SUPPLEMENTARY MATERIAL

Effect of mGluR2 Positive Allosteric Modulation on Fronto-Striatal Working Memory Activation in Schizophrenia

Daniel H. Wolf, MD, PhD[1], David Zheng BA[1], Christian Kohler, MD[1], Bruce I. Turetsky, MD[1], Kosha Ruparel, MS[1], Theodore D. Satterthwaite, MD, MA[1], Mark A. Elliott, PhD[2], Mary E. March, MS[1], Alan J. Cross, PhD[3,4], Mark A. Smith, MD, PhD[3,5,6], Stephen R. Zukin, MD[3,7], Ruben C. Gur, PhD[1,2], Raquel E. Gur, MD, PhD[1,2]

Departments of Psychiatry[1] and Radiology[2], University of Pennsylvania, Philadelphia PA 19104. AstraZeneca Pharmaceuticals LP[3], Wilmington DE, 19850. Present affiliations: Psy Therapeutics, Thornton PA 19373[4]; Vistagen Therapeutics, South San Francisco CA 94080[5]; Medical College of Georgia, Augusta, GA 30912[6]; PRA Health Sciences, Blue Bell PA 19422[7]

Contents of Supplementary Materials:

SUPPLEMENTARY METHODS

Inclusion/Exclusion Criteria

Participants had to be between the ages of 18-60 years; meet DSM-IV criteria for schizophrenia or schizoaffective depressed type; not become pregnant during the study (using adequate contraception if pre-menopausal/fertile); and possess adequate reading ability (wide range achievement test (WRAT [1]) standardized reading score >=70). Participants were excluded if they had any history of neurological disorder including head trauma with loss of consciousness; a diagnosis of substance dependence in the past 12 months or a positive urine drug screen for any unprescribed substances; an active medical condition that was deemed to affect brain function; a positive HIV or Hepatitis B test; ECT treatment in the previous 6 months; a positive urine pregnancy test; metal implants, claustrophobia, or other contraindications to MRI.

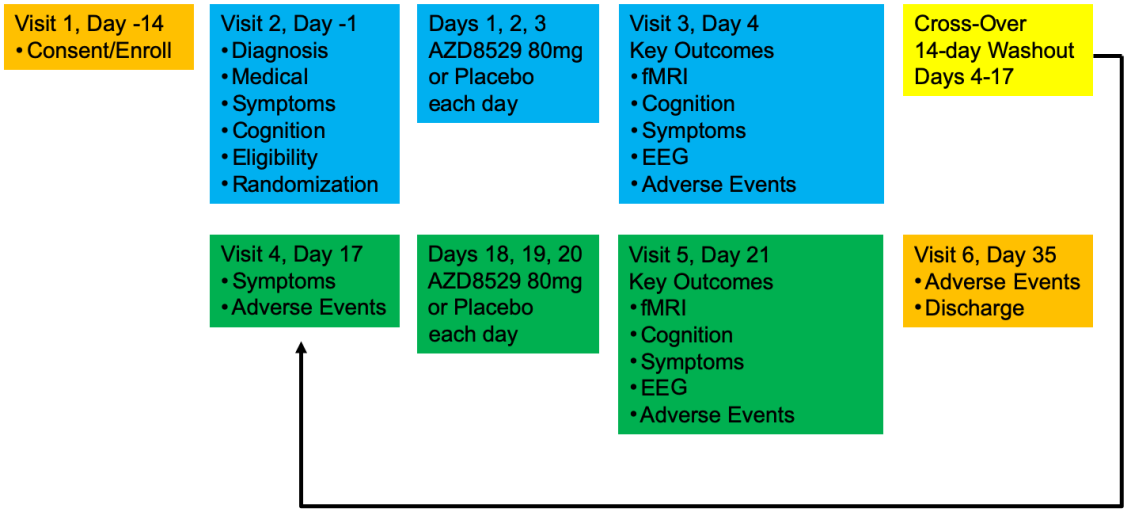
In this pilot study sample size was selected based on feasibility and typical pharmaco-fMRI sample sizes, without a pre-specified effect size. A total of 45 participants met inclusion/exclusion criteria and were randomized into remainder of the study protocol. 17 of these subjects were withdrawn after randomization for various reasons including new identification of exclusion criteria (n=4), serious adverse event (hyperglycemia, n=2), failure to take the study medication as prescribed (n=1), failure to show up to a subsequent study visit (n=10), leaving 28 subjects who completed the entire study protocol. Two subjects were subsequently excluded from fMRI analysis due to excessive motion during scanning (mean relative displacement [MRD] >3SD [0.68 mm] across all participants).

Design and Assessment Details

The study employed a randomized, double-blind, placebo-controlled, within-subject counterbalanced crossover design with a total of 6 visits (see **Figure S1** below for schematic of study design). There were 2 treatment periods: placebo phase and drug phase, each lasting 3 days. Participants were kept on their usual antipsychotics throughout the study. Subjects were enrolled and provided written informed consent on Visit 1 (Day -14). They then underwent a comprehensive clinical assessment, including a structured diagnostic interview (SCID [2]), and physical examination and laboratory screening tests (ECG, urine drug and pregnancy screens, serum chemistry and hematology). Subjects meeting specified inclusion/exclusion criteria were randomized and returned for Visit 2 within the next 14 days. A blinded consecutive-enrollment randomization schedule was prepared by the study sponsor, randomly assigning participants in a 1:1 ratio to a particular order (drug-then-placebo or placebo-then-drug). At Visit 2 (Day -1), baseline clinical and cognitive measures were obtained. Eligible participants were randomized and received 3 doses of either the active drug (80mg AZD8529) or placebo, to be taken once per day for the next 3 days (Days 1-3), as an adjunct to their normal medication regimen. On Visit 3 (Day 4), approximately 12 hours after the last dose, the laboratory screening tests, clinical and neurocognitive measures were obtained, and a blood sample was taken to determine the plasma level of AZD8529. Subjects then underwent functional magnetic resonance imaging and electrophysiology, in a randomized order. This was followed by a fourteen-day washout period, and then the intervention and assessments and intervention were repeated, with each subject crossing over from either active drug to placebo or placebo to active drug. Procedures during and between Visit 4 (Day 17) and Visit 5 (Day 21) were thus identical to those for Visits 2-3 above. Participants returned two weeks later for follow-up Visit 6 (Day 35), to be assessed for

any side effects or adverse events associated with the study, after which they were discharged from the study and instructed to continue antipsychotic medication at the discretion of their treating physician.

All study interviews were administered by trained assessors with demonstrated reliability on the relevant measures (reliability criterion 0.90 intraclass correlation). The main negative symptom (PANSS [3]) and cognitive measures (CNB, see below) analyzed here were obtained on the post-treatment visits (Visits 3 and 5), which were the key outcome days when fMRI data was collected. PANSS was also collected at visits 1, 2, and 4; CNB was also collected at visit 1. Anxiety was assessed with the State-Trait Anxiety Inventory (STAI [4]), both state and trait on visit 3, and state on visit 5. A 26-item pilot version of the Clinical Assessment Interview for Negative Symptoms (CAINS[5]) was collected as an additional exploratory measure of negative symptoms in most but not all participants, on visits 1-4. Suicidality was assessed with the Columbia-Suicide Severity Rating Scale (C-SSRS [6]) on all study visits, and extrapyramidal symptoms were assessed with the Simpson-Angus Scale (SAS [7]) on visits 2 through 5.



Supplementary Figure S1. Schematic of Study Design

Out-of-Scanner Cognitive Assessment

The Penn Computerized Neurocognitive Battery (CNB [8-10]) was used to measure cognitive performance outside of the scanner. Overall cognitive performance was calculated as an average of the z-scored accuracy measures for the following tasks: delayed word memory, delayed face memory, letter n-back (1-back and 2-back levels combined), letter continuous performance task, number continuous performance task, progressive matrices, face emotion recognition, face emotion discrimination. Due to technical errors, CNB data was not available for 3 of the 26 participants.

Image Acquisition & Processing Details

Subjects were placed in the scanner supine, earplugs were used to muffle scanner noise, head fixation was ensured by a foam-rubber device mounted on the head coil, and pulse and respiration monitors were attached. Stimuli were rear-projected to the center of the visual field using a PowerLite 7300 video projector (Epson America, Inc.; Long Beach, CA) and viewed through a head coil mounted mirror. Stimulus presentation was synchronized with image acquisition using the Presentation software package (Neurobehavioral Systems, Inc., Albany, CA). Subjects provided responses with a non-ferromagnetic response device (fORP, Current Designs, Inc., Philadelphia, PA) using their dominant hand.

Whole-brain structural data were obtained with a 5-minute 48-second magnetization-prepared, rapid acquisition gradient-echo T1-weighted image (MPRAGE, TR 1810ms, TE 3.51 ms, TI 1100ms, FOV 180x240 mm, matrix 192x256, effective voxel resolution of .94 x .94 x 1mm). Pseudo-continuous Arterial Spin-Label (pCASL) perfusion MRI was used to measure absolute CBF at rest with eyes open viewing a black screen. Forty label/control pCASL image

pairs were acquired with the following parameters: FOV=220 mm, matrix=96X96, TR=4s, TE=29ms, flip angle=90, 20 slices (5mm thick with 1mm gap), label duration 1500ms, post-label delay 1200ms. BOLD fMRI data was obtained as a single-shot gradient-echo (GE) echoplanar sequence using the following parameters: TR/TE=3000/30 ms, FOV=240 mm, matrix= 64 X 64, slice thickness/gap=3/0mm, 40 slices, effective voxel resolution of 3 x 3 x 3mm.

BOLD data were preprocessed and analyzed using FEAT, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). BOLD images were skull-stripped, motion corrected to median time point reference using tri-linear interpolation, high-pass filtered (n-back 138s, CPT 108s, EIT 100s), spatially smoothed (isotropic FWHM 6mm n-back and CPT, 4mm EIT), and grand-mean scaled using mean-based intensity normalization. The median functional image was transformed by trilinear interpolation into standard anatomical space using the T1 Montreal Neurological Institute (MNI) template with 2x2x2 mm voxel dimensions; this provided transformation parameters that were later applied to subject-level statistical images for group-level analyses. PASL perfusion image pairs were realigned to the mean image to correct for motion. Forty perfusion-weighted images were then generated via surround subtraction and converted to absolute CBF following the procedure summarized in [11], yielding a series of 40 CBF images that were averaged to produce one CBF image per participant.

Region of Interest Definitions

As in Ordaz et al. [12], unbiased ROIs were defined in MNI space using 10mm spheres centered around peak coordinates identified via Neurosynth meta-analysis (neurosynth.org [13]). The-meta-analytic search term was “working memory” (901 studies) using the forward inference/uniformity test. For bilateral ROIs, symmetric spheres were generating by taking the

peak unilateral coordinate and generating the symmetric sphere in the opposite hemisphere by flipping the sign of the x coordinate. The bilateral DLPFC ROI (1030 $2mm^3$ voxels) was centered on MNI coordinates (±46, +6,+28). The midline ACC ROI (515 $2mm^3$ voxels) was centered on MNI coordinate (-2, +18,+46). The bilateral STR ROI (1029 $2mm^3$ voxels) was centered on MNI coordinates (±10, +10,+4). These ROIs are shown in **Figure S2,** alongside images of the neurosynth metaanlysis map used to identify the peak voxels spheres were centered on. Note that the ROI we refer to as ACC extends from the superior dACC into paracingulate as does the neurosynth activation cluster; working memory tasks generally show activation in this cluster that includes both dACC and paracingulate/supplementary motor area.

Continuous Performance Task, Face Emotion Identification Task, and Perfusion

In addition to the n-back, two other fMRI tasks were also performed: a Continuous Performance Task (CPT) and an Emotion Identification Task (EIT). We focused on the n-back because the CPT and EIT were less suited than the n-back to investigating frontostriatal circuitry (see below).

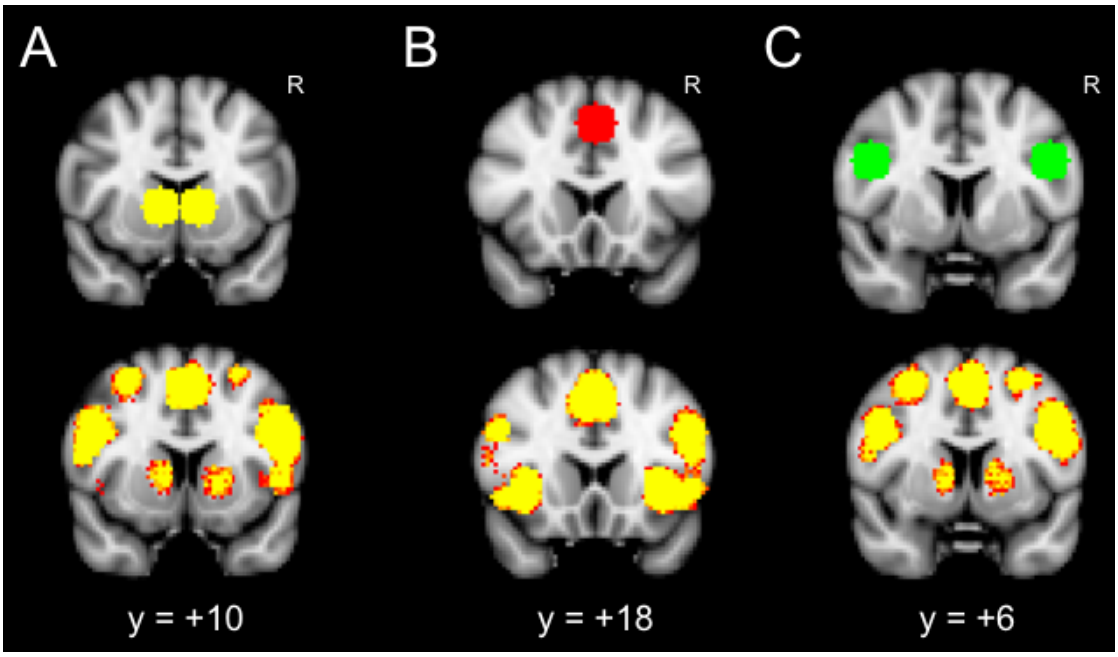
The Penn Continuous Performance Test (Penn CPT) requires participants to respond with button press whenever a 7-segment display forms a digit (number block) or letter (letter block). Here we adapted our fMRI version of this task [14] into a brief version. Stimuli are grouped into five 30 sec blocks with a 3 sec instruction period. Each stimulus appears for 300 ms followed by a black screen for 700 ms. Participants respond by pressing a button whenever the 7-segment display forms a numeral or a letter. Total task time is 2 min 33 sec, 51 time points.

The Face Emotion Identification Task (EIT) presents one face image per trial and requires the participant to select one of five emotional expressions (happy, sad, anger, fear, neutral) using a single axis response device. Facial stimuli consist of color photographs of actors

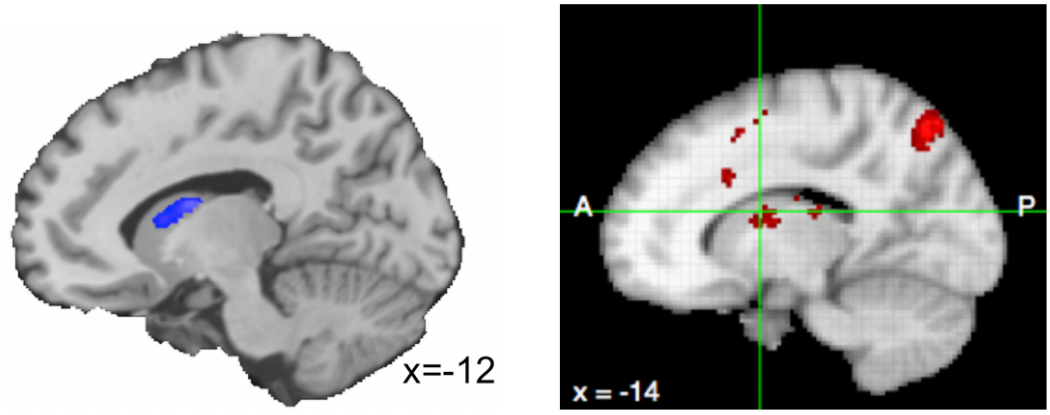
(50% female) and actresses of various ethnicities. Here we adapted our fMRI version of this task [15] to be slightly shorter. Each of the five emotions is presented on 12 trials (but with unique actors), with an event-related design. Faces are presented for 5 seconds; the baseline inter stimulus interval is variable (6-12 sec). Total task time is 12 minutes 30 seconds; 210 time points (the first 6 volumes include dummy trials which are deleted in preprocessing leaving 204 time points for analysis).

For the brief CPT, the primary task>baseline contrast did not produce robust activation in any regions outside of visual cortex, and therefore was not examined in detail. An exploratory whole-brain voxelwise analysis looking for drug-placebo effects on this task contrast did not identify any significant regions, using the same TFCE multiple-comparisons approach as applied for the n-back.

The EIT task was designed primarily to examine drug effects in the amygdala as well as orbitofrontal cortex. These regions were activated by the task, but there were no significant drug-placebo differences in these ROIs nor in a whole-brain voxelwise analysis. Perfusion data also did not reveal any significant drug-placebo differences.



Supplementary Figure S2. Regions of interest. Top Row: Unbiased ROIs in striatum (A), dorsal anterior cingulate/paracingulate (B), and dorsolateral prefrontal cortex (C), and striatum (C). Bottom Row: Corresponding images of the "working memory" term meta-analysis from Neurosynth, which were used to identify the peaks at the center of the ROI spheres.



Supplementary Figure S3. Comparison of Striatal Regions. Voxelwise image on left shows the same image as in main text Figure 3, where drug effects on striatal working memory activation correlate with drug effect on negative symptoms. This image is masked by the anatomical striatum, and is independent of the a priori spherical ROI derived from neurosynth. For comparison, the image on the right shows the activation cluster from the neurosynth metaanalysis for working memory, demonstrating correspondence of striatal regions.

Supplementary Results: PANSS Sub-Domains

As the negative symptom score was the PANSS domain of a priori interest, we focus on that in the main manuscript, but examined other domains in an exploratory fashion here. In general, greater drug-induced activation correlated with drug-induced reductions in symptoms, but the pattern varied across ROIs (Supplementary Table S1 below). The correlation with the PANSS negative symptom score was only significant in striatum. To parse the broader negative symptom domain, motivation and expressivity factor scores were calculated according to items and weights described in Fervaha et al. 2014 [16]. Correlation strength with the reduced motivation factor was similar to that seen for the overall PANSS negative symptom score in striatum, but unlike the negative symptom score was also significant in ACC and showed a statistical trend for DLPFC. The reduced expressivity factor did not show significant correlations. We also calculated a PANSS depression factor score using items and weights described in Lindenmayer et al. 1994 [17]. Change in PANSS depression scores showed a nearly-significant correlation with change in striatal activation ($r=-0.38$, $p=0.054$), and correlated significantly with effects in ACC ($r=-0.40$, $p=0.045$) but not DLPFC ($r=-0.18$, $p=0.39$).

The relationship between drug effects on striatum activation and PANSS negative symptoms identified in our primary analysis ($r=-0.42$, $p=0.03$) remained significant after controlling for PANSS positive symptoms (STR:negative PANSS partial $r= -0.39$, $p=0.049$; STR:positive PANSS partial $r=-0.22$ $p=0.24$). Controlling for PANSS general symptoms, the STR:negative PANSS relationship was nearly-significant (STR:negative PANSS partial $r= -0.36$, $p=0.059$; STR:general PANSS partial $r=-0.30$, $p=0.12$). Controlling for PANSS depression factor, the STR:negative PANSS relationship remained significant (STR:negative PANSS partial $r= -0.37$, $p=0.048$; STR:PANSS depression partial $r=-0.33$, $p=0.08$). In each case, the partial

correlation with the other symptom domain tended to be weaker than the partial correlation for the negative symptom domain when both were examined jointly (controlling for each other), and the magnitude of the STR:negative symptom correlation was not substantially reduced, indicating that the STR: negative symptom relationship is not simply explained by a confounding relationship with one of these other symptom domains. In our modest sample size we do not have the statistical power to demonstrate statistical differences in the strength of correlated correlations between different PANSS domains and ROIs shown in Table S1.

Examining intercorrelations in change scores (drug - placebo) for major PANSS scores, change in PANSS negative scores correlated significantly with change in PANSS total scores but not with change in PANSS positive or general scores (Supplementary Table S2 below). Thus it does not appear that the negative symptom-striatum relationship is simply a non-specific effect of global psychopathology, but we cannot rule out a global component or a role of other unmeasured symptom domains. Overall, the pattern across ROIs suggests either some global severity component and/or a complex pattern of symptom-region specificity which we cannot effectively parse due to study sample size and other study limitations.

	Regions of Interest		
PANSS Variables	**STR**	**ACC**	**DLPFC**
Negative	r -0.42, p 0.03	r -0.04, p 0.86	r -0.23, p 0.27
Positive	r -0.29, p 0.15	r -0.35, p 0.08	r -0.24, p 0.23
General	r -0.37, p 0.06	r -0.41, p 0.04	r -0.37, p 0.06
Total	r -0.53, p 0.006	r -0.38, p 0.053	r -0.41, p 0.04
Motivation Factor	r -0.46, p 0.02	r -0.34, p 0.09	r -0.49, p 0.01
Expressivity Factor	r -0.22, p 0.29	r 0.20, p 0.32	r 0.04, p 0.87
Depression Factor	r -0.38, p 0.054	r -0.40, p 0.045	r -0.18, p 0.39

Supplementary Table S1. Exploratory correlations of drug effects (drug-placebo) on ROI activation and PANSS domains. Pearson's r and uncorrected p-value are shown.

PANSS Domains	Negative	Positive	General	Total
Negative	-	r 0.16, p 0.42	r 0.20, p 0.33	r 0.67, p 0.0002
Positive	r 0.16, p 0.42	-	r 0.27, p 0.18	r 0.63, p 0.0005
General	r 0.20, p 0.33	r 0.27, p 0.18	-	r 0.75, p <0.0001
Total	r 0.67, p 0.0002	r 0.63, p 0.0005	r 0.75, p <0.0001	-

Supplementary Table S2. Exploratory intercorrelations of drug effects on PANSS domains. Pearson's r value and uncorrected p-value are shown.

Supplementary Results: Anxiety

To assess potential anxiolytic effects of AZD8529, we examined the STAI state anxiety scores. State anxiety did not show any difference between AZD8529 (p=0.99), nor did the drug effect on state anxiety correlate with drug effects on n-back fMRI activation in our regions of interest (STR r=-0.004, p=0.99; DLPFC r=0.2, p=0.17; ACC r=0.04, p=0.84). STAI scores were missing for two participants, but the lack of effect for anxiety does not simply reflect removal of two participants, as the correlation of drug effects between striatal and negative symptoms is unchanged by removing these two individuals (r=-0.42, p=0.04).

Supplementary Results: Extrapyramidal Symptoms

All of the patients in our sample were taking antipsychotic medication, which could contribute to secondary negative symptoms, and if the mGluR2 PAM alleviated D2 blockade that could contribute to observed effects. We therefore examined extrapyramidal symptoms as measured with the Simpson Angus Scale (SAS scores, averaged across 10 items, where items range from 0-none to 4-severe). Extrapyramidal symptoms were very low in this sample, with mean±SD across conditions 0.23±0.22, maximum 0.8 (mild). While SAS scores remained very low under both drug and placebo conditions, scores were statistically greater under the AZD8529 condition than placebo (drug 0.29±0.21; placebo 0.17±0.21; t=2.9, p=0.01). This does not support the idea

that AZD8529 reduces secondary negative symptoms by interfering with antipsychotic effects. Furthermore, drug effects on extrapyramidal symptoms did not confound the drug effects on fMRI activation, as there was no significant relationship of change in SAS scores with change in activation in our ROIs (STR r=-0.08, p=0.71; ACC r=-0.03, p=0.88; DLPFC r=-0.13, p=0.53). There was also no correlation of drug effect on SAS with drug effect on PANSS scores (PANSS negative r=-0.02; positive r=-0.03; general rr=0.11; total r=0.03; all p's>0.6).

Supplementary Results: AZD8529 Blood Levels

The main goal of obtaining AZD8529 blood levels in this study was to ensure that the dosing regimen used in fact produced average levels expected based on preclinical data to produce neurobiological effects of interest. We explored the relationship between blood levels and change in fMRI and PANSS scores. Higher blood levels correlated with higher activation in DLPFC (r=0.64, p=0.02) but not ACC (r=0.22, p=0.32) or STR (r=0.14, p=0.52). Higher blood levels did not correlate with change in PANSS symptoms (PANSS negative r=-0.04; positive r=-0.02; general r=-0.16; total r=-0.12; all p's>0.48).

Supplementary Results: Other Confound Analyses

Activations in striatum, DLPFC and ACC (across dug and placebo, or drug-placebo differences) were not significantly related to participant sex, age, education, parental education, global cognitive performance, or smoking status (p's >0.3). DLPFC did show a trend relationship of a larger drug effect in those with higher education (p=0.08). There were no effects of these variables on in-scanner n-back performance (*d'*). Older participants showed worse CNB performance (p=0.02), without a significant drug x age interaction.

Higher in-scanner motion was significantly related to lower task activation in striatum ($p=0.001$) but not DLPFC ($p=0.42$) or ACC ($p=0.14$). There was no significant effect of drug on motion ($p=0.16$), and there was no significant relationship between the drug effect on motion and the drug effect on activation in any region or the drug effects on symptoms, and key results reported in main text remained significant with motion included in the models. Activation was lower on day 2 than day 1 in DLPFC ($p<0.0001$) and ACC ($p=0.01$) but not striatum ($p=0.13$). Drug effects did not differ by day in striatum or ACC, but were smaller in DLPFC on day 2 than on day 1 ($p=0.02$). There were no day effects on d' or on overall cognitive accuracy in the CNB. Key results reported in main text remained significant with day included in the models.

Supplementary References

1. Wilkinson GS, Robertson GJ. *Wide range achievement test-fourth edition*. Psychological Assessment Resources: Lutz, FL, 2006.

2. First MB, Spitzer RL, Gibbon M, Williams JBW. *Structured clinical interview for DSM-IV-TR Axis I disorders - patient edition*. Biometrics Research, New York State Psychiatric Institute: New York, 2001.

3. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia bulletin* 1987; **13:** 261-276.

4. Spielberger C, Gorsuch R, Lusthene R. *Manual for the state-trait anxiety inventory.* Consulting Psychologist Press: Palo Alto, 1983.

5. Horan WP, Kring AM, Gur RE, Reise SP, Blanchard JJ. Development and psychometric validation of the clinical assessment interview for negative symptoms (CAINS). *Schizophrenia Research* 2011; **132:** 140-145.

6. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA et al. The columbia-suicide severity rating scale: Initial validity and internal consistency findings from three multisite studies with adolescents and adults. *American Journal of Psychiatry* 2011; **168:** 1266-1277.

7. Simpson GM, Angus JW. A rating scale for extrapyramidal side effects. *Acta Psychiatr Scand Suppl* 1970; **212:** 11-19.

8. Gur RC, Ragland JD, Moberg PJ, Turner TH, Bilker WB, Kohler C et al. Computerized neurocognitive scanning: I. Methodology and validation in healthy people. *Neuropsychopharmacology* 2001; **25:** 766-776.

9. Gur RC, Ragland JD, Moberg PJ, Bilker WB, Kohler C, Siegel SJ et al. Computerized neurocognitive scanning: Ii. The profile of schizophrenia. *Neuropsychopharmacology* 2001; **25:** 777-788.

10. Gur RC, Richard J, Hughett P, Calkins ME, Macy L, Bilker WB et al. A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: Standardization and initial construct validation. *Journal of Neuroscience Methods* 2010; **187:** 254-262.

11. Wong EC, Buxton RB, Frank LR. Quantitative imaging of perfusion using a single subtraction (QUIPSS and QUIPSS II). *Magn Reson Med* 1998; **39:** 702-708.

12. Ordaz SJ, Foran W, Velanova K, Luna B. Longitudinal growth curves of brain function underlying inhibitory control through adolescence. *J Neurosci* 2013; **33:** 18109-18124.

13. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods* 2011; **8:** 665-670.

14. Roalf DR, Ruparel K, Gur RE, Bilker W, Gerraty R, Elliott MA et al. Neuroimaging predictors of cognitive performance across a standardized neurocognitive battery. *Neuropsychology* 2014; **28:** 161-176.

15. Wolf DH, Satterthwaite TD, Loughead J, Pinkham A, Overton E, Elliott MA et al. Amygdala abnormalities in first-degree relatives of individuals with schizophrenia unmasked by benzodiazepine challenge. *Psychopharmacology* 2011; **218:** 503-512.

16. Fervaha G, Foussias G, Agid O, Remington G. Motivational and neurocognitive deficits are central to the prediction of longitudinal functional outcome in schizophrenia. *Acta Psychiatrica Scandinavica* 2014; **130:** 290-299.

17. Lindenmayer, JP, Bernstein-Hyman, R, Grochowski, S. A new five factor model of schizophrenia. Psychiatric Quarterly; 65: 299-322.