Supplementary Materials and Methods

Procedure

Encoding session

The participants encoded 80 unfamiliar photos of outdoor scenes (Peelen *et al*, 2009) which were presented for 800 ms (Clos *et al*, 2015) on the first day of the study (no medication given). The participants had to indicate whether the picture contained cars or people by pressing one of two buttons on the keyboard (Fig. 1 A). Each picture was followed by an active baseline condition (ISI 8-12 s, arrow pointing task) in which the direction of arrows presented for 800 ms had to be indicated by button press.

Working memory testing

Working memory (WM) was tested on a computer on the first day of the study prior to the encoding task (T0 baseline measurement) as well as under drug following the scanning session (T1). We measured WM performance on a self-paced complex span task (Unsworth *et al*, 2009) where the location of 2-5 squares within a 16-square pattern had to be remembered while simultaneously judging the symmetry of abstract pictures (symmetric/asymmetric). Subsequently, the participants performed a self-paced digit and block span task (Kessels *et al*, 2008). Digit sequences presented auditorily via headphones had to be entered via the keyboard after presentation of the last digit in either forward or backward order. For the forward digit span, the sequence increased from 3 digits to maximally 8 digits (depending on performance). For the backward digit span, the sequence increased from 2 digit to maximally 7 digits. Similarly, during trials of block span task, the participants saw a pattern of objects, whose location and (forward/backward) appearance order they had subsequently to indicate via mouse clicks on these objects. For the forward block span, the sequence increased from 2 objects to maximally 8 objects. For the backward block span, the sequence increased from 2 to maximally 7 objects.

Side and mood effects

All participants filled out questionnaires assessing their current mood and potential adverse effects of the medication prior to drug administration and twice during the waiting period as well as once after the scanning session (30 minutes, 2 hours and 4.5 hours post-drug). Additionally, blood pressure and pulse were controlled by a physician at these four time points.

MR Image acquisition and pre-processing

One placebo dataset was acquired only behaviorally due to scanner problems. For the remaining 53 participants, functional MR images were obtained during recognition on a 3T system Siemens Trio using single-shot echo-planar imaging with parallel imaging (GRAPPA(Griswold *et al*, 2002), in-plane acceleration factor 2) and simultaneous multi-slice acquisitions(Feinberg *et al*, 2010; Moeller *et al*, 2010; Setsompop *et al*, 2012; Xu *et al*, 2013) ("multiband", slice acceleration factor 2; TR = 1.98s, TE = 26ms, flip angle = 70°, 64 axial slices, voxel size 2 x 2 x 2 mm^3). The corresponding image reconstruction algorithm was provided by the University of Minnesota Center for Magnetic Resonance Research. In addition, an anatomical high-resolution T1-weighted image (TR = 2.3s, TE = 2.98ms, flip angle = 9°, 192 sagittal slices, voxel size 1 x 1 x 1 mm^3) and an anatomical magnetization transfer (MT) image (TR = 14ms, TE = 3.2ms, flip angle = 6°, 240 coronal slices, voxel size 1 x 1 x 1 mm^3) was acquired for each participant.

The data were pre-processed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/). The first five EPI images were discarded to allow for magnetic-field saturation. EPI images were corrected for motion and for the interaction between motion and distortion (using the unwarping procedure). Anatomical T1-weighted images were normalized to standard MNI space using DARTEL normalization. Subsequently, the EPI images and the MT image were co-registered with the normalized T1 image and the DARTEL normalization parameters were applied to the EPI images and the MT image. Finally, these normalized EPI images were spatially smoothed with a Gaussian kernel of 8 mm full width at half-maximum.

Behavioral data analysis

All behavioral data were analyzed using SPSS 21.0.0 (SPSS; Chicago, IL) and Matlab R2013a (The MathWorks; Natick, MA). Post-hoc tests were Bonferroni-corrected where appropriate and corrected p-values are reported.

Side effects and mood

The scores on the adverse medication effects questionnaire were summed together per measurement time point and analyzed relative to baseline for group differences using a repeated measures ANOVA with the factors group and time. Similarly, pulse and blood pressure measurements were analyzed relative to baseline for group differences using a repeated measures ANOVA with the factors group and time. The 16-item mood questionnaire was analyzed by reversing inverted items and log-transforming all scales before grouping the items into the three dimensions “alertness”, “calmness” and “contentment”(Bond and Lader, 1974). The resulting three dimensions were compared between the groups relative to baseline using 2 x 2 repeated measures ANOVAs with the factors group and time.

WM

Due to fatigue two haloperidol participants did not complete the post-fMRI T1 WM span tasks. Additionally, partial data loss due to technical problems affected three haloperidol and one placebo participant in the T1 WM span tasks. Working memory scores were computed for each of the WM span tests (complex span performance, block span performance forward/backward, digit span performance forward/backward) as well as for the accuracy and the RT of the symmetry rating of the complex span test. We summarized the five WM span scores into a single T0 and T1 WM span summary score using two different approaches. Firstly, z-scores of each WM span test were computed and averaged into a T0 baseline score and T1 score, respectively. Secondly, for T0 and T1 data separately, a principal components analysis (PCA) was conducted on the individual WM span scores and the resulting first component of T0 and of T1 data was used as a WM span summary score. Group differences were again examined by means of an independent t-test (T0 baseline WM span summary score) and a repeated

measures ANOVA with the factors group and time. The T0 baseline WM span summary score was moreover used as a covariate in the behavioral and fMRI analysis of memory and confidence effects to control for the potential influence of the individual dopamine baseline on medication response (Cools *et al*, 2008).

Anaylsis of fMRI data

Univariate single subject (first-level) and group (second-level) statistics were conducted using the general linear model as implemented in SPM12. On the first level, delta functions marking the picture onset trials were convolved with the canonical hemodynamic response function to create an event-related regressor for each condition. For trials with missing responses (~4% of all trials in both groups), a nuisance regressor was included in the first-level model. Low-frequency signal drifts were removed by employing a highpass filter with a cut-off period of 128 s. First, in order to test for general group differences throughout the recognition memory task, we set up a first-level model containing a regressor representing all picture onset trials. The continuous variable pleasantness was modeled as a parametric modulator. The resulting individual contrast images were compared between groups using independent-samples t-tests on the second level.

In a second first-level model, we examined effects of memory by grouping trial onsets into hit trials (correct old-responses), correct rejection trials (CR, correct new-responses), false alarm trials (FA, incorrect old-responses) and miss trials (cf. Behavioral data). For trials with missing responses (~4% of all trials in both groups), a nuisance regressor was included in the first-level model. Again, pleasantness was modeled as a parametric modulator. The corresponding four individual contrast images (hits > baseline, CR > baseline, FA > baseline, misses > baseline) of interest were fed into a second-level ANOVA. Additionally, we also set up a second-level model including pleasantness per condition. Retrieval success was evaluated by contrasting hits with CR trials using linear contrasts (Spaniol *et al*, 2009) and group differences with regard to retrieval success were evaluated by computing the group by retrieval success interactions (haloperidol hits > CR) > (placebo hits > CR) and (placebo hits > CR) >

(haloperidol hits > CR) on the second level. Moreover, we examined the data for group differences of false memory by computing the group by false memory interactions (haloperidol FA > CR) > (placebo FA > CR) and (placebo FA > CR) > (haloperidol FA > CR).

For the analysis of confidence and to simultaneously examine confidence effects per response category condition, we used a third first-level model where hit, CR, FA and miss trials were further split into high (HC), medium (MC) and low (LC) confidence trials. As before, we included a nuisance regressor for trials with missing responses and modeled pleasantness as a parametric modulator. The corresponding 12 individual contrast images (HC/MC/LC hits > baseline, HC/MC/LC CR > baseline, HC/MC/LC FA > baseline, HC/MC/LC misses > baseline) of interest were fed into a second-level ANOVA. In the second-level analysis, confidence was evaluated by contrasting HC with LC trials using linear contrasts and group differences with regard to confidence were evaluated by computing the group by confidence interactions (haloperidol HC > LC) > (placebo HC > LC) and (placebo HC > LC) > (haloperidol HC > LC).

All resulting activation maps were thresholded at $P < .05$ (family-wise error (FWE)-corrected for multiple comparisons). Given the strong a-priori hypothesis of involvement of structures with (presynaptic) dopamine receptors, we used a small volume FWE correction (SVC-FWE) at $P < .05$ based on anatomical masks (50% probability maps) of the striatum and hippocampus. The anatomical masks were created by combining the individual left and right nucleus accumbens, putamen, and caudate nucleus as well as the left and right hippocampus masks from the Harvard-Oxford subcortical structural atlases (Desikan *et al*, 2006) as implemented in FSL (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases) into one striatum and one hippocampus mask, respectively. For all other reported findings, whole-brain FWE-correction at cluster level at $P < .05$ (using a cluster-forming height threshold at voxel-level of $P < .001$ (Eklund *et al*, 2016)) was applied.

References

Bond A, Lader M (1974). The use of analogue scales in rating subjective feelings. *British Journal of Medical Psychology* **47**: 211–218.

Clos M, Schwarze U, Gluth S, Bunzeck N, Sommer T (2015). Goal- and retrieval-dependent activity in the striatum during memory recognition. *Neuropsychologia* **72**: 1–11.

Cools R, Gibbs SE, Miyakawa A, Jagust W, D'Esposito M (2008). Working memory capacity predicts dopamine synthesis capacity in the human striatum. *J Neurosci* **28**: 1208–1212.

Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, *et al* (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**: 968–980.

Eklund A, Nichols TE, Knutsson H (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci USA* **113**: 7900–7905.

Feinberg DA, Moeller S, Smith SM, Auerbach E, Ramanna S, Gunther M, *et al* (2010). Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. *PLoS ONE* **5**: e15710.

Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, *et al* (2002). Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn Reson Med* **47**: 1202–1210.

Kessels RPC, Berg E van den, Ruis C, Brands AMA (2008). The backward span of the Corsi Block-Tapping Task and its association with the WAIS-III Digit Span. *Assessment* **15**: 426–434.

Moeller S, Yacoub E, Olman CA, Auerbach E, Strupp J, Harel N, *et al* (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn Reson Med* **63**: 1144–1153.

Peelen MV, Fei-Fei L, Kastner S (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature* **460**: 94–97.

Setsompop K, Gagoski BA, Polimeni JR, Witzel T, Wedeen VJ, Wald LL (2012). Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magn Reson Med* **67**: 1210–1224.

Spaniol J, Davidson PSR, Kim ASN, Han H, Moscovitch M, Grady CL (2009). Event-related fMRI studies of episodic encoding and retrieval: meta-analyses using activation likelihood estimation. *Neuropsychologia* **47**: 1765–1779.

Unsworth N, Redick TS, Heitz RP, Broadway JM, Engle RW (2009). Complex working memory span tasks and higher-order cognition: a latent-variable analysis of the relationship between processing and storage. *Memory* **17**: 635–654.

Xu J, Moeller S, Auerbach EJ, Strupp J, Smith SM, Feinberg DA, *et al* (2013). Evaluation of slice accelerations using multiband echo planar imaging at 3 T. *Neuroimage* **83**: 991–1001.

Supplementary Figures

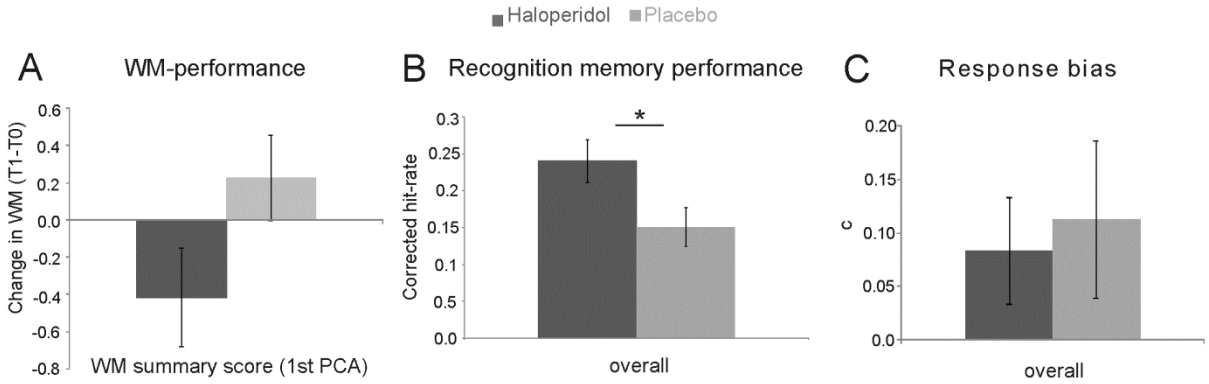

Figure S1: Additional behavioral performance results. A) The mean difference in WM performance after drug administration relative to baseline using the first principle component calculated from all available WM-span tasks was very close to significance (time x group interaction: $F(1,50) = 3.75$, $p = .058$. B) The mean corrected hit-rate (hit rate – FA rate) was significantly increased by nearly 10% in the haloperidol group ($t(52) = -2.27$, $p = .028$). C) No group differences in mean response bias c were observed (main effect of group: $F(1,50) = 0.13$, $p = .72$). Positive values of c represent a bias towards new-responses. * = significant group difference at $p < .05$. Error bars denote the SEM. HC = high confidence, MC = medium confidence, LC = low confidence.

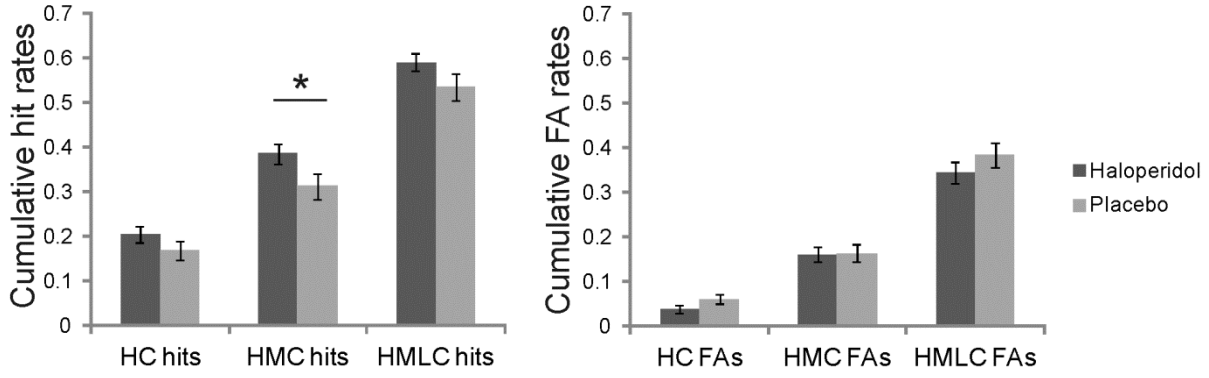

Figure S2: Cumulative hit and false alarm rates. Hit rates (left) and false alarm rates are plotted cumulatively from the most stringent to the most lax criterion. There were no group differences for the cumulative false alarm rate. For the cumulative hit rates, the haloperidol group had a higher hit rate when considering the two top levels of confidence responses ($t(52) = -2.03$, $p = .047$; not corrected for multiple comparisons). * = significant group difference at $p < .05$. Error bars denote the SEM. HC = high confidence, MC = medium confidence, LC = low confidence.

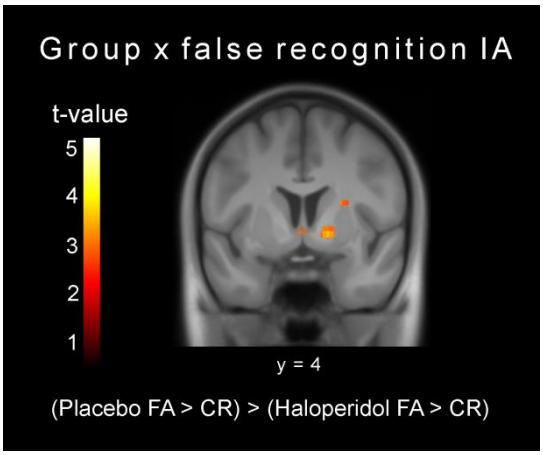

Figure S3: Exploratory analysis of drug effects on false recognition. Striatal voxels showing reduced activity in the haloperidol group for false alarms relative to correct rejection compared to the placebo group at an uncorrected threshold of p < .01.

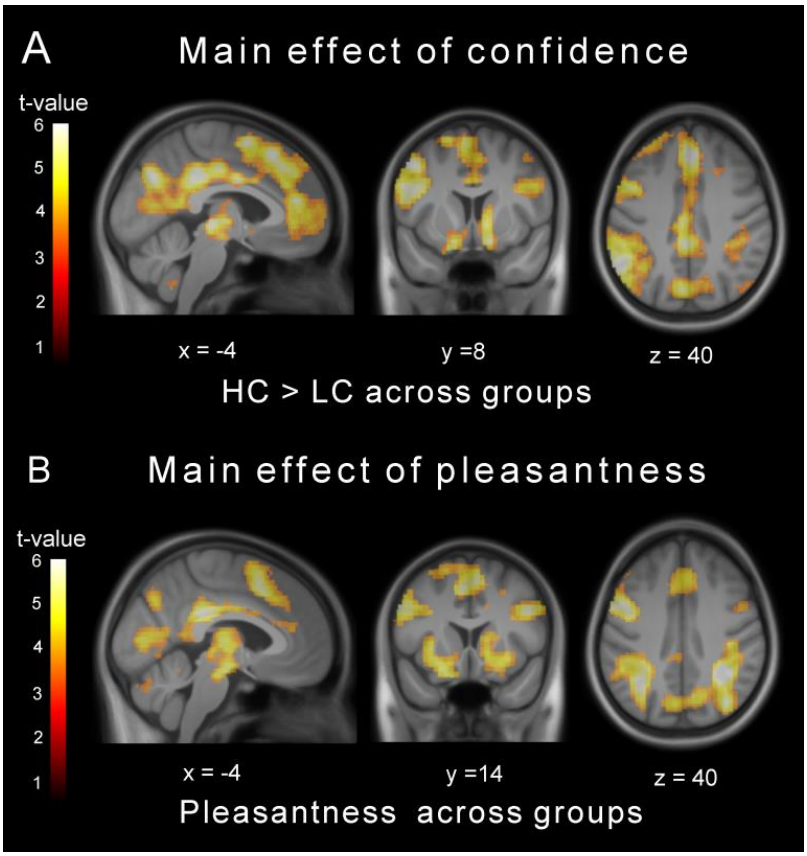

Figure S4: Additional fMRI results. A) Activation pattern for confidence across both groups. B) Activation pattern for pleasantness across both groups. No group differences were observed for pleasantness. For visualization purposes, activation maps are displayed at an uncorrected threshold of p < .001.

Supplementary Table

Table S1: Side effects and mood

	Placebo, n = 27	Haloperidol , n = 27	Statistics
Substance guess	21 placebo, 6 haloperidol	18 placebo, 9 haloperidol	$\chi^2(1) = 0.83$, p = .36
Certainty of substance guess (6 = very certain)	3.81±1.2	3.67±1.6	T(52) = 0.39, p = .70
Adverse medication effects sum score[a]	t0.5: -0.15±1.4 t2.0: -0.04±1.6 t4.5: 1.04±2.4	t0.5: -0.30±1.4 t2.0: -0.85±1.6 t4.5: 0.37±2.0	time x group interaction: F(2, 50) = 1.22, p = .31 main effect of group: F(1,51) = 1.94, p = .17
Systolic blood pressure[a]	t0.5: -0.80±4.0 t2.0: -2.00±4.8 t4.5: 0.80±6.4	t0.5: -1.67±5.4 t2.0: -2.80±6.7 t4.5: 1.85±5.9	time x group interaction: F(2, 49) = 0.74, p = .48 main effect of group: F(1,50) = 1.27, p = .27
Diastolic blood pressure[a]	t0.5: -2.00±4.1 t2.0: -1.60±3.7 t4.5: 0.40±4.5	t0.5: -0.56±2.9 t2.0: -0.56±5.1 t4.5: -0.10±4.7	time x group interaction: F(2, 49) = 1.81, p = .18 main effect of group: F(1,50) = 0.30, p = .59
Pulse[a]	t0.5: -0.38±4.9 t2.0: -4.92±4.3 t4.5: -1.77±9.3	t0.5: -1.19±6.0 t2.0: -2.59±7.0 t4.5: -1.63±7.9	time x group interaction: F(2, 50) = 2.53, p = .09 main effect of group: F(1,51) = 0.14, p = .71
Alertness[a]	t0.5: 0.01±0.02 t2.0: 0.01±0.03 t4.5: 0.03±0.04	t0.5: -0.01±0.04 t2.0: -0.02±0.04 t4.5: 0.02±0.05	time x group interaction: F(2, 50) = 2.23, p = .12 main effect of group: F(1,51) = 3.41, p = .07
Calmness[a]	t0.5: -0.01±0.04 t2.0: -0.01±0.06 t4.5: -0.02±0.05	t0.5: -0.02±0.05 t2.0: -0.01±0.05 t4.5: -0.02±0.07	time x group interaction: F(2, 50) = 0.55, p = .58 main effect of group: F(1,51) = 0.73, p = .79
Contentedness[a]	t0.5: 0.00±0.01 t2.0: 0.00±0.02 t4.5: 0.01±0.03	t0.5: -0.01±0.04 t2.0: -0.01±0.04 t4.5: 0.00±0.05	time x group interaction: F(2, 50) = 0.84, p = .44 main effect of group: F(1,51) = 0.98, p = .33

± denotes the standard variation

[a] relative to baseline measured just before tablet ingestion