cardinal et al., 2002; Guitart-Masip et al., 2010). Bi-directional NAcc-amygdala connectivity could thus be enhanced in GD subjects during presentation of addiction-relevant cues. Holmes et al. (2010) and Cardinal et al. (2002) further suggest a contribution of the orbital frontal cortex in integrating information about Pavlovian and instrumental processes, together with the striatum and amygdala. The ANDREA (affective neuroscience of decision through reward-based evaluation of alternatives) model makes similar predictions when explaining transient changes in gamble acceptance in decision-making tasks (Litt et al., 2008) (**Fig. 2**). In particular, this model suggests that the evaluation of a gamble involving possible gains and losses leads to a subjective value signal in the OFC. Amygdala inputs to OFC can modulate those subjective value representations when positively valued or salient stimuli (e.g. gambling cues) are shown in the background. Since there is some evidence that GD subjects show cue-induced changes in instrumental behavior and decision-making in response to gambling cues, putatively related to stronger behavioral PIT effects (Dixon et al., 2006; Genauck et al., 2019; Miedl et al., 2012; van Holst et al., 2012), this could mean that gambling cues increase the subjective gamble value represented in OFC via amygdala projections. We thus expected that stronger gambling-cue PIT-related functional connectivity from amygdala to OFC should help distinguish GD from HC.

In summary, we hypothesized that a neural PIT signature made up of several PIT-related fMRI contrasts could distinguish GD from HC subjects. We therefore compiled per subject a feature vector comprised of cue reactivity and PIT-related contrasts in amygdala and NAcc, and of functional connectivity parameters in a network of NAcc, amygdala and OFC. Hence the feature vector represented each subject's neural PIT signature, in the form of multiple functional magnetic resonance imaging (fMRI) aggregates (Seo et al., 2018; Whelan et al., 2014). We used all subjects' neural PIT signatures to estimate a classifier which would distinguish GD from HC subjects. We expected that PIT-related predictors would be found among the most important ones followed by the cue-reactivity predictors. Using cross-validation we assessed the generalizability of this classifier to new samples. Classifying GD and HC subjects using multivariate patterns aims to bring us closer to a clinically relevant characterization of the neural disturbances related to GD, especially when there are many relevant variables involved (Ahn et al., 2016; Ahn and Vassileva, 2016; Cerasa et al., 2018; Gabrieli et al., 2015; Guggenmos et al., 2018; Yarkoni and Westfall, 2017). To our knowledge, our study is the first one to use fMRI-based classification for investigating GD and its neural basis of increased PIT.

METHODS AND MATERIALS

Sample

The GD group consisted of subjects who were active gamblers, while the HC group consisted of subjects that had none or little experience in gambling. We recruited GD subjects via eBay classifieds, and notices in Berlin casinos and gambling halls. GD subjects were diagnosed using the German short questionnaire for gambling behavior (KFG) (cutoff ≥ 16) (Petry and Baulig, 1996). The KFG classifies subjects according to DSM-IV criteria for pathological gambling. However, in the following we use the DSM-5 term "gambling disorder" interchangeably, because the criteria largely overlap (Rodríguez-Testal et al., 2014). Any known history of a neurological disorder or a current psychological disorder (except tobacco dependence) as assessed by the Screening of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) (First et al., 2002) led to exclusion from the study. For further information on administered questionnaires, see **Supplements (Section 1.1)**. There were 13 subject dropouts due to technical errors, positive drug screenings, incidental cerebral anatomical findings or MRI contraindications. We dropped five more subjects to improve the matching of the groups on covariates of no interest (age, smoking severity, education, and see below). The final sample consisted of 30 GD and 30 HC subjects (**Tab. 1**). According to the South Oaks Gambling Screen (Lesieur and Blume, 1987; Stinchfield, 2002) (3-point Likert scales), GD subjects differed in gambling habits to HC only in frequency of playing slot machines (most frequent answer of GD: "3: once a week or more", HC: "1: not at all") ($t = 5.35$, $p < 0.001$), casinos (most frequent answer of GD: "3: once a week or more", HC: "1: not at all") ($t = 3.67$, $p = 0.001$), and sports betting (most frequent answer of GD: "2: less than once a week", HC: "1: not at all") ($t = 2.84$, $p = 0.003$). GD and HC were matched on relevant variables (net personal income, age, alcohol use), except for years in school (primary and secondary). We thus tested for stability of our classifier by adjusting for years in school.

Table 1: Sample characteristics, means and p-values calculated by two-sided permutation test.

variable	HC (30)	se	GD (30)	se	pooled se	p perm test
years in school	10.87	0.19	10.13	0.24	0.21	0.031
vocational school	2.73	0.29	2.07	0.25	0.27	0.108
net personal income	1028.61	92.27	1105.89	138.93	115.6	0.667
personal debt	8500	3396.88	24000	9590.36	6493.62	0.097
Fagerström	1.97	0.43	3.03	0.51	0.47	0.138
age	35.37	1.66	37.37	2.01	1.84	0.459
AUDIT	4.8	0.59	4.87	1.05	0.82	1
BDI-II	5.1	1.03	11.57	1.72	1.38	0.002
SOGS	1.73	0.47	8.8	0.67	0.57	<0.001
KFG	2.37	0.74	35	1.64	1.19	<0.001
BIS-15	31.8	0.99	36.33	1.08	1.03	0.004
GBQ persistence	1.96	0.2	3.28	0.19	0.2	<0.001
GBQ illusions	2.41	0.24	3.73	0.22	0.23	<0.001
ratio female	0.20	-	0.20	-	-	1.000
ratio unemployed	0.17	-	0.20	-	-	1.000
ratio smokers	0.60	-	0.77	-	-	0.262
ratio right-handed	0.97	-	0.84	-	-	0.204

*chi-square test used; se: bootstrapped standard errors; years in school: years in primary and secondary school; vocational school is vocational school and university; Fagerström: smoking severity (Heatherton et al., 1991); AUDIT: alcohol use disorders identification test (Dybek et al., 2006); BDI II: Beck's Depression Inventory (Beck et al., 1996), SOGS: South Oaks Gambling Screen (Lesieur and Blume, 1987); KFG: Kurzfragebogen zum Glückspielverhalten, Short Questionnaire Pathological Gambling, German diagnostic tool and severity measure based on the DSM-IV (Petry and Baulig, 1996); BIS-15: short version of the Barratt Impulsiveness Scale for impulsivity (Meule et al., 2011; Patton et al., 1995); GBQ persistence and GBQ illusions: from the Gamblers' Beliefs Questionnaire (Steenbergh et al., 2002)

Procedure and data acquisition

Before scanning, all subjects underwent urine drug testing to exclude any influence of cannabis, amphetamines, cocaine, methamphetamines, opiates, or benzodiazepines. They then were instructed on the task and completed the PIT task in a 3-Tesla SIEMENS Trio MRI (2 runs of about 23 minutes). EPI scans were acquired, as well as structural MRI. For further details on MRI sequences see **Supplements (Section 1.5)**.

Affective mixed-gambles task

We were inspired by established mixed-gambles decision-making tasks (Genauck et al., 2017; Tom et al., 2007) and mixed-gambles decision-making tasks with the influence of affective cues (Charpentier et al., 2015; Genauck et al., 2019). As affective cues, four sets of images were assembled: 1) 67 gambling images, showing a variety of gambling scenes, and paraphernalia (*gambling cues*); 2) 31 images showing negative consequences of gambling (*negative cues*); 3) 31 images showing positive effects of abstinence from gambling (*positive cues*); 4) 24 neutral IAPS images (*neutral cues*). For a description of the images and their categories see **Supplements (Section 1.2)**. The cues of all categories were presented in random order and each gambling cue only appeared once. For negative, positive, and neutral cue categories, we randomly drew images from each pool until we had presented 45 images of each category and each image at least once. Hence, we ran 202 trials in each subject. Subjects were each given 20€ for wagering. Every trial began with a fixation cross (inter-trial-interval, ITI, 4s to 8s). Then a cue as described above was presented and subjects were instructed to remember each cue for a paid recognition task after the experiment. After 4s (jittered), a mixed gamble, involving a possible gain and a possible loss, with probability P = 0.5 each, was superimposed on the cue. After another 4s (jittered) of decision time, we asked subjects to indicate how willing they were to accept the gamble by button press. This way we kept decision and motor processes apart. Subjects had to choose how willing they were to accept the gamble on a 4-point Likert-scale to ensure task engagement (Tom et al., 2007) (**Fig. 1**). Gambles were created by randomly drawing with replacement from a matrix with possible gambles consisting of 12 levels of gains (14, 16, …, 36) and 12 levels of losses (-7, -8, …, -18) (**Fig. 1**) (Genauck et al., 2017; Kahneman and

Tversky, 1979; Tom et al., 2007; Tversky and Kahneman, 1992). In every subject, we stratified gambles according to mean and variance of gain, loss, gamble variance, and Euclidean distance from gamble matrix (*ed*, i.e. gamble difficulty). We informed subjects that after completing the experiment five of their gamble decisions with ratings of “somewhat yes” or “yes” would be randomly chosen and played for real money.

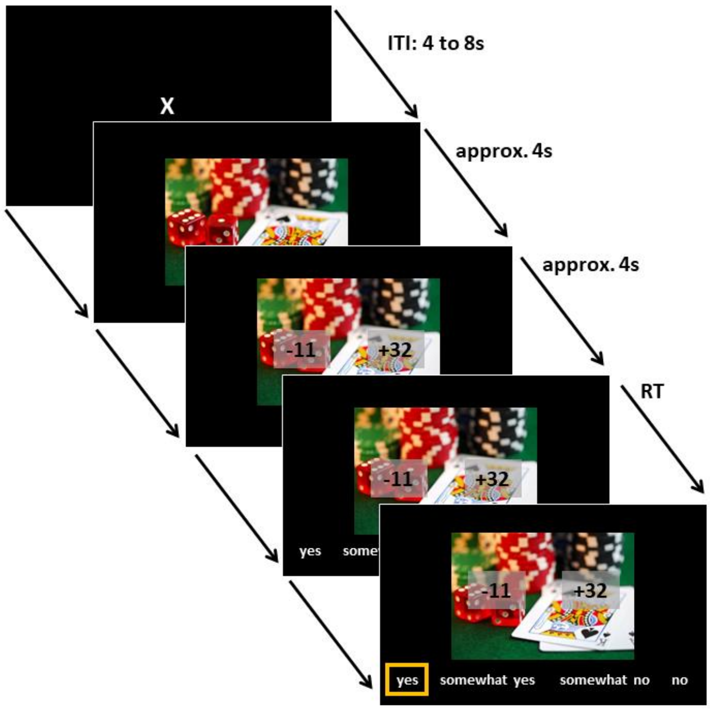

Figure 1: The affective mixed-gambles task. One trial is depicted. Subjects first saw a fixation cross with variable inter-trial-interval (ITI, 4s to 8s). Then a cue with randomly chosen affective content (67 drawn from 67 gambling related, 45 drawn from 31 with positive consequences of abstinence, 45 drawn from 31 with negative consequences of gambling, 45 drawn from 24 neutral images) was presented for about 4s. Subjects were instructed to remember the cue for a paid recognition task after all trials. Then a gamble involving a possible gain and a possible loss was superimposed on the cue (e.g. -11 and +32). Subjects were instructed to shift their attention at this point to the proposed gamble and evaluate it. Position of gain and loss was counterbalanced (left/right). Gain was indicated by a '+' sign and loss by a '-' sign. After 4s (jittered) subjects were asked to make a choice between four levels of acceptance (yes, somewhat yes, somewhat no, no; here translated from German version which used “ja, eher ja, eher nein, nein”). Direction of options (from left to right or vice versa) and side of gain amount was random. Directly after decision, the ITI started. If subjects failed to respond within 2.5s, ITI started and trial was counted as missing. RT: reaction time.

Cue ratings

After the affective mixed-gambles task, subjects rated all cues using the Self-Assessment Manikin (SAM) assessment technique (valence, arousal, dominance) (Bradley and Lang, 1994) and additional visual analogue scales. Additional questions were: 1) "How strongly does this image trigger craving for gambling?"; 2) "How appropriately does this image represent one or more gambles?"; 3) "How appropriately does this image represent possible negative effects of gambling?"; 4) "How appropriately does this image represent possible positive effects of gambling abstinence?". All cue ratings were z-standardized within subject. Cue ratings were analyzed one-by-one using linear mixed-effects regression, using lmer from the lme4 package in R (Bates et al., 2015), where cue category (and, in the respective models, clinical group) denoted the fixed effects and subjects and cues denoted the sources of random effects. Model comparisons were used to test for the effect of cue category and group and their interaction using χ^2-square difference tests. We report relevant contrast-β's only if the overall effect of the relevant factor (group, category, groupXcategory) was significant. For significance testing of those contrast-β's, we use Wald z-tests as implemented in lme4.

Behavioral data

We modeled the choice data within each subject's behavioral data by submitting dichotomized choices (somewhat no & no: 0; somewhat yes & yes: 1) into logistic regression. We dichotomized choices to increase the precision when estimating behavioral parameters, in line with previous studies (Barkley-Levenson et al., 2013; Genauck et al., 2017, 2019; Tom et al., 2007). Predictors were centralized values of gain, centralized absolute values of loss, Euclidean distance (*ed*) from gamble matrix as indicator of gamble simplicity (see **Fig. S1**) (Tom et al., 2007), and cue category (**c**). 12 steps of gain (14, 16, 18, …, 36) and 12 steps of loss (-7, -8, -9,

…, -18) formed a 12-by-12 gamble matrix, which was aggregated to 4-by-4 (e.g. gain steps 14, 16, 18 were all denoted as 16 and loss steps -18, -17, -16 were denoted as -17) as done in previous fMRI versions of this task (Genauck et al., 2017; Tom et al., 2007). We defined the gamble value (Q) on single-trial level as:

$$Q = \beta_0 + x_{gain} * \beta_{gain} + x_{loss} * \beta_{loss} + ed \cdot \beta_{ed} + c^T * \beta_c \qquad [1]$$

We call this model the **laec** model. Here c^T is a transposed column vector, denoting the dummy code of the cue's category on any given trial and β_c is a column vector holding the regression weights describing the shift in gamble value with respect to the cue category. Hence, $c^T * \beta_c$ is a scalar product describing the additive effect of cue category. We fit the logistic regression based on Eq. [1] with…

$$P(gamble\ acceptance) = 1/(1 + exp(-Q)) \qquad [2]$$

within a generalized linear mixed-effects model, using glmer from the lme4 package in R (Bates et al., 2015). Here, gain, loss, *ed*, cue category denoted the fixed effects and subjects and cues denoted the sources of random effects. To test if the groups differed in the parameters of the **laec** model, we expanded the model by an additional fixed effect of group modulating the effect of gain, loss, *ed*, and cue category (**laecg**). Statistical testing of the model comparison was performed using χ^2-square difference tests, as well as the comparison of Akaike and Bayesian information criterion (AIC, BIC). For statistical tests of single parameters in the **laecg** model, we used Wald z-tests as implemented in lme4. For more analyses of the behavioral data, please see **Supplements (Sections 1.4, 2.1)**.

FMRI data

Preprocessing and single-subject model of fMRI data

Imaging analyses were performed in SPM12 running on Matlab (R2014a). Please see **Supplements (Section 1.5)** for description of preprocessing of MRI data. We modeled the preprocessed fMRI single-subject data using three onset regressors:

1) Onset "cue" from 0s, boxcar, denoting moments of cue presentation vs. none presentation (1 vs. 0, duration: 4s plus jitter, i.e. time for showing the cue and then cue plus gamble). This onset regressor had three parametric modulators (serially orthogonalized). pmod(1): gamble cue > neutral cue, pmod(2): negative cue > neutral cue, pmod(3): positive cue > neutral cue (always coded 1 vs. -1).

2) Onset "cue plus gamble" from 4s plus jitter, boxcar, modeled the time when gamble presentation was on (1 vs. 0, duration: 4s plus jitter, i.e. the time when cue and gamble were presented but no response options available yet). This onset regressor had seven parametric modulators (serially orthogonalized). pmod(1-3): gain, loss, *ed*, mean-centered aggregated from twelve to four steps, see behavioral analysis; pmod(4): acceptance of gamble > non-acceptance (1 vs. -1); pmod(5-7): PIT modulators for the three cue categories (Garbusow et al., 2016; Schad et al., 2018). For example, the PIT regressor "acceptXgambling", pmod(5), modeled "acceptance during gambling cues vs. not accepting during gambling cues" > "accepting during neutral cues vs. not accepting during neutral cues", i.e. (1 vs. -1) > (1 vs. -1).

3) Onset "cue plus gamble plus response options" from 8s plus jitter, boxcar, modeled the time when motor response could be performed (1 vs. 0, duration: reaction time until response was made)

Missing trials were modeled with a boxcar regressor (1 vs. 0), with duration set at length of trial. Regressors were convolved with the canonical hemodynamic response function, downsampled to match the number of EPI scans and entered into a GLM (**Fig. S2**).

Extracting fMRI features for classifier building

We were interested whether PIT fMRI contrasts from a number of brain regions (regions of interest, ROIs) could predict if a subject belongs to the HC or the GD group. We hence extracted the mean activity for cue reactivity (gambling, negative, positive; pmod(1-3) of onset regressor 1) and for the PIT contrasts (acceptXgambling, acceptXnegative, acceptXpositive; pmod(5-7) of onset regressor 2) using the within-subject means from the ROIs NAcc R/L and amygdala R/L. NAcc and amygdala ROIs were taken from the Neuromorphometrics SPM12 brain atlas.

To keep in line with accounts of PIT depending on NAcc-Amy connectivity (Charpentier et al., 2015; Guitart-Masip et al., 2010) and on amygdala-OFC connectivity (Holmes et al., 2010; Litt et al., 2008) (**Fig. 2**), we also extracted functional connectivity (generalized psycho-physiological interaction, gPPI) (McLaren et al., 2012) for the PIT contrasts. We used the seeds amygdala R/L and NAcc R/L (**see Supplements, Section 1.7**). For the seeds amygdala R/L we extracted the mean from target ROIs OFC R/L (4 subregions on either side), and from target ROIs NAcc R/L. For the seeds NAcc R/L, we extracted from the target ROIs Amy R/L. Information from left medial OFC was not available due to signal loss in that region. Collecting all the extracts per subject, we had at this point for each subject a vector representing his or her specific neural PIT pattern. We z-standardized this vector for each subject. We then reduced the dimensionality of this vector for each subject by computing means (For cue reactivity: mean between respective left and right ROI; For functional connectivity: mean connectivity value between respective left and right ROI with respect to each PIT contrast, e.g. for the connectivity from NAcc to posterior OFC with respect to the PIT contrast “acceptXgambling” the mean of connectivity values from R NAcc to R posterior OFC, from L NAcc to R posterior OFC, from R NAcc to L posterior OFC, and from R NAcc to L posterior OFC).

To ensure that the task had produced meaningful signal, we checked for PIT effects in amygdala and NAcc across groups and for cue reactivity difference between groups in amygdala, NAcc and OFC using years in school as a covariate of no interest in all cases.

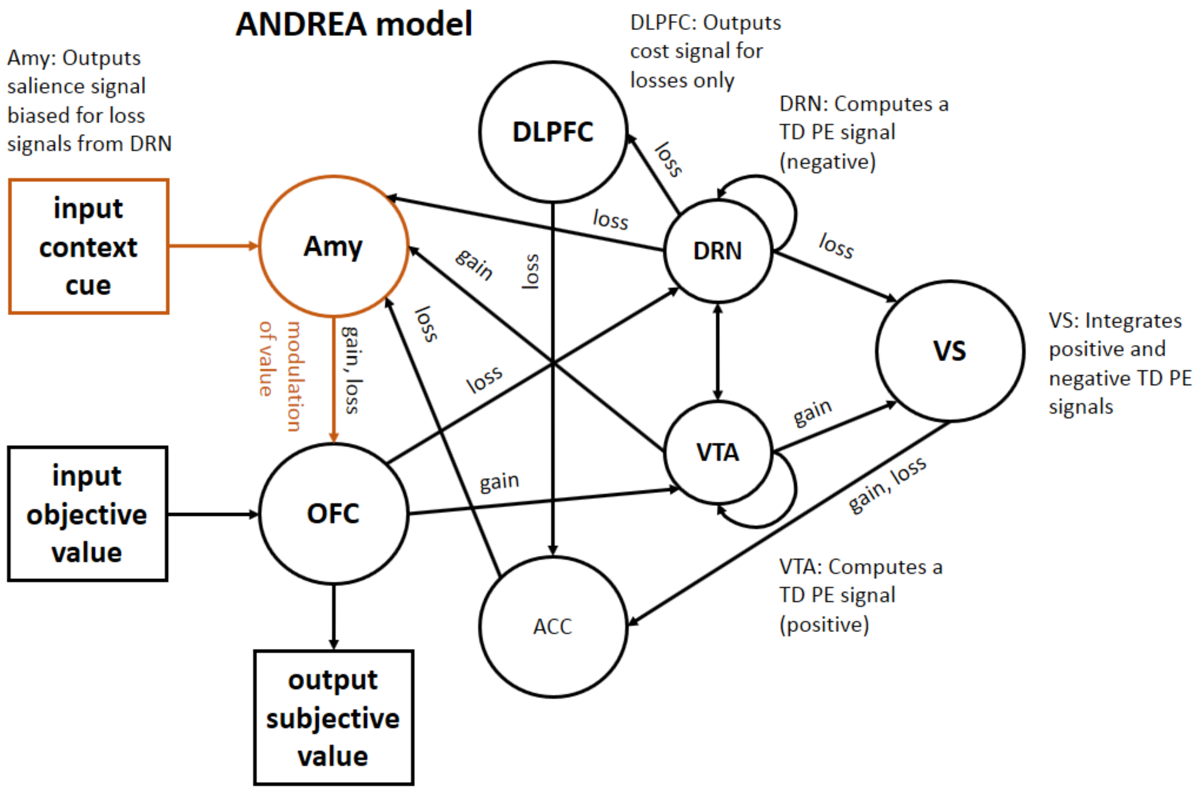

Figure 2: The ANDREA model. The model describes how loss aversion may arise in the brain during a mixed-gambles task and in addition the model makes a specific prediction how contextual cues can influence the subjective representation of gain and loss (this part of the model is highlighted in red). Namely, the amygdala is encoding and forwarding the value signal of the contextual cue, thereby modulating the subjective value representation in OFC (Litt et al., 2008). GD subjects should show a stronger functional connectivity from amygdala to OFC with respect to accepting gambles during presentation of e.g. gambling cues because this would increase the value of the gamble stored in OFC into positive direction and thus increase the likelihood of gamble acceptance.

Building the classifier based on fMRI data

The neural PIT vectors per subject were stacked into a data set. Since HC and GD were not perfectly matched on years in school, we added this variable to the data set, which was then submitted to logistic elastic net regression, with group as dependent variable. Elastic net regression is well suited for cases where there are few observations and many predictor variables that may contain groups of correlated variables (Ahn and Vassileva, 2016; Whelan et al., 2014; Zou and Hastie, 2005) (**see Supplements 1.8**). Using tuning of its two hyper-parameters (Zou and Hastie, 2005) it is also well suited to produce models that do not over-fit but generalize well to new data, especially when using cross-validation for tuning (Arlot and Celisse, 2010; Bratu et al., 2008; Varma and Simon, 2006). The algorithm tuned for optimal generalization performance on out-of-sample data using the area under the receiver-operating curve, AUC-ROC, (Ahn et al., 2016; Ahn and Vassileva, 2016; Whelan et al., 2014; Zacharaki et al., 2009). AUC-ROC ranges from 0.5 (chance) to 1 (perfect sensitivity and specificity) (Provost et al., 1998).

We assessed the generalizability of the above algorithm 1000 times via 10-fold cross-validation (Arlot and Celisse, 2010) which yielded a distribution of classifiers and thus of AUC-ROC's. Note that the cross-validation to estimate generalizability led to the cross-validations used in the elastic net regression to become *nested* (Arlot and Celisse, 2010; Bratu et al., 2008; Varma and Simon, 2006; Whelan et al., 2014). For a graphical illustration of the algorithm with cross-validation to estimate the generalization performance, see **Fig. 3**. The data and R Code can be found here: https://github.com/pransito/PIT_GD_MRI_release. We computed the mean of the obtained AUC-ROC's and estimated its p-value by performing the exact same 1000 CV rounds but each time with only "years in school" as predictor (baseline classifier). We then subtracted the AUC-ROC's of the baseline classifiers one-by-one from the 1000 AUC-ROC's of the full classifiers. This yielded a distribution of classification improvement (i.e., improvement of AUC-ROC due to using the full classifier instead of the baseline classifier). We tested this distribution against the value of classification improvement under the null-hypothesis (i.e. zero improvement) to obtain a p-value of significance of classification improvement.

After assessing the generalizability of the model by cross-validation, we then fit the model to the entire data set (no splitting in training and test data) in order to build the final interpretable and reportable classifier. Since the modelling is probabilistic, we repeated this 1000 times. We plotted the ensuing distribution of regression weight vectors as per-parameter means with 95% percentile bounds.

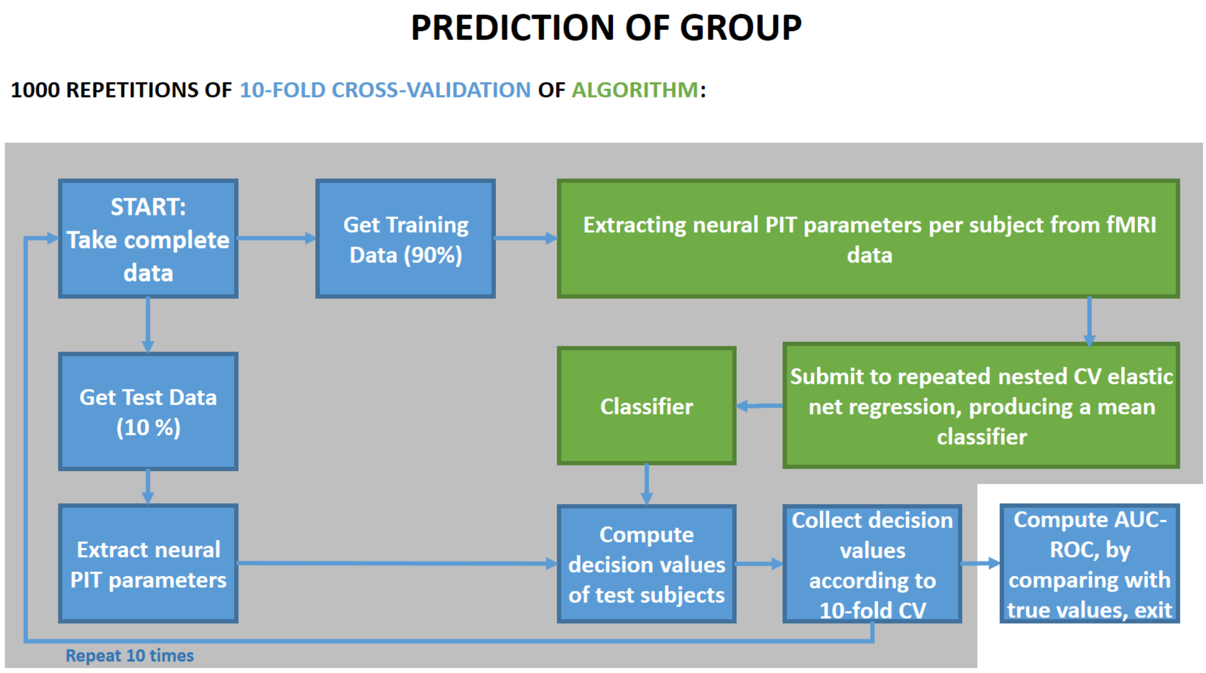

Figure 3: Classifier building algorithm with cross-validation (CV) to estimate generalization error. Nested CV was used for tuning the hyperparameters of the elastic net regression (Varma and Simon, 2006; Zou and Hastie, 2005). This was done repeatedly with different nested CV folds (10 times, 10-fold nested CV) to estimate a robust mean model within each repetition of classifier estimation.

Inspecting the classifier based on fMRI data

In order to interpret the final classifier's regression weights as an *activation pattern* (a), i.e. to know how greatly each predictor contributed to distinguishing GD from HC subjects in the classifier, we calculated:

$$\boldsymbol{a} = cov(X) * w \quad [3]$$

(Haufe et al., 2014), where w is the regression weight vector (a column vector), or in other words, the classifier. X is the matrix of predictors for all subjects and cov(X) is the covariance matrix of X. Additionally, we calculated between-group t-tests (HC vs. GD) for all predictors.

Ethics

Subjects gave written informed consent. The study was conducted in accordance with the World Medical Association Declaration of Helsinki and approved by the ethics committee of Charité - Universitätsmedizin Berlin.

RESULTS

Cue ratings

Gambling cues were seen as more appropriately representing one or more gambling games than any other cue category: gambling > neutral (β = 1.509, p < 0.001), gambling > negative (β = 1.142, p < 0.001), gambling > positive (β = 1.459, p < 0.001). They elicited more craving compared to neutral in GD subjects than in HC subjects (GD gambling > neutral: β = 1.749, HC gambling > neutral: β = 0.719, p(GD>HC) < 0.001). HC subjects indicated significantly more craving in response to gambling cues compared to neutral cues (p < 0.001). GD subjects did not rate gambling cues as more positively valenced than HC: GD > HC (β = -0.055, p < 0.712). GD subjects did not rate gambling cues as more arousal-inducing compared to HC subjects (GD gambling > neutral: 0.142, HC gambling > neutral: 0.047, p = 0.525). HC subjects did not rate gambling cues as more arousal inducing than neutral cues (p = 0.662). Gambling cues lead to higher dominance ratings overall: gambling > neutral (β = 0.368, p < 0.001). GD subjects rated gambling cues as more dominance inducing than HC subjects: GD > HC (β = 0.328, p = 0.021).

Negatively valenced cues were seen as evoking smaller valence ratings than all other categories: negative < neutral (β = 0.651, p < 0.001), negative < positive (β = 1.538, p < 0.001), negative < gambling (β = 0.977, p < 0.001). Negative cues lead to lower dominance ratings (β = -0.297, p < 0.001). There were no group differences on any rating scale with regards to the negative cues.

Negative cues were more representative of negative effects of gambling than any other group: negative > neutral (β = 1.398, p < 0.001), negative > positive (β = 1.388, p < 0.001), negative > gambling (β = 0.826, p < 0.001). GD subjects perceived negative cues as less representative for negative consequences of gambling than HC subjects (HC: 2.03, GD: 1.388, p < 0.001). Positive cues were more representative of positive effects of abstinence from gambling than any other category: positive > neutral ($\beta = 0.970$, $p < 0.001$), positive > negative ($\beta = 0.848$, $p < 0.001$) and positive > gambling ($\beta = 0.639$, $p < 0.001$), and rated as more positive (valence) than any other category: positive > neutral ($\beta = 0.886$, $p < 0.001$), positive > negative ($\beta = 1.538$, $p < 0.001$) and positive > gambling ($\beta = 0.561$, $p < 0.001$). Positive cues lead to higher dominance ratings: positive > neutral ($\beta = 0.683$, $p < 0.001$). There were no group differences on any rating scale with regards to the positive cues. (**Fig. S3**).

Describing the behavioral choice data

Here we present results comparing the **laec** model against the **laecg** model (i.e. with an effect of group onto the fixed effects of gain, loss, *ed* and category). Comparing the two models, we observed a significant χ^2 difference test result ($\chi^2 = 26.6$, df = 7, $p < 0.001$; with ΔAIC = 12.6, ΔBIC = -39.0). Inspecting the estimated parameters of the **laecg** model, we observed that acceptance rate during neutral images with all other parameters at zero (i.e. at their mean, except for *ed*, actually zero) was for HC: 59.0% and for GD: 38.8%, $p_{Wald} = 0.155$. Gambling cues were associated with stronger increase in gamble acceptance in GD subjects ($\Delta\% = 44$) than in HC subjects ($\Delta\% = -8$, $p_{Wald} = 0.003$). The same was true for negative (GD: $\Delta\% = 23$, HC: $\Delta\% = -16$, $p_{Wald} = 0.049$) and positive cues (GD: $\Delta\% = 23$, HC: $\Delta\% = 0$, $p_{Wald} = 0.030$) (**Fig. 4**). For further behavioral results, please see **Supplements (Section 2.1)**.

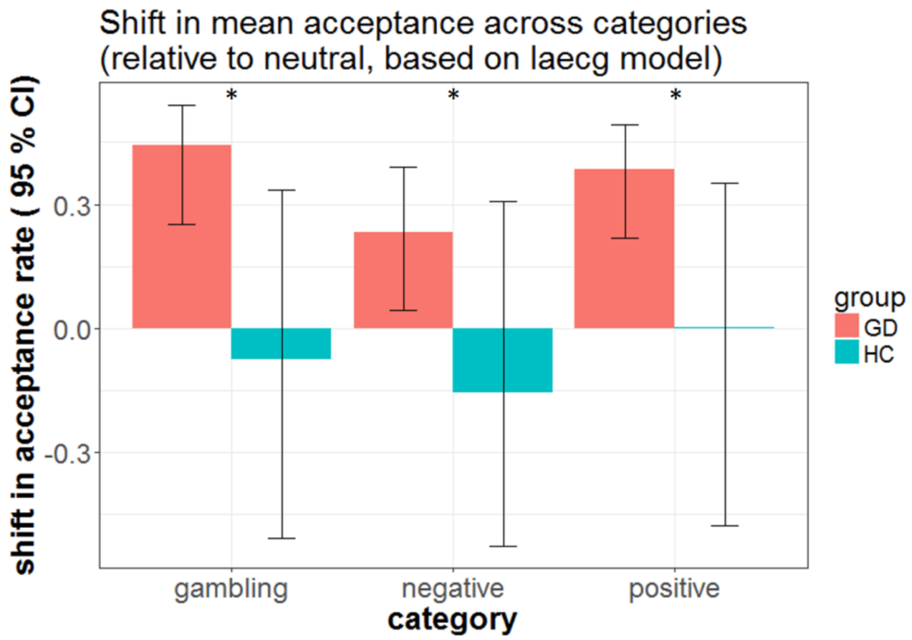

Figure 4: Shift in acceptance rate during gambles per category and group. Based on the **laecg** model. GD subjects show stronger increase in gamble acceptance (compared to neutral) in comparison to HC subjects during the presentation of all three cue categories in the background. CIs based on standard errors of parameter estimates. Stars denote significant post-hoc contrasts.

Prediction of group using fMRI data

Across groups and in line with previous findings (Garbusow et al., 2016; Guitart-Masip et al., 2010; Prévost et al., 2012; Talmi et al., 2008), there was for gambling-cues PIT a significant effect in right amygdala: [15 -6 -15], $p_{SVC} = 0.027$, $p_{uncor} = 0.003$, $k = 17$, trendwise effects in left NAcc: [-6 12 -3], $p_{SVC} = 0.082$, $p_{uncor} = 0.015$, $k = 6$, right NAcc: [6 12 -3], $p_{SVC} = 0.060$, $p_{uncor} = 0.014$, $k = 9$, and in left amygdala: [-24 -3 -18], $p_{SVC} = 0.071$, $p_{uncor} = 0.009$, $k = 31$.
In line with previous findings (Goudriaan et al., 2010; Limbrick-Oldfield et al., 2017; Potenza et al., 2003), there was for the cue reactivity contrast GD > HC (gambling cues) a trendwise effect in right amygdala: [18 -3 -15], $p_{SVC} = 0.091$, $p_{uncor} = 0.012$, $k = 5$. Further, there was for HC > GD (positive cues) a significant effect in left NAcc: [-6 6 -6], $p_{SVC} = 0.033$, $p_{uncor} = 0.005$,

k = 4, and in right NAcc: [6 9 -6], p_{SVC} = 0.035, p_{uncor} = 0.007, k = 4. For HC > GD (negative cues), there was a trendwise effect in left lateral OFC: [-36 36 -18], p_{SVC} = 0.089, p_{uncor} = 0.004, k = 24, and in right anterior OFC: [24 45 -12], p_{SVC} = 0.080, p_{uncor} = 0.005, k = 8.

The mean AUC-ROC of the full classifier using neural PIT signatures was 70.0% (mean for the baseline classifier, i.e. covariate-only classifier: 61.5%, p = 0.013) (**Fig. 5**).

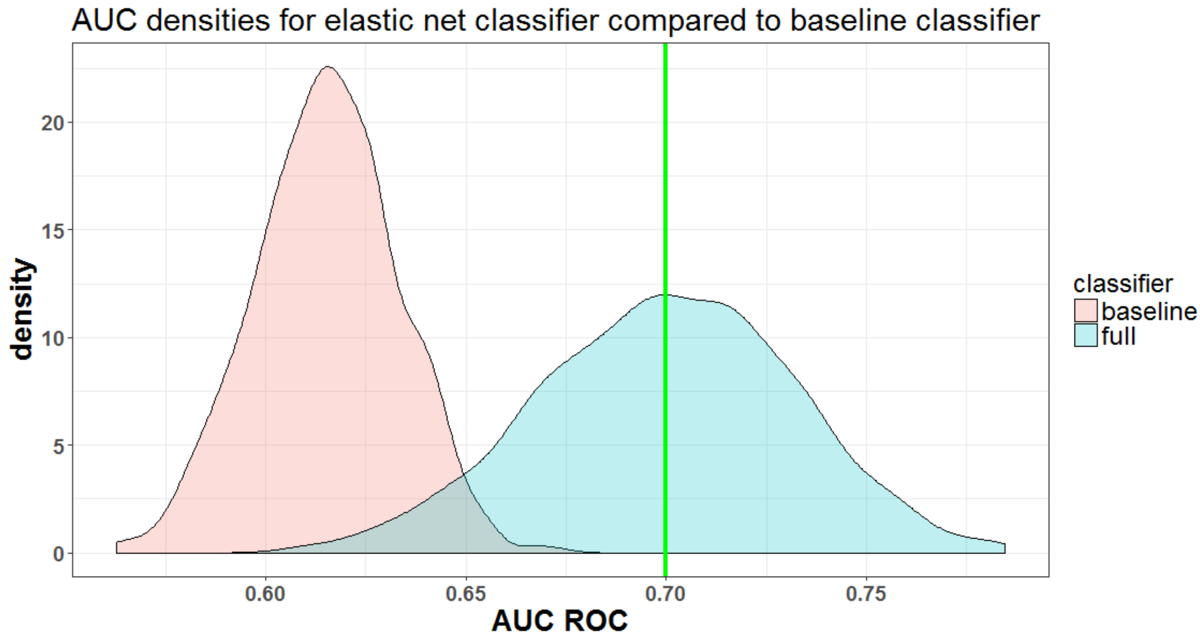

Figure 5: Classification performance of classifier using fMRI neural PIT signatures. Blue is the density plot of 1000 AUC-ROCs obtained from running 1000 repetitions of cross validation of the full classifier using neural PIT signatures. The green line shows the mean of these 1000 AUC-ROCs. In red you see the same density estimate for the baseline classifier, i.e. the covariate-only classifier, as a control condition. The full classifier performs significantly better (p = 0.013).

We ran the algorithm on the complete data set of fMRI variables. Inspecting the classifier's logistic regression weights (see **Fig. S4**) (after transformation to predictor importance, see Eq. 3, and according to t-tests), we saw that the top predictor was negative-cues-PIT-related functional connectivity from amygdala to anterior OFC, with a negative sign (**Fig. 6**). This means that the stronger not accepting a gamble was associated with increase in correlation between amygdala and anterior OFC, the *less* likely the subject was a GD person (and rather a HC subject). In other words, GD subjects showed lower such functional connectivity than HC. The next top three predictors were gambling-cues-related functional connectivity from NAcc to amygdala (positive sign), positive-cues-related functional connectivity from amygdala to lateral OFC (positive sign), and years in school (negative sign) (see **Fig. 6, 7**).

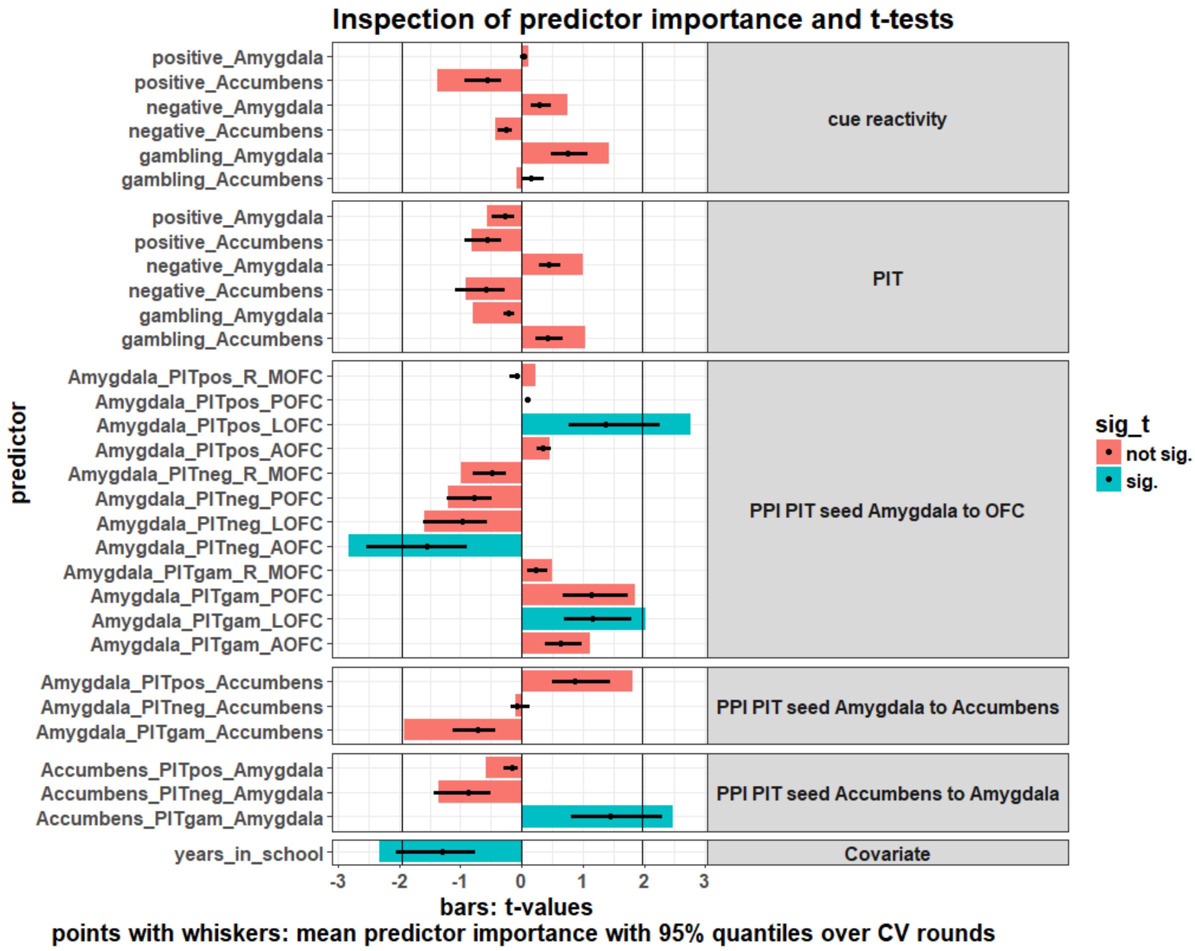

Figure 6: Estimated predictor importance. Points and quantiles are estimated predictor importance with 95%-quantiles over 1000 classifier estimation rounds. The larger the absolute size of an importance value the stronger the predictor adds to distinguishing HC from GD in the classifier. Bars show t-values of simple between- group t-tests. Significant t-tests are highlighted (Welch-test, $p < 0.05$, two-sided). Delimitations are at 1.96 and -1.96 to mark points of statistical significance for t-test. Importance values/t-values are grouped by the kind of fMRI

predictor: cue reactivity related, PIT related, Psychological-physiological-interaction (i.e. PPI) related. PPIs are further grouped by seed region and target extraction (e.g. “to OFC”). PIT: pavlovian-to-instrumental transfer; OFC: orbital frontal cortex; AOFC, LOFC, POFC, MOFC: anterior, lateral, posterior, medial orbital frontal cortex; R: right

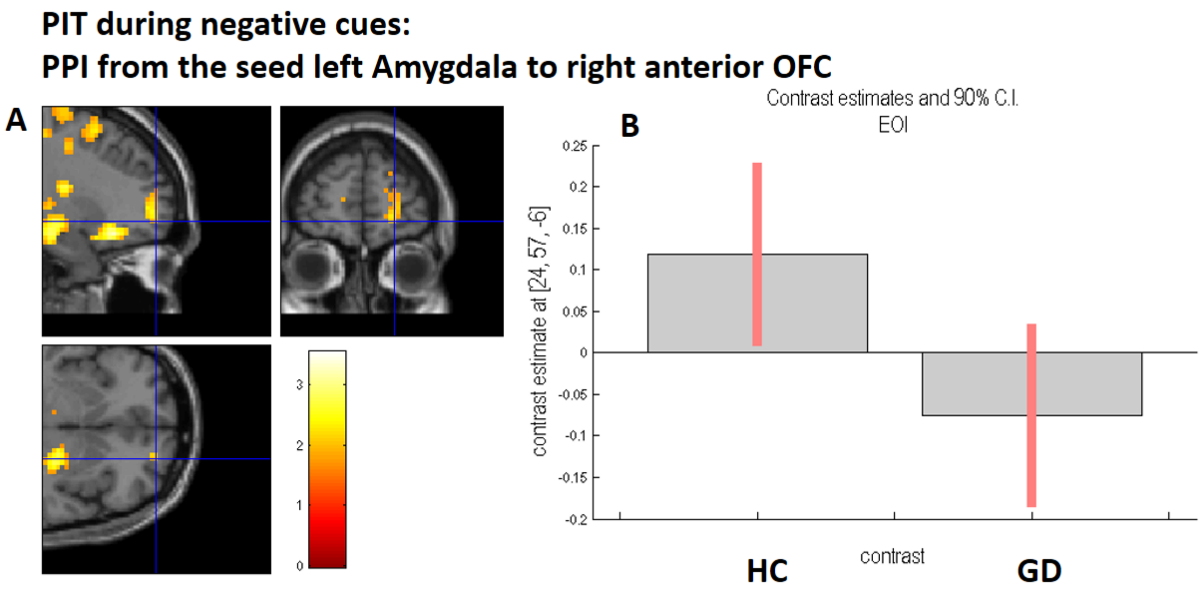

Figure 7: Generalized psycho-physiological interaction (gPPI) T-map for Pavlovian-to-instrumental transfer (PIT) during negative cues contrast (HC > GD). Displayed at $p < 0.05$. Illustration of contrast that contributed to most important predictor of the classifier. **A:** from left amygdala to right anterior orbito-frontal cortex (OFC), centered at peak within region of interest, [24, 57, -6], $p_{unc} = 0.024$, $k > 0$. Cluster extends into right superior frontal gyrus. Also visible is significant activity in right posterior OFC, which was also picked up by the classifier. **B:** plot of subjects’ beta values at peak voxel for the contrast in A. HC: healthy controls, GD: gambling disorder subjects

DISCUSSION

GD is characterized by impaired decision making (Wiehler and Peters, 2015) and craving in response to gambling associated images (Crockford et al., 2005; Goudriaan et al., 2010). There is evidence that there is significant influence of cue-induced emotional states onto value-based decision-making in GD subjects (Dixon et al., 2006; Genauck et al., 2019; Miedl et al., 2012; van Holst et al., 2012). The influence of cues onto value-based decision-making may be regarded as a form of Pavlovian-to-Instrumental Transfer (PIT), the increase of which has been associated with addictive disorders in general (Corbit and Janak, 2007; Garbusow et al., 2016; Genauck et al., 2013; Mitchell et al., 2016; Schad et al., 2018).

In the current study, we hypothesized that GD subjects should be distinguishable by neural PIT signatures based on fMRI contrasts recorded during an affective mixed-gambles task. We therefore built a classifier using fMRI PIT contrasts to distinguish GD from HC subjects focusing on brain structures known to be relevant in PIT, like amygdala and NAcc (ventral striatum). We also incorporated amygdala's connectivity to OFC, and amygdala's and NAcc's connectivity to each other. We further included neural cue reactivity contrasts as predictors. All these predictors yielded a neural PIT signature per subject which could be used to classify subjects into the GD or HC group.

Our results support our first hypothesis, showing that neural PIT signatures based on fMRI data gathered from the affective mixed-gambles task may successfully classify out-of-sample subjects into GD and HC, with a cross-validated mean AUC-ROC of 70.0% ($p = 0.013$). This performance on out-of-sample data is similar to other studies using MRI data for classification in the field of addictive disorders (Guggenmos et al., 2018; Pariyadath et al., 2014; Seo et al., 2015, 2018; Whelan et al., 2014). To our knowledge, however, the present study is the first one to use fMRI classification for investigating a behavioral addiction, namely GD, and the neural basis of increased PIT. This means that it is possible to characterize a non-substance related addiction to a considerable degree by a single neuro-functional signature, namely a neural PIT signature in a network of amygdala, NAcc and OFC, derived from PIT and SUD literature. This further implies that addictive disorders, in general, may be associated with PIT-related neural changes, independent of a substance of abuse, which means that neural PIT changes may be a product of addiction-related learning and neural plasticity or even of an innate trait (Barker et al., 2012).

Concerning the predictors in the classifier, we hypothesized that gambling-cue PIT-related functional connectivity from amygdala to OFC should be increased. We found that multiple PIT-related functional connectivities from amygdala to OFC were significant predictors in the classifier. For example, gambling-cues PIT-related functional connectivity from amygdala to OFC was increased in GD compared to HC subjects, in line with our hypothesis and in line with the general prediction that in GD subjects amygdala modulates the value computation in OFC, when addiction-related cues are presented in the background (Cardinal et al., 2002; Holmes et al., 2010; Litt et al., 2008). Furthermore, the top predictor in the classifier was PIT-related functional connectivity from amygdala to anterior OFC in trials with a *negative* cue, with a negative predictor weight. This means that the stronger the rejection of a gamble during the presentation of negative cues was associated with an increase in correlation between amygdala and anterior OFC, the *less* likely the subject was a GD person (and rather a HC subject). In other words, GD subjects showed weaker such functional connectivity than HC. GD subjects, compared to HC subjects, showed significantly more gambling during the presentation of negative cues than during the presentation of neutral cues. HC subjects may not show this effect because of stronger signal transmission related to negative cues from amygdala to OFC. Similarly, it has been found that reduced loss aversion in GD subjects was associated with reduced loss-related functional connectivity from amygdala to ventral medial prefrontal cortex in a pure mixed-gambles task (Genauck et al., 2017). This highlights that impaired decision-making in GD during a pure mixed-gambles task, as well as during an affective mixed-gambles task, may draw from the same functional neural substrate.

We looked at the next two top predictors expecting that PIT-related (as opposed to purely cue reactivity related) neural predictors should be among these. Indeed, we found that the next top predictor was gambling-cues PIT-related functional connectivity from NAcc to amygdala (positive sign), a connectivity important for cue-induced effects in mixed-gambles tasks (Charpentier et al., 2015). This means that the more gamble acceptance during presentation of gambling cues was associated with an increase in correlation between NAcc and amygdala, the *more* likely the subject was a GD person. In other words, GD subjects showed stronger such functional connectivity than HC. NAcc is seen as encoding temporal difference prediction errors, i.e. it fires when an unexpected reward signal is perceived from one moment to the next (McClure et al., 2003; Niv and Schoenbaum, 2008; O'Doherty et al., 2003; Schultz et al., 1997). GD subjects rated gambling pictures as more craving-inducing and reacted with significantly stronger gamble acceptance increase than HC when gambling-associated cues were shown in the background. We also saw an important regression weight given to gambling-cues PIT-related functional connectivity from amygdala to OFC, in line with our initial hypothesis. Therefore, it may be that gambling cues elicit a prediction error in NAcc that modulates amygdala activity, which in turn modulates the value representation in OFC in such a way that GD subjects are more inclined than HC subjects to accept the gamble at hand. This is in line with a previous study, where it has been found that GD subjects display increased functional connectivity from amygdala to posterior OFC related to increasing possible gains in a pure mixed-gambles task (Genauck et al., 2017). This highlights again that impaired decision-making in GD during a pure mixed-gambles task, as well as during an affective mixed-gambles task may draw from the same functional neural substrate. Also, it has been observed before that NAcc and amygdala seem to hold relevant signal related to PIT in healthy subjects (Prévost et al., 2012) and to increased PIT in addicted subjects (Garbusow et al., 2016). Interestingly, previous studies (Garbusow et al., 2016; Schad et al., 2018) have observed that in recently detoxified treatment-seeking AD patients, images of alcoholic beverages in the background have a suppressing effect on the instrumental task in the foreground. Contrarily, we have seen that gambling cues elicit a stronger gamble acceptance increase in GD than in HC. This may be because we have included only active non-treatment-seeking gamblers, who at that stage of disease show little activity working against their automated response towards addiction-related cues.

The third top predictor was also PIT related, in line with our hypothesis that PIT-related predictors should be more important than cue reactivity predictors. It was positive-cues PIT-related functional connectivity from amygdala to lateral OFC. This means that the stronger the acceptance of a gamble during the presentation of positive cues was associated with an increase in correlation between amygdala and OFC, the *more* likely the subject was a GD person. In other words, GD subjects showed stronger such functional connectivity than HC. This may be seen as parallel to the finding on behavioral level that GD subjects react with more gambling increase to positive pictures than HC subjects. It seems that both positive cues and gambling cues lead to increased gambling and similarly increased connectivity between amygdala and OFC in GD subjects. Also, negative cues lead to increased gambling. This is surprising because one could have expected to see decreased gambling during negative and positive cues or no effect of those cue categories (Genauck et al., 2019). On the other hand, perhaps *all three* cue categories have special salience for GD subjects, because they may have already thought about and experienced the positive and negative effects of their gambling behavior and this might conjure motivation to act. Negative cues might have even evoked a need for loss chasing and thus more gambling (Ciccarelli et al., 2019). Thus, generally, all three cue categories may lead to increased vigor/motivation, leading to a stronger propensity to accept gambles. Future studies should further explore the effect of positive and negative stimuli on gambling in GD.

Although we discussed the top-three predictors, note that our classifier is truly multivariate. Of the 30 neural PIT signature predictors, 22 received a significant regression weight (and 24 a significant activation weight), despite elastic net regression allowing for total deselection of predictors (Zou and Hastie, 2005). This means that just like on the behavioral level, where GD subjects reacted more strongly than HC to all non-neutral categories, we see that fMRI activities related to *all* non-neutral categories was relevant for characterizing GD. Cue reactivity regression weights are relatively small and the classifier heavily draws on PIT-related variables (the top-three predictors were PIT related). This emphasizes the importance of PIT as a defining marker for addictive disorders beyond cue reactivity.

We used the same cues as Genauck et al. (2019) in a new sample of GD and HC subjects and, in line with that study, we also observed that GD subjects rate the gambling cues as more craving inducing. Also, in the other categories cues were perceived as expected. The ratings and the result that neural PIT signatures successfully distinguish GD from HC subjects reinforce the notion that GD subjects' cue reactivity facilitates riskier decision-making when addiction-related cues are presented in the background of a gamble task.

Changes in NAcc's structure (Koehler et al., 2015) and function (Koehler et al., 2013; Linnet et al., 2010; Miedl et al., 2014; Reuter et al., 2005; Romanczuk-Seiferth et al., 2015) related to GD have been observed in previous studies. The same is true for amygdala's structure (Elman et al., 2012; Takeuchi et al.) and function (Genauck et al., 2017), as well as for OFC's structure

(Li et al., 2018) and function (Cavedini et al., 2002; Goudriaan et al., 2010). Our study adds to these findings by considering the functions of these structures concurrently and in a network. Our results support the notion that GD, similar to SUD, is characterized by neural incentive sensitization (Limbrick-Oldfield et al., 2017; Rømer Thomsen et al., 2014) (Limbrick-Oldfield et al., 2017; Rømer Thomsen et al., 2014) such that in GD a network of amygdala, NAcc and OFC facilitate gambling decisions in the face of gambling cues.

STRENGTHS AND LIMITATIONS

The main strength of our study is that we have used a classification approach to assess the usefulness of known neural PIT contrasts to characterize GD in out-of-sample data. Using this approach, we have estimated the single-subject relevance of these fMRI signals. Our results therefore have not only explanatory value in elucidating the basis of increased PIT in GD, but also predictive value, given that they are likely to be found in new samples of GD and matched HC subjects (Yarkoni and Westfall, 2017). Furthermore, we are to our knowledge the first to address the neural underpinnings of PIT in a behavioral addiction using a machine learning approach. Unfortunately, we have no independent validation sample to externally validate our results (Genauck et al., 2019; Guggenmos et al., 2018). Further studies are needed to collect such data. As we have laid out, there are multiple ways in which the brain may produce an overt PIT, involving at least amygdala, NAcc and OFC. To increase statistical power, we have omitted other conceptualization of PIT, e.g. as an interference task (Sommer et al., 2017), and hence any limbic-dorso-lateral-prefrontal connectivity (Bray et al., 2008). Future studies should explore this. In the current study we did not address the distinction between outcome-specific and general PIT (Bray et al., 2008; Corbit and Janak, 2007; Eder and Dignath, 2016; Hogarth and Chase, 2012; Lewis et al., 2013; Steenbergen et al., 2017). This would be a valuable next step for future studies in GD.

CONCLUSION

We have observed that it is possible to classify HC and GD subjects based on the neural correlates of PIT in a network of NAcc, amygdala and OFC. Our findings further the understanding of GD and show that PIT is relevant for characterizing non-substance-related addictive disorders also on neural level. PIT alterations at the neural level related to an addictive disorder might thus not depend on the direct effect of a substance of abuse, but rather on related learning processes or even on innate traits.

ACKNOWLEDGMENT

This study was conducted at the BCAN - Berlin Center of Advanced Neuroimaging.

ONLINE RESOURCES

R code and data (stored in an .RData file which is loaded with the R code) to run the classifier estimation and cross-validation, as well as the classical hierarchical regression analyses can be found in the following link. Further you can find there also more detailed data concerning the MRI sequences, as well as the preprocessing of MRI data and the fMRI single subject design: https://github.com/pransito/PIT_GD_MRI_release

REFERENCES

Ahn, W.-Y., Ramesh, D., Moeller, F. G., and Vassileva, J. (2016). Utility of Machine-Learning Approaches to Identify Behavioral Markers for Substance Use Disorders: Impulsivity Dimensions as Predictors of Current Cocaine Dependence. *Front. Psychiatry* 7. doi:10.3389/fpsyt.2016.00034.

Ahn, W.-Y., and Vassileva, J. (2016). Machine-learning identifies substance-specific behavioral markers for opiate and stimulant dependence. *Drug Alcohol Depend.* 161, 247–257. doi:10.1016/j.drugalcdep.2016.02.008.

American Psychiatric Association, American Psychiatric Association, and DSM-5 Task Force (2013). *Diagnostic and statistical manual of mental disorders: DSM-5.* Arlington, Va.: American Psychiatric Association.

Arlot, S., and Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Stat. Surv.* 4, 40–79. doi:10.1214/09-SS054.

Barker, J. M., Torregrossa, M. M., and Taylor, J. R. (2012). Low prefrontal PSA-NCAM confers risk for alcoholism-related behavior. *Nat. Neurosci.* 15, 1356–1358. doi:10.1038/nn.3194.

Barkley-Levenson, E. E., Van Leijenhorst, L., and Galván, A. (2013). Behavioral and neural correlates of loss aversion and risk avoidance in adolescents and adults. *Dev. Cogn. Neurosci.* 3, 72–83. doi:10.1016/j.dcn.2012.09.007.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. *R Package Version 11-8.* Available at: http://keziamanlove.com/wp-content/uploads/2015/04/StatsInRTutorial.pdf [Accessed July 28, 2015].

Beck, A. T., Steer, R. A., and Brown, Gk. (1996). BDI-II, Beck depression inventory: manual: Psychological Corp. *San Antonio TX.*

Beck, A., Wüstenberg, T., Genauck, A., Wrase, J., Schlagenhauf, F., Smolka, M. N., et al. (2012). Effect of Brain Structure, Brain Function, and Brain Connectivity on Relapse in Alcohol-Dependent PatientsRelapse in Alcohol-Dependent Patients. *Arch. Gen. Psychiatry* 69, 842–852.

Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59.

Bratu, C. V., Muresan, T., and Potolea, R. (2008). Improving classification accuracy through feature selection. in *2008 4th International Conference on Intelligent Computer Communication and Processing*, 25–32. doi:10.1109/ICCP.2008.4648350.

Bray, S., Rangel, A., Shimojo, S., Balleine, B., and O'Doherty, J. P. (2008). The Neural Mechanisms Underlying the Influence of Pavlovian Cues on Human Decision Making. *J. Neurosci.* 28, 5861–5866. doi:10.1523/JNEUROSCI.0897-08.2008.

Cardinal, R. N., Parkinson, J. A., Hall, J., and Everitt, B. J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci. Biobehav. Rev.* 26, 321–352.

Carter, B. L., and Tiffany, S. T. (1999). Meta-analysis of cue-reactivity in addiction research. *Addiction* 94, 327–340. doi:10.1046/j.1360-0443.1999.9433273.x.

Cartoni, E., Balleine, B., and Baldassarre, G. (2016). Appetitive Pavlovian-instrumental Transfer: A review. *Neurosci. Biobehav. Rev.* 71, 829–848. doi:10.1016/j.neubiorev.2016.09.020.

Cavedini, P., Riboldi, G., Keller, R., D'Annucci, A., and Bellodi, L. (2002). Frontal lobe dysfunction in pathological gambling patients. *Biol. Psychiatry* 51, 334–341. doi:10.1016/S0006-3223(01)01227-6.

Cerasa, A., Lofaro, D., Cavedini, P., Martino, I., Bruni, A., Sarica, A., et al. (2018). Personality biomarkers of pathological gambling: A machine learning study. *J. Neurosci. Methods* 294, 7–14. doi:10.1016/j.jneumeth.2017.10.023.

Charpentier, C. J., Martino, B. D., Sim, A. L., Sharot, T., and Roiser, J. P. (2015). Emotion-induced loss aversion and striatal-amygdala coupling in low-anxious individuals. *Soc. Cogn. Affect. Neurosci.*, nsv139. doi:10.1093/scan/nsv139.

Ciccarelli, M., Cosenza, M., D'Olimpio, F., Griffiths, M. D., and Nigro, G. (2019). An experimental investigation of the role of delay discounting and craving in gambling chasing behavior. *Addict. Behav.* doi:10.1016/j.addbeh.2019.02.002.

Clark, L. (2014). Disordered gambling: the evolving concept of behavioral addiction. *Ann. N. Y. Acad. Sci.* 1327, 46–61. doi:10.1111/nyas.12558.

Clark, L., Averbeck, B., Payer, D., Sescousse, G., Winstanley, C. A., and Xue, G. (2013). Pathological Choice: The Neuroscience of Gambling and Gambling Addiction. *J. Neurosci.* 33, 17617–17623. doi:10.1523/JNEUROSCI.3231-13.2013.

Corbit, L. H., and Balleine, B. W. (2005). Double Dissociation of Basolateral and Central Amygdala Lesions on the General and Outcome-Specific Forms of Pavlovian-Instrumental Transfer. *J. Neurosci.* 25, 962–970. doi:10.1523/JNEUROSCI.4507-04.2005.

Corbit, L. H., and Janak, P. H. (2007). Ethanol-associated cues produce general pavlovian-instrumental transfer. *Alcohol. Clin. Exp. Res.* 31, 766–774. doi:10.1111/j.1530-0277.2007.00359.x.

Corbit, L. H., and Janak, P. H. (2016). Habitual Alcohol Seeking: Neural Bases and Possible Relations to Alcohol Use Disorders. *Alcohol. Clin. Exp. Res.*, n/a-n/a. doi:10.1111/acer.13094.

Corbit, L. H., Janak, P. H., and Balleine, B. W. (2007). General and outcome-specific forms of Pavlovian-instrumental transfer: the effect of shifts in motivational state and inactivation of the ventral tegmental area. *Eur. J. Neurosci.* 26, 3141–3149. doi:10.1111/j.1460-9568.2007.05934.x.

Corbit, L. H., Muir, J. L., and Balleine, B. W. (2001). The Role of the Nucleus Accumbens in Instrumental Conditioning: Evidence of a Functional Dissociation between Accumbens Core and Shell. *J. Neurosci.* 21, 3251–3260.

Crockford, D. N., Goodyear, B., Edwards, J., Quickfall, J., and el-Guebaly, N. (2005). Cue-Induced Brain Activity in Pathological Gamblers. *Biol. Psychiatry* 58, 787–795. doi:10.1016/j.biopsych.2005.04.037.

de Borchgrave, R., Rawlins, J. N. P., Dickinson, A., and Balleine, B. W. (2002). Effects of cytotoxic nucleus accumbens lesions on instrumental conditioning in rats. *Exp. Brain Res.* 144, 50–68. doi:10.1007/s00221-002-1031-y.

De Tommaso, M., Mastropasqua, T., and Turatto, M. (2018). Working for beverages without being thirsty: Human Pavlovian-instrumental transfer despite outcome devaluation. *Learn. Motiv.* 63, 37–48. doi:10.1016/j.lmot.2018.01.001.

Dickinson, A., and Balleine, B. (1994). Motivational control of goal-directed action. *Anim. Learn. Behav.* 22, 1–18. doi:10.3758/BF03199951.

Dickinson Anthony, and Balleine Bernard (2002). The Role of Learning in the Operation of Motivational Systems. *Stevens Handb. Exp. Psychol.* doi:10.1002/0471214426.pas0312.

Dixon, M. R., Jacobs, E. A., and Sanders, S. (2006). Contextual Control of Delay Discounting by Pathological Gamblers. *J. Appl. Behav. Anal.* 39, 413–422. doi:10.1901/jaba.2006.173-05.

Dixon, M. R., Marley, J., and Jacobs, E. A. (2003). Delay discounting by pathological gamblers. *J. Appl. Behav. Anal.* 36, 449–458. doi:10.1901/jaba.2003.36-449.

Dybek, I., Bischof, G., Grothues, J., Reinhardt, S., Meyer, C., Hapke, U., et al. (2006). The Reliability and Validity of the Alcohol Use Disorders Identification Test (AUDIT) in a German General Practice Population Sample. *J. Stud. Alcohol Drugs* 67, 473.

Eder, A. B., and Dignath, D. (2016). Cue-elicited food seeking is eliminated with aversive outcomes following outcome devaluation. *Q. J. Exp. Psychol.* 69, 574–588. doi:10.1080/17470218.2015.1062527.

Elman, I., Becerra, L., Tschibelu, E., Yamamoto, R., George, E., and Borsook, D. (2012). Yohimbine-Induced Amygdala Activation in Pathological Gamblers: A Pilot Study. *PLOS ONE* 7, e31118. doi:10.1371/journal.pone.0031118.

Field, M., Munafò, M. R., and Franken, I. H. A. (2009). A Meta-Analytic Investigation of the Relationship Between Attentional Bias and Subjective Craving in Substance Abuse. *Psychol. Bull.* 135, 589–607. doi:10.1037/a0015843.

First, M. B., Spitzer, R. L., Gibbon, M., and Williams, J. B. (2002). Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition. *N. Y. Biom. Res. N. Y. State Psychiatr. Inst.*

Gabrieli, J. D. E., Ghosh, S. S., and Whitfield-Gabrieli, S. (2015). Prediction as a Humanitarian and Pragmatic Contribution from Human Cognitive Neuroscience. *Neuron* 85, 11–26. doi:10.1016/j.neuron.2014.10.047.

Garbusow, M., Schad, D. J., Sebold, M., Friedel, E., Bernhardt, N., Koch, S. P., et al. (2016). Pavlovian-to-instrumental transfer effects in the nucleus accumbens relate to relapse in alcohol dependence. *Addict. Biol.* 21, 719–731. doi:10.1111/adb.12243.

Garofalo, S., and Robbins, T. W. (2017). Triggering Avoidance: Dissociable Influences of Aversive Pavlovian Conditioned Stimuli on Human Instrumental Behavior. *Front. Behav. Neurosci.* 11. doi:10.3389/fnbeh.2017.00063.

Genauck, A., Andrejevic, M., Brehm, K., Matthis, C., Heinz, A., Weinreich, A., et al. (2019). Cue-induced effects on decision-making distinguish subjects with gambling disorder from healthy controls. *bioRxiv*, 564781. doi:10.1101/564781.

Genauck, A., Huys, Q. J., Heinz, A., and Rapp, M. A. (2013). Pawlowsch-Instrumentelle Transfereffekte bei Alkoholabhängigkeit. *SUCHT-Z. Für Wiss. PraxisJournal Addict. Res. Pract.* 59, 215–223.

Genauck, A., Quester, S., Wüstenberg, T., Mörsen, C., Heinz, A., and Romanczuk-Seiferth, N. (2017). Reduced loss aversion in pathological gambling and alcohol dependence is associated with differential alterations in amygdala and prefrontal functioning. *Sci. Rep.* 7, 16306. doi:10.1038/s41598-017-16433-y.

Glimcher, P. W., and Rustichini, A. (2004). Neuroeconomics: the consilience of brain and decision. *Science* 306, 447–452. doi:10.1126/science.1102566.

Goudriaan, A. E., de Ruiter, M. B., van den Brink, W., Oosterlaan, J., and Veltman, D. J. (2010). Brain activation patterns associated with cue reactivity and craving in abstinent problem gamblers, heavy smokers and healthy controls: an fMRI study. *Addict. Biol.* 15, 491–503. doi:10.1111/j.1369-1600.2010.00242.x.

Guggenmos, M., Scheel, M., Sekutowicz, M., Garbusow, M., Sebold, M., Sommer, C., et al. (2018). Decoding diagnosis and lifetime consumption in alcohol dependence from grey-matter pattern information. *Acta Psychiatr. Scand.* 137, 252–262. doi:10.1111/acps.12848.

Guitart-Masip, M., Talmi, D., and Dolan, R. (2010). Conditioned associations and economic decision biases. *Neuroimage* 53, 206–214. doi:10.1016/j.neuroimage.2010.06.021.

Hall, J., Parkinson, J. A., Connor, T. M., Dickinson, A., and Everitt, B. J. (2001). Involvement of the central nucleus of the amygdala and nucleus accumbens core in mediating Pavlovian influences on instrumental behaviour. *Eur. J. Neurosci.* 13, 1984–1992. doi:10.1046/j.0953-816x.2001.01577.x.

Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., et al. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage* 87, 96–110. doi:10.1016/j.neuroimage.2013.10.067.

Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., and Fagerström, K. O. (1991). The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. *Br. J. Addict.* 86, 1119–1127.

Heinz, A. (2017). *A New Understanding of Mental Disorders: Computational Models for Dimensional Psychiatry*. MIT Press.

Heinz, A., Löber, S., Georgi, A., Wrase, J., Hermann, D., Rey, E.-R., et al. (2003). Reward Craving and Withdrawal Relief Craving: Assessment of Different Motivational Pathways to Alcohol Intake. *Alcohol Alcohol* 38, 35–39. doi:10.1093/alcalc/agg005.

Hogarth, L., and Chase, H. W. (2012). Evaluating psychological markers for human nicotine dependence: Tobacco choice, extinction, and Pavlovian-to-instrumental transfer. *Exp. Clin. Psychopharmacol.* 20, 213.

Holmes, N. M., Marchand, A. R., and Coutureau, E. (2010). Pavlovian to instrumental transfer: A neurobehavioural perspective. *Neurosci. Biobehav. Rev.* 34, 1277–1295. doi:10.1016/j.neubiorev.2010.03.007.

Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econom. J. Econom. Soc.*, 263–291.

Koehler, S., Hasselmann, E., Wüstenberg, T., Heinz, A., and Romanczuk-Seiferth, N. (2015). Higher volume of ventral striatum and right prefrontal cortex in pathological gambling. *Brain Struct. Funct.* 220, 469–477. doi:10.1007/s00429-013-0668-6.

Koehler, S., Ovadia-Caro, S., Meer, E. van der, Villringer, A., Heinz, A., Romanczuk-Seiferth, N., et al. (2013). Increased Functional Connectivity between Prefrontal Cortex and Reward System in Pathological Gambling. *PLOS ONE* 8, e84565. doi:10.1371/journal.pone.0084565.

Krank, M. D., O'Neill, S., Squarey, K., and Jacob, J. (2008). Goal- and signal-directed incentive: conditioned approach, seeking, and consumption established with unsweetened alcohol in rats. *Psychopharmacology (Berl.)* 196, 397–405. doi:10.1007/s00213-007-0971-0.

Kühn, S., and Gallinat, J. (2011). Common biology of craving across legal and illegal drugs – a quantitative meta-analysis of cue-reactivity brain response. *Eur. J. Neurosci.* 33, 1318–1326. doi:10.1111/j.1460-9568.2010.07590.x.

Leeman, R. F., and Potenza, M. N. (2012). Similarities and Differences between Pathological Gambling and Substance Use Disorders: A Focus on Impulsivity and Compulsivity. *Psychopharmacology (Berl.)* 219, 469–490. doi:10.1007/s00213-011-2550-7.

Lesieur, H. R., and Blume, S. B. (1987). The South Oaks Gambling Screen (SOGS): a new instrument for the identification of pathological gamblers. *Am. J. Psychiatry* 144, 1184–1188. doi:10.1176/ajp.144.9.1184.

Lewis, A. H., Niznikiewicz, M. A., Delamater, A. R., and Delgado, M. R. (2013). Avoidance-based human Pavlovian-to-instrumental transfer. *Eur. J. Neurosci.* 38, 3740–3748. doi:10.1111/ejn.12377.

Leyton, M., and Vezina, P. (2012). On Cue: Striatal Ups and Downs in Addictions. *Biol. Psychiatry* 72, e21–e22. doi:10.1016/j.biopsych.2012.04.036.

Leyton, M., and Vezina, P. (2013). Striatal ups and downs: Their roles in vulnerability to addictions in humans. *Neurosci. Biobehav. Rev.* 37, 1999–2014. doi:10.1016/j.neubiorev.2013.01.018.

Li, Y., Wang, Z., Boileau, I., Dreher, J.-C., Gelskov, S., Genauck, A., et al. (2018). Altered orbitofrontal sulcogyral patterns in gambling disorder: a multicenter study. *bioRxiv*, 439034. doi:10.1101/439034.

Limbrick-Oldfield, E. H., Mick, I., Cocks, R. E., McGonigle, J., Sharman, S. P., Goldstone, A. P., et al. (2017). Neural substrates of cue reactivity and craving in gambling disorder. *Transl. Psychiatry* 7, e992. doi:10.1038/tp.2016.256.

Linnet, J., Peterson, E., Doudet, D. J., Gjedde, A., and Møller, A. (2010). Dopamine release in ventral striatum of pathological gamblers losing money. *Acta Psychiatr. Scand.* 122, 326–333. doi:10.1111/j.1600-0447.2010.01591.x.

Litt, A., Eliasmith, C., and Thagard, P. (2008). Neural affective decision theory: Choices, brains, and emotions. *Cogn. Syst. Res.* 9, 252–273. doi:10.1016/j.cogsys.2007.11.001.

Lorains, F. K., Dowling, N. A., Enticott, P. G., Bradshaw, J. L., Trueblood, J. S., and Stout, J. C. (2014). Strategic and non-strategic problem gamblers differ on decision-making under risk and ambiguity. *Addiction* 109, 1128–1137.

MacKillop, J., Amlung, M. T., Few, L. R., Ray, L. A., Sweet, L. H., and Munafò, M. R. (2011). Delayed reward discounting and addictive behavior: a meta-analysis. *Psychopharmacology (Berl.)* 216, 305–321. doi:10.1007/s00213-011-2229-0.

Madden, G. J., Petry, N. M., and Johnson, P. S. (2009). Pathological Gamblers Discount Probabilistic Rewards Less Steeply than Matched Controls. *Exp. Clin. Psychopharmacol.* 17, 283–290. doi:10.1037/a0016806.

McClure, S. M., Berns, G. S., and Montague, P. R. (2003). Temporal Prediction Errors in a Passive Learning Task Activate Human Striatum. *Neuron* 38, 339–346. doi:10.1016/S0896-6273(03)00154-5.

McLaren, D. G., Ries, M. L., Xu, G., and Johnson, S. C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *NeuroImage* 61, 1277–1286. doi:10.1016/j.neuroimage.2012.03.068.

Meule, A., Vögele, C., and Kübler, A. (2011). Psychometrische Evaluation der deutschen Barratt Impulsiveness Scale – Kurzversion (BIS-15). *Diagnostica* 57, 126–133. doi:10.1026/0012-1924/a000042.

Miedl, S. F., Büchel, C., and Peters, J. (2014). Cue-Induced Craving Increases Impulsivity via Changes in Striatal Value Signals in Problem Gamblers. *J. Neurosci.* 34, 4750–4755. doi:10.1523/JNEUROSCI.5020-13.2014.

Miedl, S. F., Peters, J., and Büchel, C. (2012). Altered neural reward representations in pathological gamblers revealed by delay and probability discounting. *Arch. Gen. Psychiatry* 69, 177–186. doi:10.1001/archgenpsychiatry.2011.1552.

Mitchell, S., Gao, J., Hallett, M., and Voon, V. (2016). The Role of Social Novelty in Risk Seeking and Exploratory Behavior: Implications for Addictions. *PLOS ONE* 11, e0158947. doi:10.1371/journal.pone.0158947.

Mucha, R. F., Geier, A., Stuhlinger, M., and Mundle, G. (2000). Appetitve effects of drug cues modelled by pictures of the intake ritual: generality of cue-modulated startle examined with inpatient alcoholics. *Psychopharmacology (Berl.)* 151, 428–432.

Niv, Y., Daw, N. D., Joel, D., and Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl.)* 191, 507–520.

Niv, Y., and Schoenbaum, G. (2008). Dialogues on prediction errors. *Trends Cogn. Sci.* 12, 265–272. doi:10.1016/j.tics.2008.03.006.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., and Dolan, R. J. (2003). Temporal Difference Models and Reward-Related Learning in the Human Brain. *Neuron* 38, 329–337. doi:10.1016/S0896-6273(03)00169-7.

Pariyadath, V., Stein, E. A., and Ross, T. J. (2014). Machine learning classification of resting state functional connectivity predicts smoking status. *Front. Hum. Neurosci.* 8. doi:10.3389/fnhum.2014.00425.

Patton, J. H., Stanford, M. S., and Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *J. Clin. Psychol.* 51, 768–774.

Petry, J., and Baulig, T. (1996). Kurzfragebogen zum Glücksspielverhalten. *Psychother. Glücksspielsucht Psychol. Verl. Union* 97.

Petry, N. M. (2012). Discounting of probabilistic rewards is associated with gambling abstinence in treatment-seeking pathological gamblers. *J. Abnorm. Psychol.* 121, 151–159. doi:10.1037/a0024782.

Petry, N. M., Blanco, C., Auriacombe, M., Borges, G., Bucholz, K., Crowley, T. J., et al. (2014). An overview of and rationale for changes proposed for pathological gambling in DSM-5. *J. Gambl. Stud. Co-Spons. Natl. Counc. Probl. Gambl. Inst. Study Gambl. Commer. Gaming* 30, 493–502. doi:10.1007/s10899-013-9370-0.

Platt, M. L., and Huettel, S. A. (2008). Risky business: the neuroeconomics of decision making under uncertainty. *Nat. Neurosci.* 11, 398–403. doi:10.1038/nn2062.

Potenza, M. N., Steinberg, M. A., Skudlarski, P., Fulbright, R. K., Lacadie, C. M., Wilber, M. K., et al. (2003). Gambling Urges in Pathological Gambling: A Functional Magnetic Resonance Imaging Study. *Arch. Gen. Psychiatry* 60, 828–836. doi:10.1001/archpsyc.60.8.828.

Prévost, C., Liljeholm, M., Tyszka, J. M., and O'Doherty, J. P. (2012). Neural Correlates of Specific and General Pavlovian-to-Instrumental Transfer within Human Amygdalar Subregions: A High-Resolution fMRI Study. *J. Neurosci.* 32, 8383–8390. doi:10.1523/JNEUROSCI.6237-11.2012.

Provost, F. J., Fawcett, T., and Kohavi, R. (1998). The case against accuracy estimation for comparing induction algorithms. in *ICML*, 445–453.

Reuter, J., Raedler, T., Rose, M., Hand, I., Gläscher, J., and Büchel, C. (2005). Pathological gambling is linked to reduced activation of the mesolimbic reward system. *Nat. Neurosci.* 8, 147–148. doi:10.1038/nn1378.

Rodríguez-Testal, J. F., Senín-Calderón, C., and Perona-Garcelán, S. (2014). From DSM-IV-TR to DSM-5: Analysis of some changes. *Int. J. Clin. Health Psychol.* 14, 221–231.

Romanczuk-Seiferth, N., Koehler, S., Dreesen, C., Wüstenberg, T., and Heinz, A. (2015). Pathological gambling and alcohol dependence: neural disturbances in reward and loss avoidance processing. *Addict. Biol.* 20, 557–569. doi:10.1111/adb.12144.

Romanczuk-Seiferth, N., van den Brink, W., and Goudriaan, A. E. (2014). From Symptoms to Neurobiology: Pathological Gambling in the Light of the New Classification in DSM-5. *Neuropsychobiology* 70, 95–102. doi:10.1159/000362839.

Rømer Thomsen, K., Fjorback, L. O., Møller, A., and Lou, H. C. (2014). Applying incentive sensitization models to behavioral addiction. *Neurosci. Biobehav. Rev.* 45, 343–349. doi:10.1016/j.neubiorev.2014.07.009.

Saddoris, M. P., Stamatakis, A., and Carelli, R. M. (2011). Neural correlates of Pavlovian-to-instrumental transfer in the nucleus accumbens shell are selectively potentiated following cocaine self-administration. *Eur. J. Neurosci.* 33, 2274–2287. doi:10.1111/j.1460-9568.2011.07683.x.

Schacht, J. P., Anton, R. F., and Myrick, H. (2013). Functional neuroimaging studies of alcohol cue reactivity: a quantitative meta-analysis and systematic review. *Addict. Biol.* 18, 121–133. doi:10.1111/j.1369-1600.2012.00464.x.

Schad, D. J., Garbusow, M., Friedel, E., Sommer, C., Sebold, M., Hägele, C., et al. (2018). Neural correlates of instrumental responding in the context of alcohol-related cues index disorder severity and relapse risk. *Eur. Arch. Psychiatry Clin. Neurosci.*, 1–14. doi:10.1007/s00406-017-0860-4.

Schulreich, S., Gerhardt, H., and Heekeren, H. R. (2016). Incidental fear cues increase monetary loss aversion. *Emot. Wash. DC* 16, 402–412. doi:10.1037/emo0000124.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Sebold, M., Schad, D. J., Nebe, S., Garbusow, M., Jünger, E., Kroemer, N. B., et al. (2016). Don't Think, Just Feel the Music: Individuals with Strong Pavlovian-to-Instrumental Transfer Effects Rely Less on Model-based Reinforcement Learning. *J. Cogn. Neurosci.* 28, 985–995. doi:10.1162/jocn_a_00945.

Seo, S., Beck, A., Matthis, C., Genauck, A., Banaschewski, T., Bokde, A. L. W., et al. (2018). Risk profiles for heavy drinking in adolescence: differential effects of gender. *Addict. Biol.* doi:10.1111/adb.12636.

Seo, S., Mohr, J., Beck, A., Wüstenberg, T., Heinz, A., and Obermayer, K. (2015). Predicting the future relapse of alcohol-dependent patients from structural and functional brain images. *Addict. Biol.* 20, 1042–1055. doi:10.1111/adb.12302.

Sescousse, G., Barbalat, G., Domenech, P., and Dreher, J.-C. (2013). Imbalance in the sensitivity to different types of rewards in pathological gambling. *Brain*, awt126. doi:10.1093/brain/awt126.

Sommer, C., Garbusow, M., Jünger, E., Pooseh, S., Bernhardt, N., Birkenstock, J., et al. (2017). Strong seduction: impulsivity and the impact of contextual cues on instrumental behavior in alcohol dependence. *Transl. Psychiatry* 7, e1183. doi:10.1038/tp.2017.158.

Steenbergen, H. van, Watson, P., Wiers, R. W., Hommel, B., and Wit, S. de (2017). Dissociable corticostriatal circuits underlie goal-directed vs. cue-elicited habitual food seeking after satiation: evidence from a multimodal MRI study. *Eur. J. Neurosci.* 46, 1815–1827. doi:10.1111/ejn.13586.

Steenbergh, T. A., Meyers, A. W., May, R. K., and Whelan, J. P. (2002). Development and validation of the Gamblers' Beliefs Questionnaire. *Psychol. Addict. Behav. J. Soc. Psychol. Addict. Behav.* 16, 143–149.

Stinchfield, R. (2002). Reliability, validity, and classification accuracy of the South Oaks Gambling Screen (SOGS). *Addict. Behav.* 27, 1–19.

Takeuchi, H., Tsurumi, K., Murao, T., Mizuta, H., Murai, T., and Takahashi, H. Amygdala volume is associated with risky probability cognition in gambling disorder. *Addict. Biol.* 0. doi:10.1111/adb.12640.

Talmi, D., Seymour, B., Dayan, P., and Dolan, R. J. (2008). Human Pavlovian–Instrumental Transfer. *J. Neurosci.* 28, 360–368. doi:10.1523/JNEUROSCI.4028-07.2008.

Tom, S. M., Fox, C. R., Trepel, C., and Poldrack, R. A. (2007). The Neural Basis of Loss Aversion in Decision-Making Under Risk. *Science* 315, 515–518. doi:10.1126/science.1134239.

Tversky, A., and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *J. Risk Uncertain.* 5, 297–323.

van Holst, R. J., van Holstein, M., van den Brink, W., Veltman, D. J., and Goudriaan, A. E. (2012). Response Inhibition during Cue Reactivity in Problem Gamblers: An fMRI Study. *PLoS ONE* 7, e30909. doi:10.1371/journal.pone.0030909.

Varma, S., and Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* 7, 91. doi:10.1186/1471-2105-7-91.

Vezina, P., and Leyton, M. (2009). Conditioned cues and the expression of stimulant sensitization in animals and humans. *Neuropharmacology* 56, Supplement 1, 160–168. doi:10.1016/j.neuropharm.2008.06.070.

Whelan, R., Watts, R., Orr, C. A., Althoff, R. R., Artiges, E., Banaschewski, T., et al. (2014). Neuropsychosocial profiles of current and future adolescent alcohol misusers. *Nature* 512, 185–189. doi:10.1038/nature13402.

Wiehler, A., and Peters, J. (2015). Reward-based decision making in pathological gambling: The roles of risk and delay. *Neurosci. Res.* 90, 3–14. doi:10.1016/j.neures.2014.09.008.

Wölfling, K., Mörsen, C. P., Duven, E., Albrecht, U., Grüsser, S. M., and Flor, H. (2011). To gamble or not to gamble: at risk for craving and relapse--learned motivated attention in pathological gambling. *Biol. Psychol.* 87, 275–281. doi:10.1016/j.biopsycho.2011.03.010.

Yarkoni, T., and Westfall, J. (2017). Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspect. Psychol. Sci.* 12, 1100–1122. doi:10.1177/1745691617693393.

Zacharaki, E. I., Wang, S., Chawla, S., Yoo, D. S., Wolf, R., Melhem, E. R., et al. (2009). Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. *Magn. Reson. Med.* 62, 1609–1618. doi:10.1002/mrm.22147.

Zou, H., and Hastie, T. (2005). Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 67, 301–320.