Biological Psychology 127 (2017) 1–9

Contents lists available at ScienceDirect

Biological Psychology

journal homepage: www.elsevier.com/locate/biopsycho

Striatal dopaminergic modulation of reinforcement learning predicts reward—oriented behavior in daily life

Zuzana Kasanova[a,*], Jenny Ceccarini[b], Michael J. Frank[c], Thérèse van Amelsvoort[d], Jan Booij[e], Alexander Heinzel[f], Felix Mottaghy[f], Inez Myin-Germeys[a]

[a] Center for Contextual Psychiatry, Department of Neuroscience, KU Leuven – Leuven University, Leuven, Belgium
[b] Division of Nuclear Medicine and Molecular Imaging, Department of Imaging & Pathology, University Hospitals Leuven, Leuven, Belgium
[c] Department of Cognitive, Linguistic and Psychological Sciences, Brown University, Providence, USA
[d] Department of Psychiatry and Neuropsychology, Maastricht University, Maastricht, The Netherlands
[e] Department of Nuclear Medicine, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
[f] Department of Nuclear Medicine, University Hospital RWTH Aachen University, Aachen, Germany

ARTICLE INFO

Keywords:
Dopamine
Reward
Reinforcement learning
PET
Ecological momentary assessments
Motivation

ABSTRACT

Much human behavior is driven by rewards. Preclinical neurophysiological and clinical positron emission tomography (PET) studies have implicated striatal phasic dopamine (DA) release as a primary modulator of reward processing. However, the relationship between experimental reward-induced striatal DA release and responsiveness to naturalistic rewards, and therefore functional relevance of these findings, has been elusive.

We therefore combined, for the first time, a DA $D_{2/3}$ receptor [^{18}F]fallypride PET during a probabilistic reinforcement learning (RL) task with a six day ecological momentary assessments (EMA) of reward-related behavior in the everyday life of 16 healthy volunteers. We detected significant reward-induced DA release in the bilateral putamen, caudate nucleus and ventral striatum, the extent of which was associated with better behavioral performance on the RL task across all regions. Furthermore, individual variability in the extent of reward-induced DA release in the right caudate nucleus and ventral striatum modulated the tendency to be actively engaged in a behavior if the active engagement was previously deemed enjoyable. This study suggests a link between striatal reward-related DA release and ecologically relevant reward-oriented behavior, suggesting an avenue for the inquiry into the DAergic basis of optimal and impaired motivational drive.

1. Introduction

Rewards are those stimuli or affective states that elicit approach behavior, increase frequency of such behavior, and thus maintain motivated action and adaptive learning (Schultz, 2010). For instance, we investigated the components of reward-oriented behavior in the everyday life of the general population, and demonstrated that positive affect experienced during physical or social activities significantly increased the odds of engaging in similar activities in the near future (Wichers et al., 2015). Importantly, deviations from the normative reward-oriented behavior can result in addiction on the one extreme (Volkow et al., 2010), and motivational deficits of depression or schizophrenia (Barch, Pagliaccio, & Luking, 2016) on the other.

Advances in molecular neuroscience allowed to examine the neurochemical modulators of reward-directed behavior and have brought the mesolimbic dopamine (DA) system to the fore. In vivo microdialysis studies revealed that DA levels were increased in the striatum upon lever pressing for rewards (e.g. food) (Schultz, Dayan, & Montague, 1997;Salamone, Cousins, McCullough, Carriero, & Berkowitz, 1994), occasioning reward learning. Corroborating positron emission tomography (PET) studies of DA $D_{2/3}$ receptor binding, with [^{11}C]raclopride displacement as an index of DA release (Laruelle, 2000) in the human striatum, showed increased DA signaling during active reward learning condition (Zald et al., 2004; Weiland et al., 2014). Meanwhile, passive reward delivery failed to evoke changes in baseline DA firing (Hakyemez, Dagher, Smith, & Zald, 2008), confirming that striatal DAergic phasic firing modulates reward learning by representing the imminent reward initiated by a cue.

In nature, however, the probability that a reward-approach behavior will result in a reward varies. The striatum has been shown to be exquisitely sensitive to the violation of the predicted outcome of a behavior. In experimental animals, unexpected reward delivery (positive prediction error) evokes DA bursts, while unexpected reward omission (negative prediction error) elicits DA dips (Schultz, 2010;

* Corresponding author.
E-mail address: zuzana.kasanova@kuleuven.be (Z. Kasanova).

http://dx.doi.org/10.1016/j.biopsycho.2017.04.014
Received 7 March 2017; Received in revised form 12 April 2017; Accepted 28 April 2017
Available online 29 April 2017

Schultz et al., 1997; Schultz, 2016). Translated to humans, increased DA release to unexpected rewards was observed in the ventral striatum (VST) (Pappata, 2002; Yoder et al., 2009; Martin-Soelch et al., 2011), and medial caudate nucleus (Zald et al., 2004) of healthy volunteers. While these accounts converge on probabilistic reward-induced striatal DA release, their functional relevance in terms of associations with reward-oriented behavior remains sporadic. One study in human volunteers reported significant relationships between DA signaling change from baseline to rewarded task and reaction time in the task, and another detected that boosting DA levels using levodopa increased risky choices of potential gains, and happiness resulting from the gains (Rutledge, Skandali, Dayan, & Dolan, 2015).

While various studies have linked striatal BOLD covariates of reward prediction error to reinforcement learning, (Jonasson et al., 2014; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006), to our knowledge, no PET study to date has directly investigated the relationship between reward-induced striatal DA release and concurrent acquisition of reward contingencies in humans. Moreover, there has been little effort to establish whether either experimental measures of reinforcement learning or striatal DA release are related to daily-life reward-oriented behavior.

We therefore explored the striatal DAergic modulation of reward learning *in vivo,* and combined it with individual variability in reward-oriented behavior in the everyday life. Specifically, we performed a single day protocol [^{18}F]fallypride PET scan (Alpert, Badgaiyan, Livni, & Fischman, 2003) during an active control condition and a commonly studied probabilistic reward task designed to elicit robust reward prediction errors and the associated DAergic activity in the striatum of healthy volunteers. The same participants also underwent an ecological momentary assessment (EMA) study intended to capture the extent to which the enjoyment of being active increased the odds of being active in the near future, throughout a 6-day period. Based on the existing data, we predicted a significant increase in striatal DA release from control to reward condition. Additionally, we expected the various indices of reward function to co-vary in individuals: greater reward-induced DA release will be associated with higher reward-oriented behavior, as assessed concurrently by the probabilistic reward task, and separately by EMA in the everyday life.

2. Methods

2.1. Sample and demographics

The medical ethics committee of the Maastricht University and of the RWTH Aachen University approved the study. Approval for performing the PET study was additionally granted by the national authority for radiation protection in humans in Germany (Bundesamt für Strahlenschutz, BfS). All participants signed written informed consent before entering the study.

A total of 18 healthy volunteers were recruited to participate in this study via digital and newspaper advertisements. The inclusion criteria were age between 18 and 60 years, good general health and compliance with study procedures. The general exclusion criteria were *i*) lifetime history of Axis I or II disorders as determined by the Mini-International Neuropsychiatric Interview (M.I.N.I.) (Sheehan et al., 1998); *ii*) having a first- or second-degree relative with a diagnosed psychotic disorder or major depressive disorder as determined by the Family Interview for Genetic Studies (FIGS); *iii*) current use of neuroleptics, steroids, thyroid medication, and lifetime use of illicit hard drugs > 5 times, soft drugs > 20 times, alcohol > 7 units per week, as confirmed by the substance abuse module of the Composite International Diagnostic Interview (CIDI) (Robins et al., 1988) and by urinalysis on the day of the PET scan; *iv*) history of any neurological condition, epilepsy or head injury; *v*) non-removable metal elements in or on the body; *vi*) vision or hearing impairments affecting the performance on the task; *vii*) pregnancy, which was confirmed by a urine test on the day of the scan.

Nicotine use in all participants was ascertained using the CIDI, the IQ of the sample was estimated using the Dutch Adult Reading Test (DART), and any potential symptoms of psychopathology and psychological distress were measured using the Symptom Checklist—Revised (SCL-90-R) (Derogatis, Rickels, & Rock, 1976).

One participant was excluded based on non-compliance with the study procedures and another one due to positioning difficulties during scanning that resulted in a large part of the cerebellum missing from the images. The final analyses were thus performed on 16 participants (12 women; mean age = 38.06 years, SD = 15.61).

2.2. General procedures

Upon inclusion into the study, the demographic and lifestyle questionnaires as well as neuropsychological and symptom assessments took place. During the second session, the [^{18}F]fallypride PET scan was performed while the participants were engaged in an active control and a probabilistic stimulus selection task. On a separate occasion, the ecological momentary assessment (EMA) method (Myin-Germeys, Birchwood, & Kwapil, 2011) was explained in detail and the week of daily-life assessments commenced. Once data collection was completed, the participants were compensated with gift certificates in the value of 125 Euros, and an additional amount that they won in the reward task, which was always rounded up to 15 Euros.

2.3. Imaging data acquisition and analysis

2.3.1. General scanning procedures

On the day of the [^{18}F]fallypride PET scan, participants first received a structural Magnetic Resonance Imaging (MRI) scan (specified below), followed by the placement of a catheter into the left antecubital vein for the tracer administration. A minimal of 90-min break after catheterization was introduced to allow for any possible experience of discomfort or stress to dissipate. Then, participants were positioned on the PET scanner, a 30-inch computer monitor was adjusted approximately 100 cm in front of their eyes, and a response box with two buttons was placed on their side under the right hand to be used during the upcoming tasks. To minimize head movement during the acquisition, the head was fixated using a firm strap across the forehead. After positioning, a 10-min ^{68}Ge/^{68}Ga-transmission scan was performed, followed immediately by the radiotracer injection. At that moment, the [^{18}F]fallypride PET control condition was initiated, lasting exactly 80 min. Then, participants were removed from the PET scanner for a 15-min break. After repositioning using the localization system of the scanner, a 25-min baseline rest condition without any stimulation was completed, followed by the experimental probabilistic stimulus selection task (PSST) that was initiated exactly at 120 min post-injection, and was terminated at the end of the [^{18}F]fallypride PET dynamic acquisition at 180 min post-injection. Thereafter, the catheter was removed, and participants were debriefed and compensated for the completion of the study.

2.3.2. MRI

T1-weighted MRI scans were acquired on a Siemens 3T scanner (Siemens Healthcare. Munich, Germany) using the Magnetization Prepared Rapid Acquisition Gradient-Echo (MP-RAGE) sequence, with TE = 2.52 ms, TR = 1900 ms, matrix dimensions = 256 × 256, slice thickness = 1 mm, slice number = 176.

Tracer preparation: The radiosynthesis of [^{18}F]fallypride was a high-yield modification of the synthesis method for ^{18}F-dEMAethoxyfallypride, described in detail elsewhere (Lataster et al., 2011).

2.3.3. PET acquisition

Dynamic [^{18}F]fallypride PET measurements were performed in three-dimensional mode on a Siemens ECAT EXACT HR+ scanner (Siemens-CTY, Knoxville, TN, USA). [^{18}F]fallypride data were collected in a single

session (Alpert et al., 2003), starting immediately after a single bolus administration of [^{18}F]fallypride (mean injected dose = 185 MBq in 60 s frames during the first 6 min and 120 s frames thereafter). The first segment corresponded to the control and baseline condition and the second segment to the experimental condition (please see above in *Procedures*). Regarding the PET reconstruction, 63 slices of 2.4 mm slice thickness (pixel size = 2 × 2 mm) were reconstructed per time frame by filtered back projection (Hamming filter) after Fourier rebinning into two-dimensional sinograms. Data sets were corrected for random coincidences, scatter radiation and attenuation (10 min ^{68}Ge/^{68}Ga-transmission scan).

2.3.4. PET data analysis

For each subject, the dynamic PET images were first realigned to correct for potential effects of head movement using SPM2 (Wellcome Trust, UK). All PET processing procedures were then performed according to an automatic protocol using the PMOD brain PNEURO tool (v. 3.6, PMOD Technologies, Zurich, Switzerland). Realigned PET images were first rigidly co-registered to individual T1 MRI. Then the individual MR images were spatially normalized nonlinearly co-registered to the standard Montreal Neurological Institute (MNI) space in PMOD. Subsequently, the same was done for the PET images using the same spatial transformation as the registered MR images. For each subject, MR images were segmented into grey matter, white matter and cerebrospinal fluid within native MRI space. Automatic delineation of the left and right cerebellum (reference region, see below) and deep nuclei (including caudate nucleus [caudate], putamen and VST), was performed using the Brain Parcellation in the PMOD PNEURO tool. All normalized co-registered and segmented images were visually checked for accuracy. The fit of the delineated regions to the co-registered PET was then visually checked for accuracy, and if necessary, manually adjusted.

Subsequently, [^{18}F]fallypride PET data were analyzed using a modified simplified reference region model (SRRM), the linear extension of the SRTM (LSRRM) (Alpert et al., 2003; Badgaiyan, 2013), in accordance with previous endogenous DA displacement-type experiments (Lataster et al., 2011; Christian et al., 2006; Ceccarini et al., 2012; Hernaus et al., 2015; Kuepper et al., 2013). Reward-induced [^{18}F]fallypride displacement, reflecting DA release (Ceccarini et al., 2012), was quantified using time activity curves (TAC) and receptor kinetic parameter estimates (Alpert et al., 2003) obtained for each region of interest (ROI).

The significance of the regional DAergic activation was assessed using two kind of tests. The first approach uses the magnitude of the DA activation, and tends to detect high intensity signals (e.g. the peak height of a cluster in SPM (Friston, Frith, Liddle, & Frackowiak, 1991)). Precisely, for each individual, the LSSRM estimates the amplitude of reward-induced [^{18}F]fallypride displacement in each ROI − γ, based on the assumption that changes in competition between DA release and radioligand binding are reflected in the estimation of γ (Alpert et al., 2003), which has been previously considered an accurate approximation of stimulus-induced changes in DA release (Alpert et al., 2003; Badgaiyan, 2013; Christian et al., 2006). γ was calculated over an exponential decay function $h(t) = \exp[-t(t - T)]$, where t = measurement time, T = time of experimental condition initiation (120 min in the current activation paradigm) and τ controls the rate at which activation effects die away (dissipation rate set to $\tau = 0.03\ \text{min}^{-1}$).

However, task-induced activations do not necessarily coincide with a sharp DA peak, but can appear as more broad spatially distributed DA activation events. For this reason, the second approach is based on the spatial extent of regions defined by a simple thresholding of the statistical map, a test that is generally more sensitive to extended signals (Poline, Worsley, Evans, & Friston, 1997). In this particular case, consistent with prior [^{18}F]fallypride DAergic modulation studies (Lataster et al., 2011; Christian et al., 2006; Ceccarini et al., 2012; Hernaus et al., 2015; Kuepper et al., 2013; Vrieze et al., 2013), as an additional outcome measure of reward-induced DAergic activity to the γ parameter, the percentage of statistically significant voxels [surviving p(/number of total voxels) = 0.05] within each ROI was calculated to detect the spatial extent of the significant DA release induced by the task.

2.4. Probabilistic stimulus selection task

The experimental condition consisted of a version of a probabilistic stimulus selection task (PSST) (Pessiglione et al., 2006; Frank, Seeberger, & O'Reilly, 2004), behavioral performance which has been previously shown to be sensitive to DA manipulation, modified for PET imaging. It was administered using E-prime (Psychology Software Tools), presented on a 30-inch screen. The task was self-paced and consisted of 6 independent learning blocks. In each block 3 pairs of items below a picture of an actor were presented 40 times in a random order, for a total of 120 trials per block. Every trial started with the presentation of a picture of the actor with a neutral expression above a pair of items that illustrated the actor's hobbies (e.g. left item: basketball, right item: bicycle helmet) or profession (e.g. left item: stethoscope for medicine, right item: ruler for mathematics), depending on the block. The same actor was always presented with the same pair of items. A new set of 3 actors + pairs of items was presented in every block, requiring the participants to learn new set of contingencies (Fig. 1). The images of actors and items were selected randomly from a large pool of validated photographs and were fully counterbalanced across participants. The participants were instructed to learn which picture belonged to each actor by choosing either the left or right item (pressing either the L or R key on the response box) and receiving a feedback: the actor's smile and a win of 5 euro cents following a correct choice, and a frown and the loss of 5 euro cents after an incorrect choice. Each pair of items was associated with different probabilities of reinforcement: 90:10, 80:20 and 70:30. For instance, the choice of the correct item of the 80:20 pair led to a smile and +5 euro cents on 80% of the trials and to a frown and a −5 euro cents on 20% of trials. A tally of total money earned was always present in the middle of the screen. All participants were told beforehand that they would keep the money they earned in the task.

The performance on the PSST was primarily quantified as the total amount of money each participant won in the task. The secondary performance outcome was the average proportion of correct choices, defined as choices of the more frequently rewarded stimulus (90, 80, 70% chance of being rewarded) over its pair across all blocks .

2.5. Control task

The control task was designed to contain all features of the PSST, except for the main manipulation, the associative learning from feedback. Similar to the PSST, there were 6 blocks of 120 trials in which the participants were presented with two choice items below a photograph of an actor with a neutral expression. The two choices described some visual feature of the actor, e.g. dark/light hair, oval/long face etc. The participant was required to choose one of the items by pressing the L or R key on the response box, and wait for another one to appear, until all 18 actors were presented 40 times, lasting approximately 10 min per block. There was a 4 s inter-trial interval during which the previous image and items were still visible on the screen. No feedback and therefore no learning occurred in this task as participants were simply selecting the item that they thought described the actors better. In total, this task contained the same number of presentations of faces, choices of one of two items, and presses of the response box keys as the PSST, thus controlling for its visuomotor stimulation of the DA system.

2.6. Ecological momentary assessments

Within a week from the PET scan, each participant received an

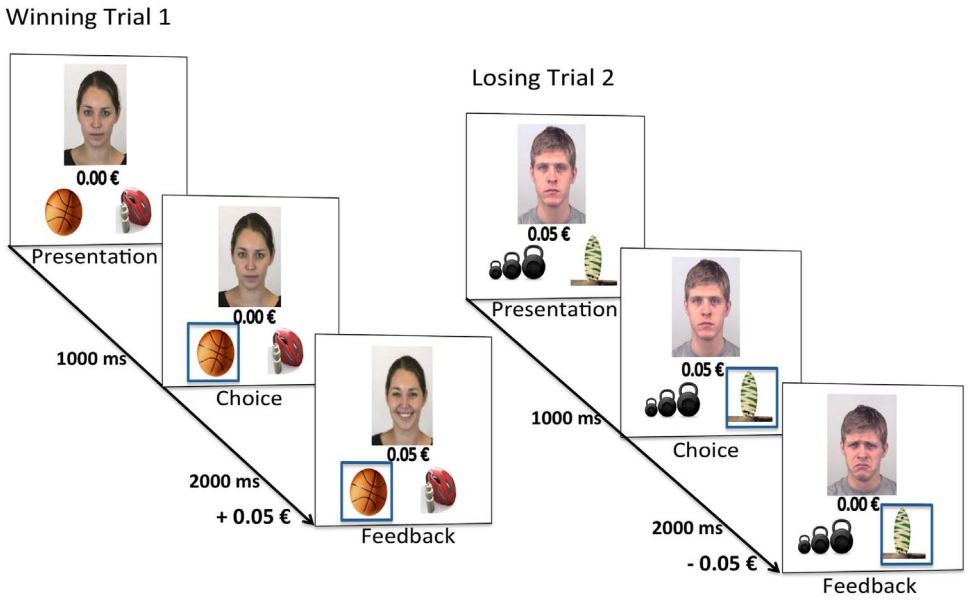

Fig. 1. Probabilistic stimulus selection task.Participants were presented with a pair of items displayed side-by-side on the screen, below a colour photograph of an actor with a neutral expression. The participants were instructed to choose the better of the two stimuli by pressing either the L or R key on the response box. Each pair of items was associated with different probabilities of reinforcement (90:10, 8:20, 70:30); on a valid trial, the choice of the correct stimulus led to a reward: the actor's smile and a win of 5 Euro cents. On an invalid trial, the choice of the correct stimulus led to a frown and the loss of 5 Euro cents.

electronic portable touch-screen device—PsyMate® and extensive training on how to use it. For the duration of the next 6 days the participants carried around the PsyMate®, which was programmed to beep 10 times per day at unexpected moments between 7:30 and 22:30. Each beep was a prompt to fill out a brief questionnaire with items appraising, among others, the current engagement in activities ("I am actively engaged in something") and the pleasantness of the current activity ("I like doing this"), rated on a 7-point Likert scale (1 = *Not at all*, 7 = *Very much*). Previous EMA research defined reward–oriented behavior as the extent to which positive affect (PA) experienced during activities at tn − 1 predicts engagement in this activity at t0. Accordingly, in the present study it was operationalized as the extent to which being actively engaged in something at t − 1 interacts with the rating of pleasantness of the activity at t − 1, in predicting being engaged in an activity at t0. In other words, reward-oriented behavior is here defined as the tendency to remain actively engaged in behavior, if said engagement was previously deemed pleasant. This association is purposefully left aspecific in that at each moment participants were asked to rate the extent to which they were actively engaged in an activity (mentally, physically or both), and how much they liked it, without specifying the activity. Therefore, the active engagement itself is the behavior of interest, and the ability of the enjoyment of active engagement to increase the odds of being actively engaged in the next moment, regardless of whether it was the same activity, is operationalized as reward-driven behavior in this context.

2.7. Statistical analyses is missing

2.7.1. Association between reward-induced tracer displacement and reward task performance

All final analyses were performed in STATA 11.2 (StataCorp, 2009). To investigate the association between reward-induced tracer displacement in all ROIs and performance on the PSST, regression analyses were conducted with total winnings as the outcome variable, and the spatial extent of reward-induced tracer displacement in each ROI as the predictor. The regression analysis was then repeated with proportion of correct choices in the task as the outcome variable.

2.7.2. Association between reward-induced tracer displacement and daily-life reward—oriented behavior

The EMA data have a hierarchical character, with up to 60 observations – beeps – nested within participants. Lagged multilevel regression models were applied to account for this structure. To assess the association between reward-induced DA release and daily life reward-oriented behavior, multilevel regression was performed with level of active engagement in behavior at t0 as the outcome variable, and active engagement at t–1 × enjoyment of the current activity at

Table 1
Demographics, psychopathology symptoms, PSST performance, EMA, Spatial extent and amplitude of reward-induced tracer displacement (i.e., DA release) per ROI.

N = 16	Mean	SD
Demographics and subclinical symptoms of psychopathology		
Age	38.06	15.61
IQ	103.75	8.14
Depression	0.07	0.13
Anxiety	0.03	0.07
Psychosis: Positive symptoms	0.27	0.20
Psychosis: Negative symptoms	0.31	0.17
Probabilistic stimulus selection task performance		
Winnings	12.46	2.89
Proportion of correct choices	0.83	0.06
Ecological Momentary Assessments		
Number of assessments (max. 60)	50.49	8.49
Active engagement in a behavior	3.80	1.05
Enjoyment of current activity	5.22	1.48
Spatial extent of reward-induced tracer displacement (% voxels)		
R Caudate	17.168	16.883
L Caudate	20.586	20.881
R Putamen	19.323	21.715
L Putamen	16.59	19.229
R Ventral striatum (VST)	23.786	27.807
L Ventral striatum (VST)	17.022	16.096
Reward-induced tracer displacement (gamma)		
R Caudate	0.00165	0.00264
L Caudate	0.00109	0.00314
R Putamen	0.00623	0.01947
L Putamen	0.00046	0.00237
R Ventral striatum (VST)	0.00099	0.00240
L Ventral striatum (VST)	0.00061	0.00247

SD = standard deviation; R = right; L = left.

t–1 × spatial extent of reward-induced tracer displacement in each ROI as the predictors. The same analyses were then repeated with the magnitude of reward-induced tracer displacement.

2.7.3. Association between reward task performance and daily-life reward-oriented behavior

To assess the association between reward task performance and daily life reward-oriented behavior, one multilevel regression was performed with level of active engagement in behavior at t0 as the outcome variable, and active engagement at t–1 × enjoyment of the current activity at t–1 × total winnings. This multilevel regression was then performed with proportion of correct choices as the predictor indexing the task performance.

To control for individual variability in age, gender, smoking status and IQ, these variables were entered into the regressions as additional predictors in all of the abovementioned analyses. In addition, to control for covariance of the observations within each participant, the structure of the matrix was set to covariance(unstructured).

3. Results

3.1. Sample demographics

The sample is described in terms of demographic, psychopathology and reward sensitivity measures in Table 1. The participants endorsed minimal levels of subclinical psychopathology that could have affected the indices of reward responsiveness.

Exploratory EMA analyses revealed high compliance rate, with all participants filling out on average 84.3% of all assessment beeps, a number that far exceeds the minimal sufficient response rate (Csikszentmihalyi and Larson, 1987).

3.2. Reward-induced subcortical dopamine release

Significant amplitude and spatial extent of reward-induced tracer displacement were detected in bilateral caudate, putamen and VST (Table 1, Fig. 2)

3.2.1. Association between reward-induced striatal DA release and reward task performance

As evidenced by Fig. 3, the group showed satisfactory performance on the task, reaching a proportion of correct choices above chance level within the first 10 trials of each pair, and surpassing the accuracy of 0.8 by the end of the task (Fig. 3).

As detailed in Table 2 and shown in Fig. 4, we detected a qualitatively positive association between reward-induced spatial extent of DA release and reward sensitivity as measured by the experimental PSST—total winnings/proportion of correct choices in all ROIs (Table 2, Fig. 4). Higher winnings in the task were significantly associated with more extensive task-induced DA release in right caudate, and a trend for significant association in left putamen (Table 2). Higher proportion of correct choices was significantly related to more extensive task-induced DA release in right caudate and bilateral putamen, with a trend emerging in left caudate and right VST (Table 2, Fig. 4). Higher winnings and proportion of correct choices in the task were statistically significantly associated with the magnitude of reward-induced DA release in the right VST only (Table 2).

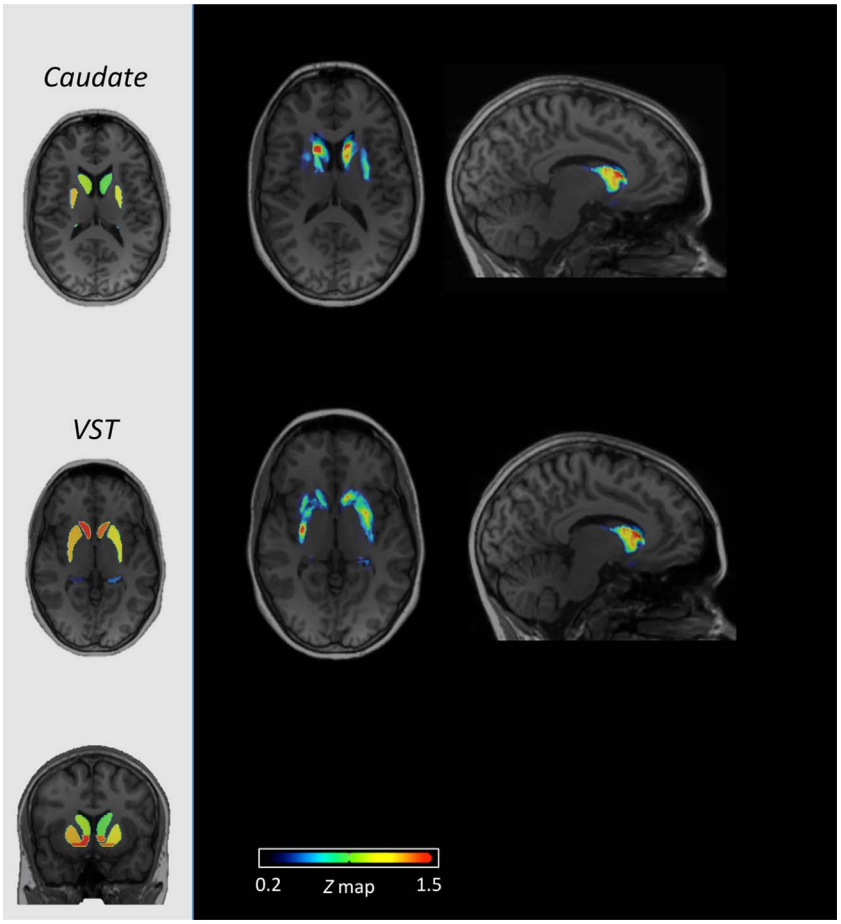

Fig. 2. Striatal reward-induced dopaminergic release.Average statistical parametric Z-map of γ representing the striatal DA release induced by the reward learning task shown in transverse, coronal, and sagittal sections overlaid on T1-weighted MRI template. The images visualize the significant reward-induced [^{18}F]fallypride displacement in the caudate and ventral striatum (VST). The left panel visualizes the masks used to delineate the caudate nucleus and ventral striatum.

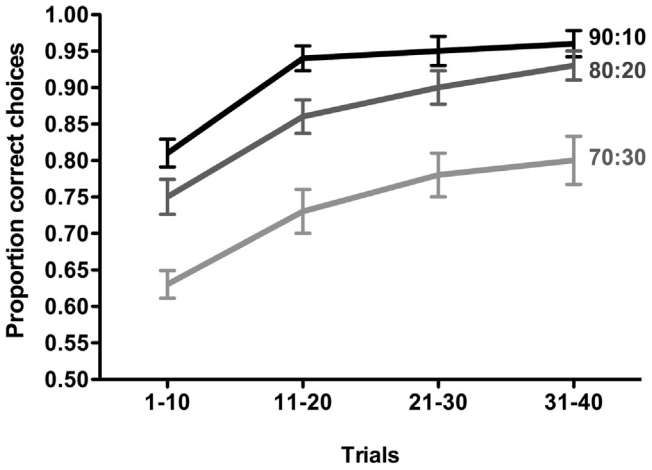

Fig. 3. Performance on the probabilistic stimulus selection task on each pair of stimuli.The y-axis presents the average accuracy on each pair of stimuli, computed per four 10-trial segments (x-axis).

3.2.2. Association between subcortical DA release and daily-life reward—oriented behavior

Exploratory EMA analyses revealed high compliance rate, with all participants filling out on average 84.3% of all assessment beeps, a number that far exceeds the minimal sufficient response rate (Csikszentmihalyi and Larson, 1987). Lagged multilevel regression analyses revealed a significant three-way interaction between the extent of reward-induced DA release in right caudate and right VST, activity engagement in behavior at t − 1 and enjoyment of that activity at t-1, in predicting the level of active engagement at t0 (Table 2). Multilevel regression analyses revealed significant positive association between reward-oriented behavior in the daily life and the extent of reward-induced DA release in the right caudate and right VST. Significant positive association between daily-life reward oriented behavior and magnitude of reward-induced DA release was detected in the right VST (Table 2).

3.2.3. Association between reward task performance and daily-life reward-oriented behavior

Lagged multilevel regressions revealed a trend for a three-way interaction between proportion of correct choices in the task × enjoyment of activity at t–1 × level of engagement in activity at t-1 in predicting the level of active engagement at t0 ($B = 0.55$, $z = 1.81$ $p = 0.07$). That is, performance in the reward task (weakly) moderated the tendency to be engaged in an activity if it was previously enjoyable to be active. The interaction between winnings in the task x enjoyment of activity at t–1 × level of engagement in activity at t − 1 in predicting the level of active engagement at t0 did not reach statistical significance ($B = 0.011$, $z = 1.16$, $p = 0.110$).

4. Discussion

The present PET study combines, for the first time, the functional molecular imaging account of DA release during reward processing and real-world reward function. Specifically, we investigated the DAergic activity during probabilistic reward learning, an essential requisite of motivated action (Huys, Pizzagalli, Bogdan, & Dayan, 2013), in combination with individual variability in reward-oriented behavior in the everyday life. Firstly, we detected DA release during the task in the dorsal and ventral striatum of healthy volunteers. Furthermore, more extensive reward-induced DA release in dorsal striatum was associated with better performance on the task, consistent with literature implicating striatal DA in probabilistic reinforcement learning (Cox et al., 2015; Deserno, Boehme, Heinz, & Schlagenhauf, 2013), but shown here for the first time in terms of task-evoked DA release. This finding mirrors animal neurophysiological literature implicating burst activity of DA neurons to reward-predictive stimuli and errors in predictions thereof (Schultz et al., 1997; Schultz, 2016), and align with computational models of DA-dependent striatal associative learning mechanisms in humans (Doll, Bath, Daw, & Frank, 2016).

Additionally, and perhaps clinically most importantly, reward-induced DA release in the right caudate and VST was found to modulate the tendency to be actively engaged in an activity if it was previously deemed enjoyable, a pattern captured throughout the everyday life of the participants. It is noteworthy that this effect was a-specific, meaning that reward-induced DA release modulated the extent to which positive experience of active engagement in any behavior increased the odds of future active engagement, regardless of whether one remained engaged in the same, or became engaged in a different activity. In terms of reinforcement contingencies, in this context the enjoyment of the activity is the reward, and the increase in active engagement in this

Table 2
Associations between reward-induced [^{18}F]fallypride displacement (i.e., DA release) in all ROIs and reward learning, measured by the probabilistic stimulus selection task, and daily-life reward—oriented behavior assessed by the ecological momentary assessments.

	Probabilistic stimulus selection task						Ecological momentary assessments		
	Winnings			Proportion correct choices			Reward-oriented behavior		
ROI	p	B	t	p	B	t	p	B	z
				Spatial extent of reward-induced tracer displacement (% voxels)					
R Caudate	0.039**	0.094	2.37	0.011**	0.0023	3.11	0.047**	0.0028	1.99
L Caudate	0.198	0.052	1.38	0.081*	0.0015	1.94	0.288	0.0012	1.06
R Putamen	0.134	0.057	1.63	0.045**	0.0016	2.29	0.563	0.0007	0.58
L Putamen	0.052*	0.083	2.21	0.016**	0.0021	2.88	0.211	0.002	1.25
R VST	0.104	0.052	1.79	0.084*	0.0012	1.92	0.014**	0.0021	2.46
L VST	0.158	0.072	1.52	0.061*	0.0020	2.11	0.325	0.0016	0.98
				Reward-induced tracer displacement (gamma)					
R Caudate	0.310	330.8	1.07	0.188	9.10	1.41	0.486	7.08	0.70
L Caudate	0.850	57.46	0.19	0.803	1.63	0.26	0.398	8.12	0.85
R Putamen	0.087*	621.4	1.90	0.150	11.52	1.56	0.575	5.68	0.56
L Putamen	0.118	580	1.71	0.185	10.78	1.42	0.770	4.45	0.29
R VST	0.027*	828.7	2.60	0.04**	16.88	2.37	0.034**	22.28	2.12
L VST	0.161	493.3	1.51	0.248	8.92	1.23	0.823	2.789	0.22

R = right, L = left; VST = ventral striatum, p = p-value, B = beta coefficient; t = t-statistic, z = z-statistic.
* Trend for significance.
** Significance at $p < 0.05$.

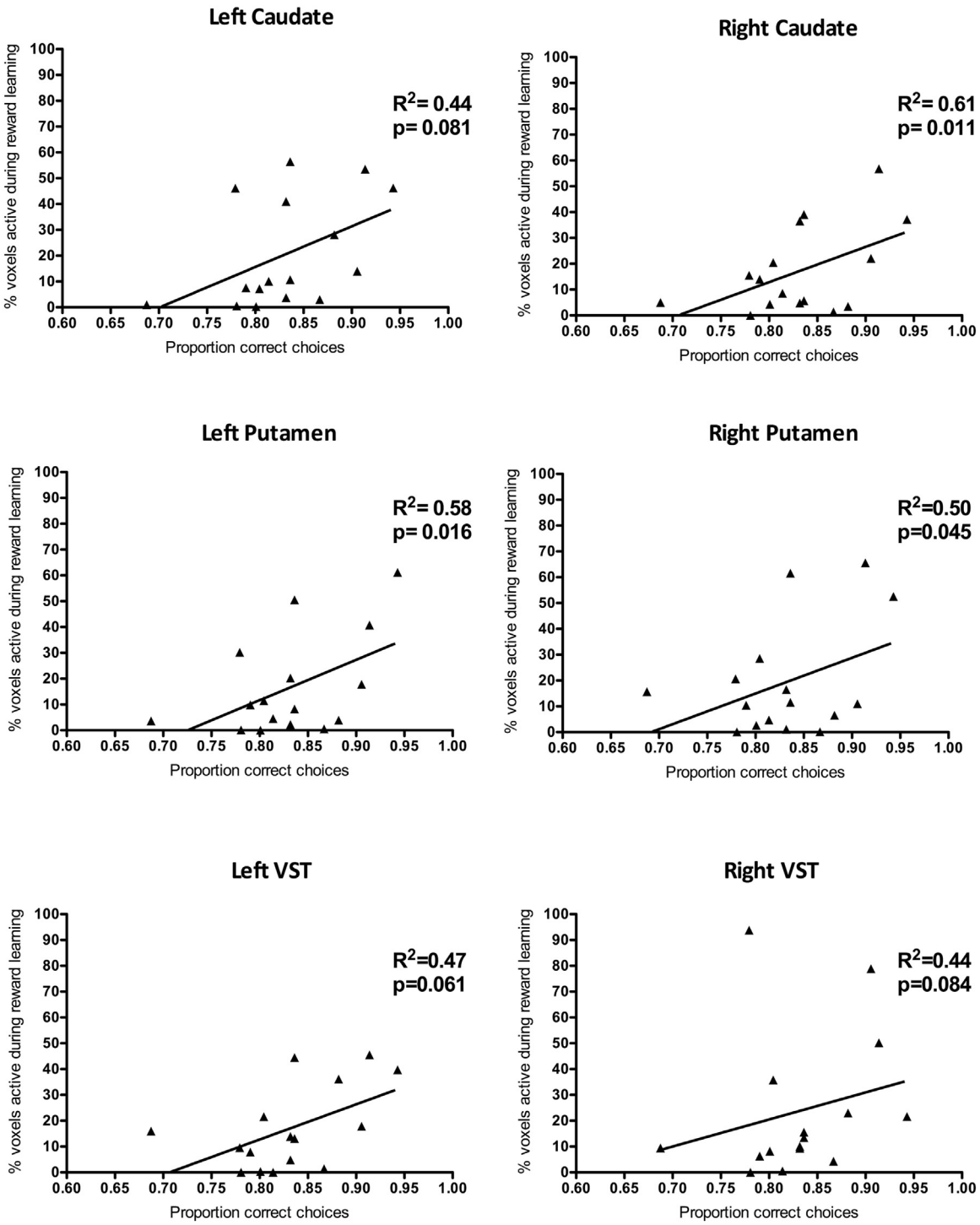

Fig. 4. Associations between spatial extent of reward-induced [^{18}F]fallypride displacement (i.e., DA release) in bilateral caudate nucleus, putamen and VST and performance on the reward task.The y-axis represents the spatial extent of reward-induced DA release, and the x-axis shows performance on reward task. The x-axis represents the proportion of correct choices of the more frequently rewarded stimulus over its alternative. The strength of the associations (R^2) and statistical significance level (p) between performance on the task and spatial extent of DA release are indicated.

or any other activity is the reinforced behavior. This lack of specificity in our assessments and analyses was deliberate, since most adults may not be free to choose their activities throughout their work day and family time, but they do have control over the level of their engagement in their tasks at hand. Two behavioral mechanisms could plausibly contribute to our findings: the enjoyment of activities in which people feel engaged into might increase the tendency to select more activities that require higher engagement (i.e. social interaction, physical activity), and/or it might increase engagement in whatever task that was already scheduled (i.e. attending a meeting, commuting). These findings build on our previous report that in a large general population sample, affective experience that was paired with physical activity and social context at previous measurements modified the likelihood to show similar behaviors at next moments (Wichers et al., 2015). Here, we offer initial evidence for the involvement of the striatal DA system in the pathway from affective experience to motivated action.

Moreover, this study elaborates on the accounts relating striatal DA function, such as DA $D_{2/3}$ receptor occupancy and binding and DA

release in response to amphetamine challenge, to behaviors such as drug craving (Wong et al., 2006), opportunistic eating (Guo, Simmons, Herscovitch, Martin, & Hall, 2014) and effort-based decision-making (Treadway et al., 2012), respectively. (Treadway et al., 2012) In the current study, however, the individual variability in a state-like striatal DA responsiveness to rewards was associated with a trait-level tendency to demonstrate behavior oriented toward laboratory as well as naturalistic rewards. It is important to note, however, that these associations were not present across the entire striatum bilaterally. This may be due to the relatively small sample size, which has been shown to be sufficient for detecting significant striatal task-induced DA release in the present and comparable studies (Zald et al., 2004; Anderson et al., 2016), but may be underpowered when lagged EMA data is incorporated into analyses.

The current results should also be interpreted with due consideration of other limitations of the study. Firstly, the assumptions of the model used to analyze the PET imaging data constrain the order of the conditions so that the control condition is always followed by the experimental condition. This design might have affected the results due to increased fatigue towards the end of the scan when the reward condition was administered. Nonetheless, we detected significant increase in striatal DA release in all ROIs during the second reward-inducing part of the scan.

Another limitation pertains to the inclusion of habitual nicotine users into the study because it could potentially affect DAergic transmission in general and reward-induced DA release in particular (Mansvelder and McGehee, 2000). We attempted to minimize its impact on the results while maintaining a sample that is representative of the general population by controlling for smoking status in all analyses, and asking participants to refrain from smoking on the day of the PET scan.

In conclusion, the present PET study ties the neurochemical index of reward responsiveness to its behavioral counterpart, thus elucidating its functional relevance. The current results also confirm that [^{18}F]fallypride PET is suitable for measurements of task-dependent changes in striatal DA release, and integrations with its pertinent behaviors as they unfold in the everyday life. Decidedly, the ability to study DA release associated with ecologically relevant rewards is essential in order to elucidate the dopaminergic basis of goal-oriented behavior and its role in motivational impairments.

Acknowledgements

This work was supported by an ERC consolidator grant to Prof. Dr. Inez Myin-Germeys (ERC-2012-StG, project 309767—INTERACT), and by the Research Foundation—Flanders (FWO) postdoctoral fellowship to Dr. Jenny Ceccarini. The authors thank Rayyan Tutunji, Nele Soons, Dr. Siamak Mohammadkhani Shali, Dr. Ye Rong, Dr. Oliver Winz, Wendy Beuken, Bernward Oedekoven and Ron Mengelers.

References

Alpert, N. M., Badgaiyan, R. D., Livni, E., & Fischman, A. J. (2003). A novel method for noninvasive detection of neuromodulatory changes in specific neurotransmitter systems. *Neuroimage*, *19*(July (3)), 1049–1060.

Anderson, B. A., Kuwabara, H., Wong, D. F., Gean, E. G., Rahmim, A., BrasićJ, R., et al. (2016). The role of dopamine in value-based attentional orienting. *Current Biology*, *26*(4), 550–555.

Badgaiyan, R. D. (2013). Detection of dopamine neurotransmission in real time. *Frontiers in Neuroscience*, *7*, 125.

Barch, D. M., Pagliaccio, D., & Luking, K. (2016). Mechanisms underlying motivational deficits in psychopathology: similarities and differences in depression and schizophrenia. *Current Topics in Behavioral Neurosciences*, *27*, 411–449.

Ceccarini, J., Vrieze, E., Koole, M., Muylle, T., Bormans, G., Claes, S., et al. (2012). Optimized in vivo detection of dopamine release using 18 F-fallypride PET. *Journal of Nuclear Medicine: Official Publication Society of Nuclear Medicine*, *53*(October (10)), 1565–1572.

Christian, B. T., Lehrer, D. S., Shi, B., Narayanan, T. K., Strohmeyer, P. S., Buchsbaum, M. S., et al. (2006). Measuring dopamine neuromodulation in the thalamus: using F-18 fallypride PET to study dopamine release during a spatial attention task. *Neuroimage*, *31*(May (1)), 139–152.

Cox, S. M., Frank, M. J., Larcher, K., Fellows, L. K., Clark, C. A., Leyton, M., et al. (2015). Striatal D1 and D2 signaling differentially predict learning from positive and negative outcomes. *Neuroimage*, *109*, 95–101.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *The Journal of Nervous and Mental Disease*, *175*(September (9)), 526–536.

Derogatis, L. R., Rickels, K., & Rock, A. F. (1976). The SCL-90 and the MMPI: a step in the validation of a new self-report scale. *British Journal of Psychiatry*, *128*, 280–289.

Deserno, L., Boehme, R., Heinz, A., & Schlagenhauf, F. (2013). Reinforcement learning and dopamine in schizophrenia: dimensions of symptoms or specific features of a disease group? *Frontiers in Psychiatry*, *4*, 172.

Doll, B. B., Bath, K. G., Daw, N. D., & Frank, M. J. (2016). Variability in dopamine genes dissociates model-based and model-free reinforcement learning. *Journal of Neuroscience*, *36*(4), 1211–1222.

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science*, *306*(5703), 1940–1943.

Friston, K. J., Frith, C. D., Liddle, P. F., & Frackowiak, R. S. (1991). Comparing functional (PET) images: the assessment of significant change. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, *11*(July (4)), 690–699.

Guo, J., Simmons, W. K., Herscovitch, P., Martin, A., & Hall, K. D. (2014). Striatal dopamine D2-like receptor correlation patterns with human obesity and opportunistic eating behavior. *Molecular Psychiatry*, *19*(October (10)), 1078–1084.

Hakyemez, H. S., Dagher, A., Smith, S. D., & Zald, D. H. (2008). Striatal dopamine transmission in healthy humans during a passive monetary reward task. *Neuroimage*, *39*(4), 2058–2065.

Hernaus, D., Collip, D., Kasanova, Z., Winz, O., Heinzel, A., van Amelsvoort, T., et al. (2015). No evidence for attenuated stress-induced extrastriatal dopamine signaling in psychotic disorder. *Translational Psychiatry*, *5*, e547.

Huys, Q. J., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biology of Mood & Anxiety Disorders*, *3*(1), 12.

Jonasson, L. S., Axelsson, J., Riklund, K., Braver, T. S., Ogren, M., Backman, L., et al. (2014). Dopamine release in nucleus accumbens during rewarded task switching measured by (1)(1)Craclopride. *Neuroimage*, *99*, 357–364.

Kuepper, R., Ceccarini, J., Lataster, J., van Os, J., van Kroonenburgh, M., van Gerven, J. M., et al. (2013). Delta-9-tetrahydrocannabinol-induced dopamine release as a function of psychosis risk: 18 F-fallypride positron emission tomography study. *PLoS One*, *8*(7), e70378.

Laruelle, M. (2000). Imaging synaptic neurotransmission with in vivo binding competition techniques: a critical review. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, *20*(March (3)), 423–451.

Lataster, J., Collip, D., Ceccarini, J., Haas, D., Booij, L., van Os, J., et al. (2011). Psychosocial stress is associated with in vivo dopamine release in human ventromedial prefrontal cortex: a positron emission tomography study using (1)(8) Ffallypride. *Neuroimage*, *58*(4), 1081–1089.

Mansvelder, H. D., & McGehee, D. S. (2000). Long-term potentiation of excitatory inputs to brain reward areas by nicotine. *Neuron*, *27*(August (2)), 349–357.

Martin-Soelch, C., Szczepanik, J., Nugent, A., Barhaghi, K., Rallis, D., Herscovitch, P., et al. (2011). Lateralization and gender differences in the dopaminergic response to unpredictable reward in the human ventral striatum. *European Journal of Neuroscience*, *33*(May (9)), 1706–1715.

Myin-Germeys, I., Birchwood, M., & Kwapil, T. (2011). From environment to therapy in psychosis: a real-world momentary assessment approach. *Schizophrenia Bulletin*, *37*(March (2)), 244–247.

Pappata, S. (2002). In vivo detection of striatal dopamine release during reward: a PET study with 11 CRaclopride and a single dynamic scan approach. *NeuroImage*, *16*(4), 1015–1027.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042–1045.

Poline, J. B., Worsley, K. J., Evans, A. C., & Friston, K. J. (1997). Combining spatial extent and peak intensity to test for activations in functional imaging. *Neuroimage*, *5*(February (2)), 83–96.

Robins, L. N., Wing, J., Wittchen, H. U., Helzer, J. E., Babor, T. F., Burke, J., et al. (1988). The Composite International Diagnostic Interview: an epidemiologic Instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry*, *45*(December (12)), 1069–1077.

Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2015). Dopaminergic modulation of decision making and subjective well-being. *Journal of Neuroscience*, *35*(27), 9811–9822.

Salamone, J. D., Cousins, M. S., McCullough, L. D., Carriero, D. L., & Berkowitz, R. J. (1994). Nucleus accumbens dopamine release increases during instrumental lever pressing for food but not free food consumption. *Pharmacology Biochemistry and Behavior*, *49*(September), 25–31.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599.

Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behavioral and Brain Functions*, *6*, 24.

Schultz, W. (2016). Dopamine reward prediction-error signalling: a two-component response. *Nature Reviews Neuroscience*, *17*(March (3)), 183–195.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., et al. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, *59*(Suppl. 20), 22–33.

Treadway, M. T., Buckholtz, J. W., Cowan, R. L., Woodward, N. D., Li, R., Ansari, M. S., et al. (2012). Dopaminergic mechanisms of individual differences in human effort-based decision-making. *Journal of Neuroscience*, *32*(18), 6170–6176.

Volkow, N. D., Wang, G. J., Fowler, J. S., Tomasi, D., Telang, F., & Baler, R. (2010). Addiction: decreased reward sensitivity and increased expectation sensitivity conspire to overwhelm the brain's control circuit. *Bioessays*, *32*(September (9)), 748–755.

Vrieze, E., Ceccarini, J., Pizzagalli, D. A., Bormans, G., Vandenbulcke, M., Demyttenaere, K., et al. (2013). Measuring extrastriatal dopamine release during a reward learning task. *Human Brain Mapping*, *34*(March), 575–586.

Weiland, B. J., Heitzeg, M. M., Zald, D., Cummiford, C., Love, T., Zucker, R. A., et al. (2014). Relationship between impulsivity, prefrontal anticipatory activation, and striatal dopamine release during rewarded task performance. *Psychiatry Research*, *223*(3), 244–252.

Wichers, M., Kasanova, Z., Bakker, J., Thiery, E., Derom, C., Jacobs, N., et al. (2015). From affective experience to motivated action: tracking reward-seeking and punishment-avoidant behaviour in real-life. *PLoS One*, *10*(6), e0129722.

Wong, D. F., Kuwabara, H., Schretlen, D. J., Bonson, K. R., Zhou, Y., Nandi, A., et al. (2006). Increased occupancy of dopamine receptors in human striatum during cue-elicited cocaine craving. *Neuropsychopharmacology*, *31*(December (12)), 2716–2727.

Yoder, K. K., Morris, E. D., Constantinescu, C. C., Cheng, T. E., Normandin, M. D., O'Connor, S. J., et al. (2009). When what you see isn't what you get: alcohol cues, alcohol administration, prediction error, and human striatal dopamine. *Alcoholism: Clinical and Experimental Research*, *33*(1), 139–149.

Zald, D. H., Boileau, I., El-Dearedy, W., Gunn, R., McGlone, F., Dichter, G. S., et al. (2004). Dopamine transmission in the human striatum during monetary reward tasks. *Journal of Neuroscience*, *24*(17), 4105–4112.