



Published in final edited form as:

Schizophr Res. 2017 March ; 181: 107–116. doi:10.1016/j.schres.2016.10.011.

Schizophrenia is associated with a pattern of spatial working memory deficits consistent with cortical disinhibition

Martina Starc^{a,b,1}, John D. Murray^{a,c,1}, Nicole Santamauro^a, Aleksandar Savic^{a,d}, Caroline Diehl^{a,b}, Youngsun T. Cho^a, Vinod Srihari^a, Peter T. Morgan^a, John H. Krystal^{a,h}, Xiao-Jing Wang^c, Grega Repovs^{b,1}, and Alan Anticevic, Ph.D.^{a,e,f,g,i,*1}

^aDepartment of Psychiatry, Yale University School of Medicine, 300 George Street, New Haven, CT 06511, USA

^bDepartment of Psychology, University of Ljubljana, Ljubljana, Slovenia

^cCenter for Neural Science, New York University, New York, NY 06510, USA

^dUniversity Psychiatric Hospital Vrapce, University of Zagreb, Zagreb 10000, Croatia

^eAbraham Ribicoff Research Facilities, Connecticut Mental Health Center, New Haven, CT 06519, USA

^fInterdepartmental Neuroscience Program, Yale University, New Haven, CT 06520, USA

^gNIAAA Center for the Translational Neuroscience of Alcoholism, New Haven, CT 06519, USA

^hDepartment of Neurobiology, Yale University, New Haven, CT, USA

ⁱDepartment of Psychology, Yale University, 2 Hillhouse Avenue, CT 06520, USA

Abstract

Schizophrenia is associated with severe cognitive deficits, including impaired working memory (WM). A neural mechanism that may contribute to WM impairment is the disruption in excitation-inhibition (E/I) balance in cortical microcircuits. It remains unknown, however, how these alterations map onto quantifiable behavioral deficits in patients. Based on predictions from a

*Corresponding author at: Yale University, Department of Psychiatry, 34 Park St., New Haven, CT 06519, USA. alan.anticevic@yale.edu (A. Anticevic).

¹Designates equal contributions.

Financial conflicts of interest

J.H.K. consults for several pharmaceutical and biotechnology companies with compensation less than \$10,000 per year. He also has stock options in two companies, each valued less than \$2000 and three patents for pharmacotherapies for psychiatric disorders. None of these financial interests are directly related to this paper. All other authors declare that they have no conflict of interest.

Financial disclosures

Dr. John H. Krystal consults for several pharmaceutical and biotechnology companies with compensation less than \$10,000 per year. Dr. Alan Anticevic serves on the Scientific Advisory Board and consults BlackThorn Therapeutics. Dr. John D. Murray consults for BlackThorn Therapeutics.

Contributors

AA, GR & JDM conceptualized and designed the study. AA, NS, PM, AS, CD, Y.T.C & VS collected the data. MS and AA performed data analyses. AA examined and interpreted the results in consultation with MS and JDM. MS wrote the first draft of the manuscript, which AA, JHK and JDM edited. JDM and X-JW performed the computational modeling aspect of the study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.schres.2016.10.011.

validated microcircuit model of spatial WM, we hypothesized two key behavioral consequences: i) increased variability of WM traces over time, reducing performance precision; and ii) decreased ability to filter out distractors that overlap with WM representations. To test model predictions, we studied $N=27$ schizophrenia patients and $N=28$ matched healthy comparison subjects (HCS) who performed a spatial WM task designed to test the computational model. Specifically, we manipulated delay duration and distractor distance presented during the delay. Subjects used a high-sensitivity joystick to indicate the remembered location, yielding a continuous response measure. Results largely followed model predictions, whereby patients exhibited increased variance and less WM precision as the delay period increased relative to HCS. Schizophrenia patients also exhibited increased WM distractibility, with reports biased toward distractors at specific spatial locations, as predicted by the model. Finally, the magnitude of the WM drift and distractibility were significantly correlated, indicating a possibly shared underlying mechanism. Effects are consistent with elevated E/I ratio in schizophrenia, establishing a framework for translating neural circuit computational model of cognition to human experiments, explicitly testing mechanistic behavioral hypotheses of cellular-level neural deficits in patients.

Keywords

Schizophrenia; Working memory; Computational modeling; Disinhibition; Cognitive deficits; Excitation/inhibition balance

1. Introduction

Schizophrenia (SCZ) is associated with profound cognitive deficits, such as memory and executive function (Barch and Ceaser, 2012; Kalkstein et al., 2010), which are among the best predictors of vocational and social disability (Green, 2006; Nuechterlein et al., 2011). Problems with working memory (WM) – the ability to transiently maintain and manipulate information internally – are particularly prominent in SCZ (Forbes et al., 2009; Lee and Park, 2005) and have been proposed as a core cognitive deficit in this illness (Barch and Ceaser, 2012; Goldman-Rakic, 1994; MacDonald et al., 2006; Silver et al., 2003; Van Snellenberg and de Candia, 2009). While SCZ patients show WM deficits across modalities (Quee et al., 2010), spatial WM is particularly amenable for clinical translation as it can be studied across animal (Wang et al., 2013), pharmacological (Driesen et al., 2013), computational modeling (Compte et al., 2000) and patient studies (Driesen et al., 2008) using comparable paradigms. Here we examined specific mechanisms of spatial WM deficits based on explicit predictions of a biophysically-based computational WM model (Murray et al., 2014) informed by primate physiology experiments.

Decades of primate and human studies have implicated prefrontal cortex (PFC) neural circuits in WM maintenance and manipulation, which are impaired in SCZ (Anticevic et al., 2013b; Barch and Ceaser, 2012; Metzak et al., 2011). The cellular basis of spatial WM maintenance implicates persistent firing of location-selective PFC pyramidal cells (Funahashi et al., 1989)—extensively characterized by animal electro-physiology (Rao et al., 2000; Wang et al., 2013) and computational modeling (Compte et al., 2000; Durstewitz et al., 1999; Durstewitz and Seamans, 2002; Wang, 2006; Wang, 2010; Wang et al., 2004).

The models propose that WM is supported by the interplay between recurrent excitation (E) among pyramidal neurons (which sustains persistent activity over the delay) and lateral inhibition (I) mediated by interneurons (which stabilizes WM representations and reduces the impact of external distraction) (Compte et al., 2000; Murray et al., 2014; Wang et al., 2004). Specific alterations in optimal E/I balance disrupt the ability to represent information and shield WM from interference (Rao et al., 2000). One such potential alteration — disruption of inhibitory interneurons, leading to cortical disinhibition — has been implicated in SCZ neuropathology (Lewis et al., 2005; Marin, 2012). It remains unknown, however, how such cellular hypotheses map onto quantifiable behavioral WM deficits. Here we tested the behavioral consequences of altered E/I balance on spatial WM performance in SCZ.

Biophysically realistic computational modeling offers one strategy to quantify the impact of altered E/I balance on WM (Anticevic et al., 2013a; Anticevic et al., 2015; Murray et al., 2014). The consequences of cortical disinhibition on WM are well-characterized by a spiking local circuit model comprised of *E*- and *I*-cells (Murray et al., 2014). *E*-cells interact through horizontal connections mediating recurrent excitation via *N*-methyl-D-aspartate receptors (NMDAR) and a pool of *I*-cells mediates feedback synaptic inhibition. Prior work modeled cortical disinhibition by reducing NMDAR conductance for *E*-*I* connections (Kotermanski and Johnson, 2009), thought to occur in SCZ by the well-established NMDAR hypo-function hypothesis (Anticevic et al., 2012a; Krystal et al., 2003a; Macdonald and Chafee, 2006).

We used this model architecture to derive testable qualitative behavioral predictions generated by NMDAR hypo-function on *I*-cells. First, the model predicted that following disinhibition, WM representational fields are broadened, resulting in increased random drift over time. This effect was observed behaviorally as a reduction in WM precision over longer delays. Second, the model predicted distorting effects of intervening distractors on WM, whereby under disinhibition, a broader WM representation exhibits a wider window in which distractors can interfere with WM (Murray et al., 2014). Finally, the model predicted that both increased response variability and increased distractor sensitivity stem from a shared underlying mechanism — altered E/I balance. We tested these three model-derived hypotheses in patients diagnosed with SCZ. Collectively, this ‘computational psychiatry’ study translates a neural circuit computational model of cognition to test behavioral hypotheses of cellular-level neural deficits in SCZ patients.

2. Methods

2.1. Subjects

We recruited $N=27$ SCZ patients from outpatient clinics of the Department of Psychiatry, Yale University and $N=28$ healthy comparison subjects (HCS) from the local community (Table 1). Subjects were independently diagnosed by two trained clinicians using the Structured Clinical Interview (SCID) for DSM-IV (First et al., 2001). All subjects provided informed consent approved by Yale Institutional Review Board. Patients met DSM-IV diagnostic criteria for SCZ or schizoaffective disorder, but no other Axis I diagnosis or drug abuse/dependence at the time of recruitment. Prior and current nicotine and alcohol use was permitted. HCS met the following inclusion criteria: i) no current or lifetime Axis I disorder

(determined by a trained PhD-level clinician); and ii) no history of psychotic, mood or other Axis I disorders in first-degree relatives (reported by detailed family history). Subjects were excluded if they had: i) history of other neurological conditions (e.g. epilepsy, migraine, head trauma, loss of consciousness); ii) any MRI contraindications; or iii) any concomitant major medical disorder. HCS were demographically matched to SCZ patients. However, groups differed in education attainment and measures of verbal and non-verbal intelligence, as expected in severe mental illness (Glahn et al., 2006) (Table 1). Critically, adjusting for these variables did not alter results (see Supplement).

2.2. Current symptoms & medication

Symptom severity was evaluated using the Scale for Assessment of Positive and Negative Symptoms (SAPS/SANS) (Andreasen, 1983b) and the Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1987). 85% of patients (23/27) were receiving antipsychotics, which we converted to chlorpromazine (CPZ) equivalents (Andreasen et al., 2010) (Table 1). None of the identified effects correlated with CPZ equivalents and did not change when we co-varied for medication dose (see Supplement).

2.3. Computational model

Full details of the computational model implementation were reported previously (Murray et al., 2014). For the purposes of the current experimental predictions we modeled a cortical circuit that performs spatial WM through stimulus-selective persistent activity. The circuit contains recurrently connected E pyramidal cells and I interneurons. Pyramidal cells are tuned to angular location. Stimulus inputs transiently excite a corresponding subset of *E*-cells, and a persistent activity pattern encodes stimulus location through the delay. Cortical disinhibition was implemented through a reduction of excitatory NMDA conductance on interneurons (G_{EI}), a site implicated in the pathology of SCZ (Anticevic et al., 2012a; Belforte et al., 2010; Krystal et al., 2003b). We used the population vector approach to decode the behavioral report location from the neural WM activity pattern. We characterized two aspects of model performance: i) time-dependent decay of WM precision by computing the across-trial variability of the decoded location as a function of delay duration. ii) behavioral impact of external distractors on WM report. Distractors were identical to the initial cue, with the same intensity and duration but with a different stimulus position (Fig. 2D).

2.4. Experimental design

Subjects completed two delayed spatial WM paradigms, designed to mimic primate physiology experiments (Goldman-Rakic, 1995) and the implemented computational model architecture (Compte et al., 2000). The tasks manipulated: i) delay period duration (testing if the SCZ group exhibits greater WM response variability as a function of delay duration), and ii) distance between the WM cue and distractor presented during the delay period (testing if the SCZ group exhibits a differential response bias in the direction of the distractor across two distractor distances, Fig. 1; see Supplement for comprehensive detail). Subjects also completed a control 'motor' task to verify that differences between groups are not driven exclusively by lower motor skill in patients. Subjects were instructed to keep their eyes fixed on the middle of the screen throughout the task and were monitored for

compliance by the experimenter (see Limitations for considerations surrounding eye tracking). Response expectations were controlled for by insuring that the number of trials decreased with delay duration (from 60 to 20; see Supplement for complete details).

2.5. Behavioral analyses

To maintain quality control, inadequate responses (3%) were excluded from the analysis (see Supplement). Testing for differences in the proportion of excluded trials did not yield any significant between-group effects (see Supplement & Figs. S1–4), suggesting that outlier trials did not influence results.

All remaining trials were included in the main analysis. To ensure that different numbers of trials and possible differences in measurement precision were not driving the results, a random subsample of 20 trials from each condition (one trial for each of the 20 used locations) were selected for additional analyses (see Supplement & Fig. S5). Sub-sampling did not change results in any way.

We examined two different dependent measures: i) angular displacement of the probe response relative to the cue; ii) standard deviation of angular displacement. SD is sensitive to overall spatial dispersion of responses relative to the cue (i.e. lower WM precision as a function of longer delay, model prediction 1), whereas angular displacement is sensitive to a directional bias in the probe response (i.e. toward distractors, model prediction 2).

Model-derived predictions were tested at the group level with a i) 2×6 mixed-measures ANOVAs with *Delay Duration* (motor, 0 s, 5 s, 10s, 15 s, 20s) as a within- and *Diagnosis* (SCZ, HCS) as a between-groups factor and ii) 2×3 mixed-measures ANOVAs with *Distractor Location* (none, 20° , 50°) as a within- and *Diagnosis* (SCZ, HCS) as a between-groups factor. Greenhouse–Geisser-corrected p -values are reported in cases where ANOVA sphericity assumptions were violated, which were determined with a Mauchly's test for sphericity ($p < 0.05$). We opted for this test since our sample size fell in the 'moderate' range by recommended standards (i.e. when Mauchly's tends to perform well which was the case in the reported analyses). Thus, approach was biased toward a more conservative side of the Mauchly's test (i.e. when sample size is sufficient). Furthermore, in such cases we always report the Greenhouse–Geisser epsilon value (GG ϵ) and the GG-corrected p -values to allow examination of sphericity deviations. Between groups post-hoc t-tests with Holm-Bonferroni corrections were computed where appropriate. All experimental analyses, statistics and visualization were implemented in the R statistical computing environment (<http://www.r-project.org>).

3. Results

3.1. Computational modeling results and predictions

Here we describe key results and qualitative behavioral predictions of the spiking circuit spatial WM model, which directly informed present experimental design and testing of model predictions (Murray et al., 2014). The model behavior shown here is generated with parameter values from our prior modeling study (Murray et al., 2014) without any attempt to adjust parameters to quantitatively fit empirical data. The model therefore provides

qualitative predictions for how WM behavior is altered under elevated E/I ratio, which motivated experimental design and planned analyses. Fig. 2A shows a spatiotemporal plot of the *E*-cell activity during WM. Disinhibition, resulting in an elevated E/I ratio, increases the width of the WM activity profile compared to the control condition (Fig. 2A) and predicts two key behavioral deficits — namely WM drift and susceptibility to distraction.

First, background noise in the model caused the WM activity pattern to undergo random drift during delay, degrading WM precision over time, in line with experimental findings (Ploner et al., 1998; White et al., 1994). Disinhibition increased the rate of this drift (Fig. 2B), which suggests that SCZ will be associated with higher drift-related variability in WM reports.

Second, we tested the model's ability to resist WM interference by external distraction. In line with human psychophysics (Herwig et al., 2010), the model exhibited a pattern of distractibility that depended on the similarity between the distractor and memorandum (Fig. 2C). Specifically, the model predicted that SCZ patients would exhibit increased bias of WM reports toward the distractor (Fig. 2D).

Finally, drift-induced WM variability and WM distractibility are positively correlated in the model, as they arise from a common underlying mechanism of disinhibition (Fig. 2E). Both measures are therefore predicted to correlate across individuals experimentally, irrespective of diagnostic category.

3.2. Testing effects of delay duration on spatial WM performance

We first examined if SCZ patients exhibit more variable response accuracy, especially as a function of increased delay duration. Response clouds in Fig. 3A indicate that the spread of WM responses increases with delay duration across groups. However, SCZ patients exhibited wider spread at all delay durations. Furthermore, the spread increased more for the SCZ group with longer delays. We tested this effect formally by calculating the SD of angular displacement (Fig. 3B). Response variability increased as a function of *Delay Duration* [$F(5265) = 150.665$, $GGe = 0.571$, $p < 0.001$], and was greater in the SCZ group (*Diagnosis* main effect [$F(1,53) = 11.395$, $p = 0.001$], Cohen's $d = 0.58$ – 1.10 , Fig. 3D). Critically, the slope of WM drift as a function of delay duration was steeper in SCZ group, as predicted by the model (*Delay Duration* \times *Diagnosis* interaction [$F(5265) = 6.155$, $GGe = 0.571$, $p < 0.001$]) (Fig. 2B). For each subsequent second of the delay period the angular displacement, SD increased on average for 0.15° in HCS group and 0.26 in SCZ group [$t(51.58) = 3.730$, $p < 0.001$, one-tailed, Cohen's $d = 1.01$] (Fig. 3E). Importantly, increased WM drift was not associated with a particular directional bias. That is, variability of responses increased across both groups (Fig. 3B). However, the average distribution remained centered on the WM cue (Fig. 3C), indicating a 'random' spatial drift. We observed no significant angular bias effects for *Diagnosis*, [$F(1,53) = 2.693$, $p = 0.107$], *Delay Duration* [$F(5265) = 0.173$, $GGe = 0.468$, $p = 0.872$], or *Delay Duration* \times *Diagnosis* [$F(5265) = 1.775$, $GGe = 0.468$, $p = 0.167$].

3.3. Testing effects of distractor position of spatial WM performance

Next, we assessed whether SCZ patients exhibit increased distractibility, especially as a function of distractor location. The increased spread of responses in SCZ group was replicated in the distractor task (*Diagnosis* main effect [$F(1,50) = 8.669, p = 0.005$]) (Fig. 4A-B). Results also revealed a main effect of *Distractor Location* [$F(2,100) = 11.846, GGe = 0.73, p < 0.001$], but no *Diagnosis* \times *Distractor Location* interaction [$F(2,100) = 2.972, GGe = 0.73, p = 0.073$]. In the proximal distractor condition, the spread of responses was non-specific. However, there was a net bias toward the distractor for the distal condition. Put differently, in addition to response variability, distractors affected the average angular displacement of response location (Fig. 4C). Proximal distractors caused a subtle angular shift away from the distractor in both groups equally. In contrast, distal distractors caused a shift toward the distractor in SCZ patients only [$t(26.86) = 2.110, p = 0.044$], confirmed by a *Diagnosis* \times *Distractor Location* [$F(2,100) = 4.032, GGe = 0.685, p = 0.036$] and a main *Distractor Location* effect [$F(2,100) = 24.705, GGe = 0.685, p < 0.001$]. Main effects of *Diagnosis* was not significant [$F(1,50) = 3.336, p = 0.074$].

We next tested the possibility that the response distribution in the SCZ group in the distal distraction condition might be bi-modal, with most responses centered on the target but some centered on the distractor. This may be due to a complete loss of the correct target memory trace rather than a slight bias toward the distractor. To rule out this possibility we performed Hartigans' dip test for uni-multimodality on the pattern of participant's responses. The test on all the responses in the SCZ group revealed that the distribution is not bimodal ($D = 0.008, p = 0.990$). Additionally, we tested for uni-multimodality within each SCZ subject specifically, to examine if there are any subjects that exhibit a distribution of responses in the distal distraction condition that shows a bimodal tendency. None of the tests were significant (all $Ds < 0.07$, all $ps > 0.15$). Further conservative tests, which exclude subjects with large report displacements, also support that the primary effect of the distal distractor induces a shift of the response distribution toward the distractor position. This result is consistent with model predictions, rather than a bi-modal response distribution with reports divided between being centered on the target or distractor (see Supplement for complete treatment and analysis).

3.4. Relationship between memory drift and distraction

A key model prediction suggests that increased response variability and increased distractor sensitivity stem from the same underlying mechanism—altered E/I balance via disinhibition, which broadens the WM representation. This mechanism should operate irrespective of diagnosis and therefore, to maximize power, we collapsed the analysis across all subjects. To test this, we quantified if WM drift rate in the delay condition (the slope of the relationship between delay duration and SD of angular displacement) predicts displacement toward the distractor in the distal condition (Fig. 5). As predicted, results revealed a significant relationship [$r = 0.40, t(48) = 3.002, p = 0.004, 2$ -tailed]: the greater the rate of WM drift, the greater the bias toward the distal distractor.

4. Discussion

While WM deficits in SCZ are well-established, it remains unknown how hypotheses of neural circuit dysfunction in SCZ map onto quantifiable WM deficits and patient behavior. To probe this question we developed a biophysically-grounded computational models of WM that generated mechanistic and testable behavioral predictions related to E/I imbalance (Murray et al., 2014). Our results verified two key model predictions: i) loss of inhibitory tuning in a cortical circuit model was associated with increased WM drift over time, and ii) loss of WM precision was linearly related to distortions in WM representation in the presence of distractors. Collectively, this ‘computational psychiatry’ study illustrates the interplay between biophysically-informed computational modeling of cortical microcircuit function and behavioral experimental results in SCZ patients.

4.1. Linking cellular-level hypotheses to behavioral deficits in schizophrenia via computation

The precise neural mechanisms underlying WM deficits in SCZ remain unknown, especially in relation to cellular-level pathophysiological models of the illness (Gonzalez-Burgos et al., 2010). Based on pharmacological and post-mortem work, several hypotheses proposed disrupted cortical E/I balance, possibly due to abnormal NMDA or GABA signaling between pyramidal cells and interneurons, which may lead to disinhibition (Lewis and Moghaddam, 2006; Nakazawa et al., 2012). Recently, we characterized the impact of synaptic disinhibition mechanisms in a computational microcircuit model of WM, which we adapted here to generate novel behavioral predictions (Murray et al., 2014). Specifically, we computationally modeled disruptions in E/I balance as a reduction of NMDAR current onto I-cells, effectively weakening feedback inhibition onto pyramidal cells. This synaptic manipulation caused elevated E/I ratio and generated a ‘disinhibited’ model regime. Consequently, background noise was more likely to influence firing throughout the activated neuronal pool that would otherwise get hyper-polarized by lateral inhibition. A broadened WM activity pattern absorbed more intrinsic noise and caused greater variability of WM responses. Furthermore, the circuit was more susceptible to distractor input that overlapped with the flanks of the WM representation, which would otherwise be filtered via intact lateral inhibition. Here we quantified model behavior, via population-level neural activity readout, to generate two novel predictions concerning WM drift and distractibility. In turn, we designed two behavioral experiments to quantify these effects in SCZ patients. This study demonstrates the possible interplay between microcircuit-level models of neural function (Wong and Wang, 2006) with behavioral experiments explicitly built around cellular-level models (Anticevic et al., 2015).

4.2. Schizophrenia is associated with elevated random WM drift over time

By manipulating the WM delay period duration, in conjunction with the continuous response measure of WM performance, we examined three behavioral patterns that point to possibly distinct computational mechanisms: i) Patients might exhibit, on some fraction of trials, complete loss of WM representations due to collapse of WM activity (Wang, 2001), or form ‘false’ memories of locations that were not encoded initially (Cano-Colino and Compte, 2012); ii) alternatively, patients may maintain WM representations, but exhibit greater WM

‘drift’ due to disinhibition (as the delay period increases. These dissociable outcomes yield distinct error patterns and point to dissociable neural mechanisms: i) The first two possibilities would imply that the cue representation in WM is lost, or obstructed by a randomly formed false memory, and thus the error pattern would be fully random, distributed uniformly across angles; ii) The alternative possibility, predicted by elevated E/I ratio, causes a broadened WM representation, which produces greater variability of responses over time. Experimentally arbitrating these possibilities is vital to understand underlying neural mechanisms of WM deficits in SCZ. Put differently, while the existing schizophrenia literature strongly supports lower working memory on average, the precise patterns of predicted errors by the computational model may be less intuitive and needed to be verified.

Experimental results were generally more in line with the second and largely inconsistent with the first scenario. We observed lower overall WM spatial precision in SCZ than HCS – an effect that increased over time. However, patients did not exhibit random responses on a higher percentage of trials, as would be predicted by the first scenario (Cano-Colino and Compte, 2012). Instead, they exhibited lower spatial precision at all delay durations, consistent with prior reports (Badcock et al., 2008). Moreover, variability of SCZ responses increased as a function of delay duration, as predicted by the model (Murray et al., 2014). Finally, as noted, to maintain quality control, we discarded trials where subjects failed to initiate a response or that were displaced $>90^\circ$ degrees in any angle from the cue position. Thus, it may be possible that patients exhibit very large angular displacement on some trials, which appears to be consistent with the possibility that patients might have lost WM representation. Two results argue against this possibility: i) Only 2% of all trials met these criteria (see Supplement) and ii) Even when examining results for this small percentage, there was no significant *Delay Duration* \times *Diagnosis* interaction (see Fig. S1), arguing against this effect changing over longer delays. This response pattern in patients largely implicates increased ‘drift’ in the WM representations, possibly due to elevated internal noise, resulting in enhanced degradation of WM representations over time. To summarize, if patients exhibit a loss of WM precision (but not a complete loss of the representation) then the error pattern would represent a greater scatter around the probe location. Conversely, if patients exhibit a complete loss of the WM representation then the error pattern would represent a random scatter at all locations. We observed a pattern of errors consistent with the first scenario but not the second, as predicted by the model.

A related question concerns the initial precision in WM encoding, which is reduced in SCZ (Lee and Park, 2005). Here we observed SCZ deficits in the pure encoding condition (i.e. 0-s delay condition), also consistent with prior studies (Lee and Park, 2005). Furthermore, deficits may occur in the final behavioral motor output, either due to inherent motor abnormalities (independent of WM), deficits in the transformation of the WM representation to a motor action plan, or lower speed of performance. We observed subtle motor deficits in SCZ when mapping the probe to the cue (i.e. no encoding or delay requirement). However, response variability increased with WM delay, providing evidence that WM deficits are compounded over time, consistent with predictions.

4.3. Schizophrenia is associated with increased range of distractor influence on WM

SCZ has been repeatedly associated with increased susceptibility to distraction (Anticevic et al., 2011; Anticevic et al., 2012b; Hahn et al., 2010; Oltmanns and Neale, 1975), especially during WM performance (Hahn et al., 2010), which may be associated with deficits in PFC function (Anticevic et al., 2011). However, microcircuit hypotheses of such distractibility during WM remain unexplored. We tested one specific model-derived prediction—the possibility that some distractor filtering deficits operate at the level of cortical microcircuits involved in WM maintenance. We modeled distractor effects by presenting a ‘distractor’ input at a distinct spatial location from that of the initial WM cue (Fig. 2C–D). We found that elevated E/I ratio (inducing disinhibition), effectively broadened the WM representation and expanded the ‘distractibility window’ in the model. The increase in errors was preferential, occurring when the distractor was presented at locations overlapping the flanks of WM representations. This effect emerges because of the overlap between WM representation and the distractor position, effectively biasing the response toward the distractor location. Experimental results revealed that SCZ patients exhibit a greater angular bias toward the distractor, especially for the distal condition. Interestingly, responses to proximal distractors were subtly biased in the direction away from the distractor, which is not captured by the circuit model and may involve a possible compensatory strategy that warrants further study. Critically, the model predicted that both WM drift and magnitude of angular distractor bias would be positively correlated irrespective of diagnosis, which was confirmed empirically. This final test implies that the same (or closely related) neural mechanism may govern both drift and distractibility during spatial WM.

4.4. Experimental limitations and future directions

Prior meta-analysis did not report an increase in effect size of group differences over time (Lee and Park, 2005), which seems to argue against differences in WM drift observed here. However, meta-analytic measure of effect size is affected both by the mean difference and by variability for each group (Cohen’s d = difference in means, divided by pooled SD). We observed that both mean differences and variability (individual differences within groups) increase with longer delays. Therefore, quantifying the between-group effect via a formal effect size across delay durations that are pooled from distinct studies cannot, by definition, statistically fully establish if the magnitude of the WM drift is in fact greater in SCZ as the delay duration increases. Put simply, the relative effect size remains unchanged due to elevated variance across both groups. This statistical situation that ‘masks’ the possibility of observing greater WM drift over time if examined across studies as both the between-group effect and variability change for a given study. Therefore, this question requires an explicit parametric experimental test in the same sample of subjects – a key goal of this study. Furthermore, the absence of a delay effect in the reported meta-analyses is further complicated because most reported studies use different modalities (e.g. verbal rehearsal, colors, shapes), which may involve distinct neural mechanisms and that produce different sensitivities to delay duration (Gold et al., 2010; Lee and Park, 2005; Rothmay et al., 2007).

Also, we did not parametrically vary the distractor location at multiple positions given the excessive burden/fatigue placed on the patients from such a long experiment. Future studies should include multiple distractor distances. In particular, the model predicts that beyond a

certain cue-distractor separation (i.e. outside the distractibility window), neither SCZ nor HCS subjects will exhibit bias toward the distractor location (Herwig et al., 2010). This remains to be experimentally tested.

An important limitation is the absence of eye-tracking. While we did not explicitly track eye position, an experimenter ensured that subjects continuously attended and maintained fixation. That said, patients could perhaps use eye position as a compensatory strategy to remember certain spatial locations. However, if that were the case we would not expect a differential pattern of errors to emerge across delay and distracter conditions. Put simply, use of eye position as a guide would work against (not toward) reported delay effects in either group. Therefore, one might argue that by not enforcing stringent eye fixation established an even stronger case for presented experimental results. Moreover, precisely because we controlled expectation of delay duration neither controls nor patients could improve their WM precision by anticipating trial length (and therefore using eye position as a strategy). Nevertheless, future studies should explicitly replicate and generalize these effects using eye tracking to examine if eye position may interact in any way with presented effects.

Of note, we designed our paradigm to probe WM precision deficits and not abnormalities in WM stability or buffer size (Erickson et al., 2014; Gold et al., 2010). There are compelling experimental data to suggest degradations in both aspects of WM in SCZ. Thus, present findings serve predominately to implicate loss of WM precision over time rather than to rule out other forms of WM impairment (Erickson et al., 2014). Also, we focused on one modality in a very defined experimental framework – namely visuo-spatial WM. Prior studies with healthy controls have shown that visuo-spatial WM exhibits delay-dependent drift (Ploner et al., 1998; White et al., 1994), and similarity-dependent distractibility (Herwig et al., 2010). Present findings reveal that both effects are altered in SCZ. It remains to be determined if these effects generalize to other WM-related representations beyond visuo-spatial location. For instance, Gold and colleagues studied SCZ effects using a multiple-item change-detection task based on color WM with delays of 1 or 4 s. They found no main effect of delay duration on WM precision, and no group interaction for SCZ vs. HCS (Gold et al., 2010). Future studies may seek to manipulate color, semantic or even facial similarity (via face-morphing software) to systematically probe whether the cue-distractor similarity effects generalize across experimental contexts.

Lastly, as is the case with most clinical studies, we cannot definitely rule out the effects of *anti*-psychotic medication or more long-term poly-pharmacy. Future studies should examine reported effects in fully unmedicated or prodromal samples and incorporate additional diagnostic categories (e.g. bipolar illness) to examine cross-diagnostic generalizability.

4.5. Model limitations and future directions

While the core qualitative predictions of the model were largely confirmed, our analyses revealed aspects of WM behavior that are not fully captured by the current model, which should be addressed in future studies. For the model results presented here, we used the same parameters in our prior modeling study (. We did not perform any explicit fitting of model parameters to capture empirical data. Therefore, the model generates qualitative predictions for effects of altered E/I balance, rather than quantitative fits to presented

empirical data. Improving the quantitative fit to empirical data is possible through adjustment of model parameters, although rigorous fitting is impractical due to the computational cost of model simulation. For instance, the WM drift rate can be reduced by increasing the number of neurons in the network (. The magnitude of distractor effects can be scaled by changing the strength of the stimulus (. The bump width, which affects the distractibility window, can be altered through synaptic connectivity parameters (. Moreover, post-hoc ‘fine-tuning’ of the model to better fit the data would be circular. Thus, here we explicitly generated qualitative yet parsimonious and testable predictions.

Specifically, the model was developed to explore deficits in WM maintenance under E/I imbalance (. In addition to WM maintenance deficits, SCZ subjects showed deficits in motor responses and in WM encoding. These components of the WM task were not explicitly included in the model. Future modeling studies could address these aspects by modeling the upstream input to the WM circuit and the downstream readout for motor response. Additionally, the current model was designed for single-item visuospatial WM, matching our task design. Prior modeling studies have extended this framework to study multiple-item visuo-spatial WM (. Future experimental and modeling studies should explore how the clinical phenomena studied here extend to multiple-item WM.

The primary qualitative discrepancy between model behavior and empirical data is that both subject groups displayed a repulsive bias away from the proximal distractor. To our knowledge, this class of WM circuit models is not capable of showing a pattern of distractor effects, with repulsion at short distances and attraction at medium distances. For instance, in a model of multiple-item WM, items show attraction at short distances and repulsion at medium distances (. Future experimental and modeling studies are needed to address whether the repulsion at short distances we observed is a task-dependent compensatory strategy, or a general feature of WM whose mechanism may occur within WM circuits.

5. Conclusion

We report two experimental tests in SCZ patients of qualitative predictions generated by a biophysically-informed cortical microcircuit computational model: i) elevated WM drift over time, and ii) broadened window of WM distractibility. In doing so, this ‘computational psychiatry’ study provides a provisional clinical translation of cellular-level computational modeling to behavioral experiments in SCZ. Such findings can in turn be applied to neuroimaging data or to other symptoms to test mechanistic hypotheses and guide more precise development of therapeutics for cognitive deficits in SCZ (Murray et al., 2014).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial support for this study was provided by the National Institutes of Health (NIH) grants DP5OD012109-03 (PI: AA), R01MH103831-01 (PI: VHS), and R01-MH062349 (to X-J.W. & J.D.M.), the National Alliance for Research on Schizophrenia and Depression (PI: AA), the Fulbright Foundation (AS), and the Yale Center for Clinical Investigation (PI: AA).

Role of the funding source

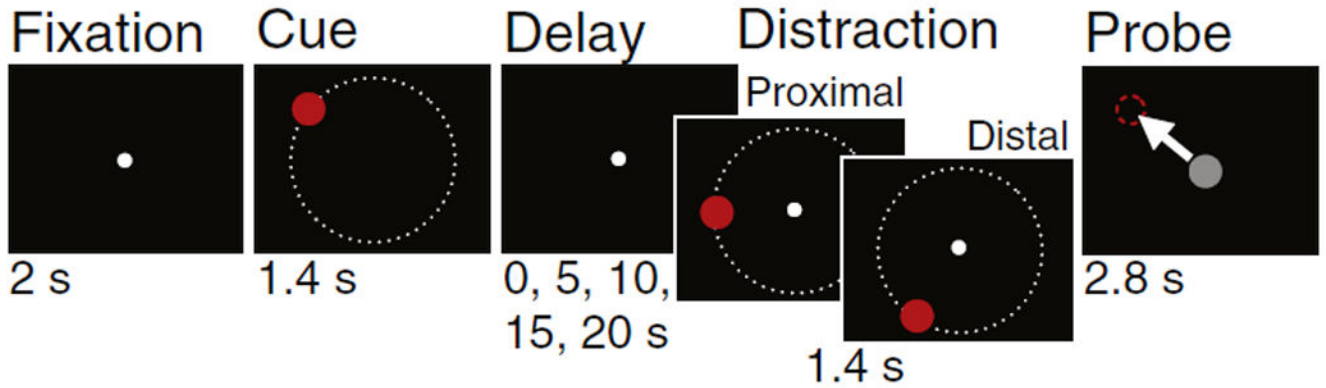
The funding source had no further role in the current study with regard to data collection, data analysis and interpretation of findings or in manuscript preparation and the submission decision.

References

- Andreasen, NC. The Scale for the Assessment of Negative Symptoms (SANS). University of Iowa; Iowa City: 1983a.
- Andreasen, NC. The Scale for the Assessment of Positive Symptoms (SAPS). University of Iowa; Iowa City: 1983b.
- Andreasen NC, Pressler M, Nopoulos P, Miller D, Ho B-C. Antipsychotic dose equivalents and dose-years: a standardized method for comparing exposure to different drugs. *Biol Psychiatry*. 2010; 67(3):255–262. [PubMed: 19897178]
- Anticevic A, Repovs G, Corlett PR, Barch DM. Negative and non-emotional interference with visual working memory in schizophrenia. *Biol Psychiatry*. 2011; 70(12):1159–1168. [PubMed: 21861986]
- Anticevic A, Gancsos M, Murray JD, Repovs G, Driesen NR, Ennis DJ, Niciu MJ, Morgan PT, Surti TS, Bloch MH, Ramani R, Smith M, Wang XJ, Krystal JH, Corlett PR. NMDA receptor function in large-scale anti-correlated neural systems with implications for cognition and schizophrenia. *Proc Natl Acad Sci U S A*. 2012a; 109(41):16720–16725. [PubMed: 23012427]
- Anticevic A, Repovs G, Krystal JH, Barch DM. A Broken Filter: Prefrontal Functional Connectivity Abnormalities in Schizophrenia During Working Memory Interference. *Schizophr Res Epub ahead of print*. 2012b
- Anticevic A, Cole MW, Repovs G, Savic A, Driesen NR, Yang G, Cho YT, Murray JD, Glahn DC, Wang X-J, Krystal JH. Connectivity, pharmacology, and computation: toward a mechanistic understanding of neural system dysfunction in schizophrenia. *Frontiers in Psychiatry*. 2013a; 4(169)
- Anticevic A, Repovs G, Barch DM. Working memory encoding and maintenance deficits in schizophrenia: neural evidence for activation and deactivation abnormalities. *Schizophr Bull*. 2013b; 39(1):168–178. [PubMed: 21914644]
- Anticevic A, Murray JD, Barch DM. Bridging levels of understanding in schizophrenia through computational modeling. *Clinical Psychological Science*. 2015; 3(3):433–459. [PubMed: 25960938]
- Badcock JC, Badcock DR, Read C, Jablensky A. Examining encoding imprecision in spatial working memory in schizophrenia. *Schizophr Res*. 2008; 100(1–3):144–152. [PubMed: 17804202]
- Barch DM, Ceaser A. Cognition in schizophrenia: core psychological and neural mechanisms. *Trends Cogn Sci*. 2012; 16(1):27–34. [PubMed: 22169777]
- Belforte JE, Zsiros V, Sklar ER, Jiang Z, Yu G, Li Y, Quinlan EM, Nakazawa K. Postnatal NMDA receptor ablation in corticolimbic interneurons confers schizophrenia-like phenotypes. *Nat Neurosci*. 2010; 13(1):76–83. [PubMed: 19915563]
- Cano-Colino M, Compte A. A computational model for spatial working memory deficits in schizophrenia. *Pharmacopsychiatry*. 2012; 45(Suppl 1):S49–S56. [PubMed: 22565235]
- Compte A, Brunel N, Goldman-Rakic PS, Wang XJ. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex*. 2000; 10(9):910–923. [PubMed: 10982751]
- Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Med*. 2013; 11:126. [PubMed: 23672542]
- Driesen NR, Leung HC, Calhoun VD, Constable RT, Gueorguieva R, Hoffman R, Skudlarski P, Goldman-Rakic PS, Krystal JH. Impairment of working memory maintenance and response in schizophrenia: functional magnetic resonance imaging evidence. *Biol Psychiatry*. 2008; 64(12):1026–1034. [PubMed: 18823880]
- Driesen NR, McCarthy G, Bhagwagar Z, Bloch MH, Calhoun VD, D'Souza DC, Gueorguieva R, He G, Leung HC, Ramani R, Anticevic A, Suckow RF, Morgan PT, Krystal JH. The impact of NMDA receptor blockade on human working memory-related prefrontal function and connectivity. *Neuropsychopharmacology*. 2013; 38(13):2613–2622. [PubMed: 23856634]

- Durstewitz D, Seamans JK. The computational role of dopamine D1 receptors in working memory. *Neural Netw.* 2002; 15(4–6):561–572. [PubMed: 12371512]
- Durstewitz D, Kelc M, Gunturkun O. A neurocomputational theory of the dopa-minergic modulation of working memory functions. *J Neurosci.* 1999; 19(7):2807–2822. [PubMed: 10087092]
- Erickson M, Hahn B, Leonard C, Robinson B, Luck S, Gold J. Enhanced vulnerability to distraction does not account for working memory capacity reduction in people with schizophrenia. *Schizophrenia Research: Cognition.* 2014; 1(3):149–154. [PubMed: 25705590]
- First, MB., Spitzer, RL., Gibbon, M., Williams, JBW. *Structured Clinical Interview for the DSM-IV-TR Axis I Disorders.* American Psychiatric Press; Washington, D.C.: 2001.
- Forbes NF, Carrick LA, McIntosh AM, Lawrie SM. Working memory in schizophrenia: a meta-analysis. *Psychol Med.* 2009; 39(6):889–905. [PubMed: 18945379]
- Funahashi S, Bruce CJ, Goldman-Rakic PS. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J Neurophysiol.* 1989; 61(2):331–349. [PubMed: 2918358]
- Glahn DC, Bearden CE, Bowden CL, Soares JC. Reduced educational attainment in bipolar disorder. *J Affect Disord.* 2006; 92:309–312. [PubMed: 16524631]
- Gold JM, Hahn B, Zhang WW, Robinson BM, Kappenman ES, Beck VM, Luck SJ. Reduced capacity but spared precision and maintenance of working memory representations in schizophrenia. *Arch Gen Psychiatry.* 2010; 67(6):570–577. [PubMed: 20530006]
- Goldman-Rakic PS. Working memory dysfunction in schizophrenia. *Journal of neuropsychiatry.* 1994; 6(4):348–357.
- Goldman-Rakic PS. Cellular basis of working memory. *Neuron.* 1995; 14:477–485. [PubMed: 7695894]
- Gonzalez-Burgos G, Hashimoto T, Lewis DA. Alterations of cortical GABA neurons and network oscillations in schizophrenia. *Curr Psychiatry Rep.* 2010; 12(4):335–344. [PubMed: 20556669]
- Green MF. Cognitive impairment and functional outcome in schizophrenia and bipolar disorder. *The Journal of clinical psychiatry.* 2006; 67(10):e12. [PubMed: 17107235]
- Hahn B, Robinson BM, Kaiser ST, Harvey AN, Beck VM, Leonard CJ, Kappenman ES, Luck SJ, Gold JM. Failure of schizophrenia patients to overcome salient distractors during working memory encoding. *Biol Psychiatry.* 2010; 68(7):603–609. [PubMed: 20570242]
- Herwig A, Beisert M, Schneider WX. On the spatial interaction of visual working memory and attention: evidence for a global effect from memory < guided saccades. *J Vis.* 2010; 5(8):1–10.
- Kalkstein, S., Hurford, I., Gur, R. *Neurocognition in Schizophrenia.* Swerdlow, NR., editor. Springer Berlin Heidelberg; Berlin, Heidelberg: 2010. p. 373-390.
- Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull.* 1987; 13:261–276. [PubMed: 3616518]
- Kotermanski SE, Johnson JW. Mg²⁺ imparts NMDA receptor subtype selectivity to the Alzheimer's drug Memantine. *J Neurosci.* 2009; 29(9):2774–2779. [PubMed: 19261873]
- Krystal JH, D'Souza DC, Mathalon D, Perry E, Belger A, Hoffman R. NMDA receptor antagonist effects, cortical glutamatergic function, and schizophrenia: toward a paradigm shift in medication development. *Psychopharmacology.* 2003a; 169(3–4):215–233. [PubMed: 12955285]
- Krystal JH, D'Souza DC, Mathalon D, Perry E, Belger A, Hoffman R. NMDA receptor antagonist effects, cortical glutamatergic function, and schizophrenia: toward a paradigm shift in medication development. *Psychopharmacology.* 2003b; 169(3–4):215–233. [PubMed: 12955285]
- Lee J, Park S. Working memory impairments in schizophrenia: a meta-analysis. *J Abnorm Psychol.* 2005; 114(4):599–611. [PubMed: 16351383]
- Lewis DA, Moghaddam B. Cognitive dysfunction in schizophrenia: convergence of gamma-aminobutyric acid and glutamate alterations. *Arch Neurol.* 2006; 63(10):1372–1376. [PubMed: 17030651]
- Lewis DA, Hashimoto T, Volk DW. Cortical inhibitory neurons and schizophrenia. *Nat Rev Neurosci.* 2005; 6(4):312–324. [PubMed: 15803162]
- Macdonald AW, Chafee MV. Translational and developmental perspective on N-methyl-D-aspartate synaptic deficits in schizophrenia. *Dev Psychopathol.* 2006; 18(3):853–876. [PubMed: 17152404]

- MacDonald AW 3rd, Becker TM, Carter CSn. Functional magnetic resonance imaging study of cognitive control in the healthy relatives of schizophrenia patients. *Biol Psychiatry*. 2006; 60(11): 1241–1249. [PubMed: 16945345]
- Marin O. Interneuron dysfunction in psychiatric disorders. *Nat Rev Neurosci*. 2012; 13(2):107–120. [PubMed: 22251963]
- Metzak PD, Riley JD, Wang L, Whitman JC, Ngan ETC, Woodward TS. Decreased efficiency of task-positive and task-negative networks during working memory in schizophrenia. *Schizophr Bull*. 2011; 38(4):803–813. [PubMed: 21224491]
- Murray JD, Anticevic A, Corlett PR, Gancsos M, Krystal JH, Wang X-J. Linking microcircuit dysfunction to cognitive impairment: effects of disinhibition associated with schizophrenia in a cortical working memory model. *Cereb Cortex*. 2014; 24(4):859–872. [PubMed: 23203979]
- Nakazawa K, Zsiros V, Jiang Z, Nakao K, Kolata S, Zhang S, Belforte JE. GABAergic interneuron origin of schizophrenia pathophysiology. *Neuropharmacology*. 2012; 62(3):1574–1583. [PubMed: 21277876]
- Nuechterlein KH, Subotnik KL, Green MF, Ventura J, Asarnow RF, Gitlin MJ, Yee CM, Gretchen-Doorly D, Mintz J. Neurocognitive predictors of work outcome in recent-onset schizophrenia. *Schizophr Bull*. 2011; 37(suppl 2):S33–S40. [PubMed: 21860045]
- Oltmanns TF, Neale JM. Schizophrenic performance when distractors are present: attentional deficit or differential task difficulty? *J Abnorm Psychol*. 1975; 84(3):205–209. [PubMed: 1133249]
- Ploner CJ, Gaymard B, Rivaud S, Agid Y, Pierrot-Deseilligny C. Temporal limits of spatial working memory in humans. *Eur J Neurosci*. 1998; 10(2):794–797. [PubMed: 9749746]
- Quee PJ, Eling PA, van der Heijden FM, Hildebrandt H. Working memory in schizophrenia: a systematic study of specific modalities and processes. *Psychiatry Res*. 2010; 185(1–2):54–59. [PubMed: 20510466]
- Rao SG, Williams GV, Goldman-Rakic PS. Destruction and creation of spatial tuning by disinhibition: GABA (A) blockade of prefrontal cortical neurons engaged by working memory. *J Neurosci*. 2000; 20(1):485–494. [PubMed: 10627624]
- Rothmay C, Baumann O, Endestad T, Rutschmann SM, Magnussen S, Greenlee MW. Dissociation of neural correlates of verbal and non-verbal visual working memory with different delays. *Behav Brain Funct*. 2007; 3:56. [PubMed: 17958919]
- Silver H, Feldman P, Bilker W, Gur RC. Working memory deficit as a core neuropsychological dysfunction in schizophrenia. *Am J Psychiatry*. 2003; 160(10):1809–1816. [PubMed: 14514495]
- Van Snellenberg JX, de Candia T. Meta-analytic evidence for familial coaggregation of schizophrenia and bipolar disorder. *Arch Gen Psychiatry*. 2009; 66(7):748–755. [PubMed: 19581566]
- Wang XJ. Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci*. 2001; 24(8):455–463. [PubMed: 11476885]
- Wang X-J. Toward a prefrontal microcircuit model for cognitive deficits in schizophrenia. *Pharmacopsychiatry*. 2006; 39(Suppl 1):S80–S87. [PubMed: 16508903]
- Wang X-J. Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol Rev*. 2010; 90(3):1195–1268. [PubMed: 20664082]
- Wang X-J, Tegnér J, Constantinidis C, Goldman-Rakic PS. Division of labor among distinct subtypes of inhibitory neurons in a cortical microcircuit of working memory. *Proc Natl Acad Sci U S A*. 2004; 101(5):1368–1373. [PubMed: 14742867]
- Wang M, Yang Y, Wang CJ, Gamo NJ, Jin LE, Mazer JA, Morrison JH, Wang XJ, Arnsten AF. NMDA receptors subserve persistent neuronal firing during working memory in dorsolateral prefrontal cortex. *Neuron*. 2013; 77(4):736–749. [PubMed: 23439125]
- White JM, Sparks DL, Stanford TR. Saccades to remembered target locations: an analysis of systematic and variable errors. *Vis Res*. 1994; 34(1):79–92. [PubMed: 8116271]
- Wong KF, Wang XJ. A recurrent network mechanism of time integration in perceptual decisions. *J Neurosci*. 2006; 26(4):1314–1328. [PubMed: 16436619]

**Fig. 1.**

Working memory paradigm. Subjects were asked to remember the position of circles ($d = 125\text{px}$) that were presented at 20 pseudo-randomly chosen angles along a hidden radial grid ($r = 415\text{px}$). This was done to mimic the ‘ring’ structure of the biophysically-based computational model motivating the design (Compte et al., 2000). After a delay subjects used a high-sensitivity joystick to indicate the remembered location, providing a parametric index of accuracy (as opposed to a forced-choice yes/no answer). The delay period varied parametrically such that subjects were asked to hold the location in memory for 0 s (i.e. immediate recall), 5 s, 10 s, 15 s, or 20s (60–20 trials). Subjects also completed a series of trials that contained a distractor, such that an additional circle appeared that subjects did not have to remember. During the distractor task the delay period was always 10s and the distractor appeared in the middle of the delay (after 4.3 s). There were two types of distractors, appearing at either 20° (proximal distractors) or 50° (distal distractors) from the original cue position (40 trials each). Lastly, subjects completed a control motor task (not shown) where cue circle and probe circle appeared simultaneously, requiring subjects to place the probe on top of the cue circle which necessitated a motor response but no WM maintenance or recall (20 trials).

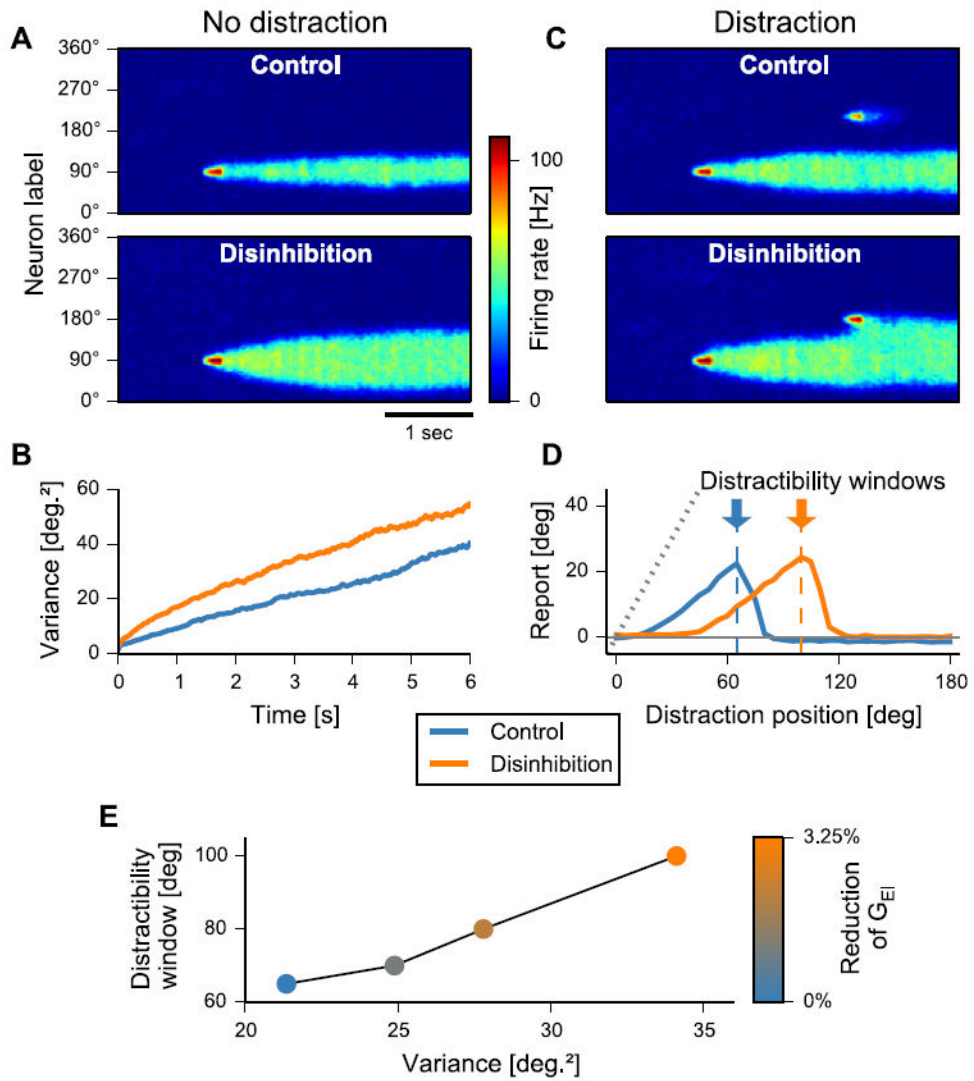


Fig. 2. Computational modeling results and experimental predictions. A. In the cortical circuit model of spatial WM, a brief stimulus excites a subset of neurons, which encode the stimulus across the delay through a pattern of persistent activity that is shaped by recurrent excitation and lateral inhibition. Disinhibition is implemented via reduction of NMDAR conductance from pyramidal cells onto inhibitory interneurons (G_{EI}) (Murray et al., 2014). Disinhibition results in a broadened WM representation. B. The variability of WM report increases with delay duration. Disinhibition increases the rate at which variability increases, constituting a deficit in WM maintenance. C. Distractors are modeled as an intervening stimulus inputs during the delay. Under the control condition, there is no overlap between distractor and WM representations, and the report is unperturbed by the distractor. Under disinhibition, there is overlap and the report is shifted toward the distractor location. D. The mean WM report is shifted in the direction of the distractor, with dependence on the angular separation between distractor and cue. Disinhibition shows a larger distractibility window, i.e. the separation with maximal impact on WM report. E. The distractibility window and

WM variability (chosen here at 3-s delay, no-distractor condition) are correlated as both smoothly increase with the strength of disinhibition (reduction of G_{EI}). Panels A-D adapted from (Murray et al., 2014) with permission.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

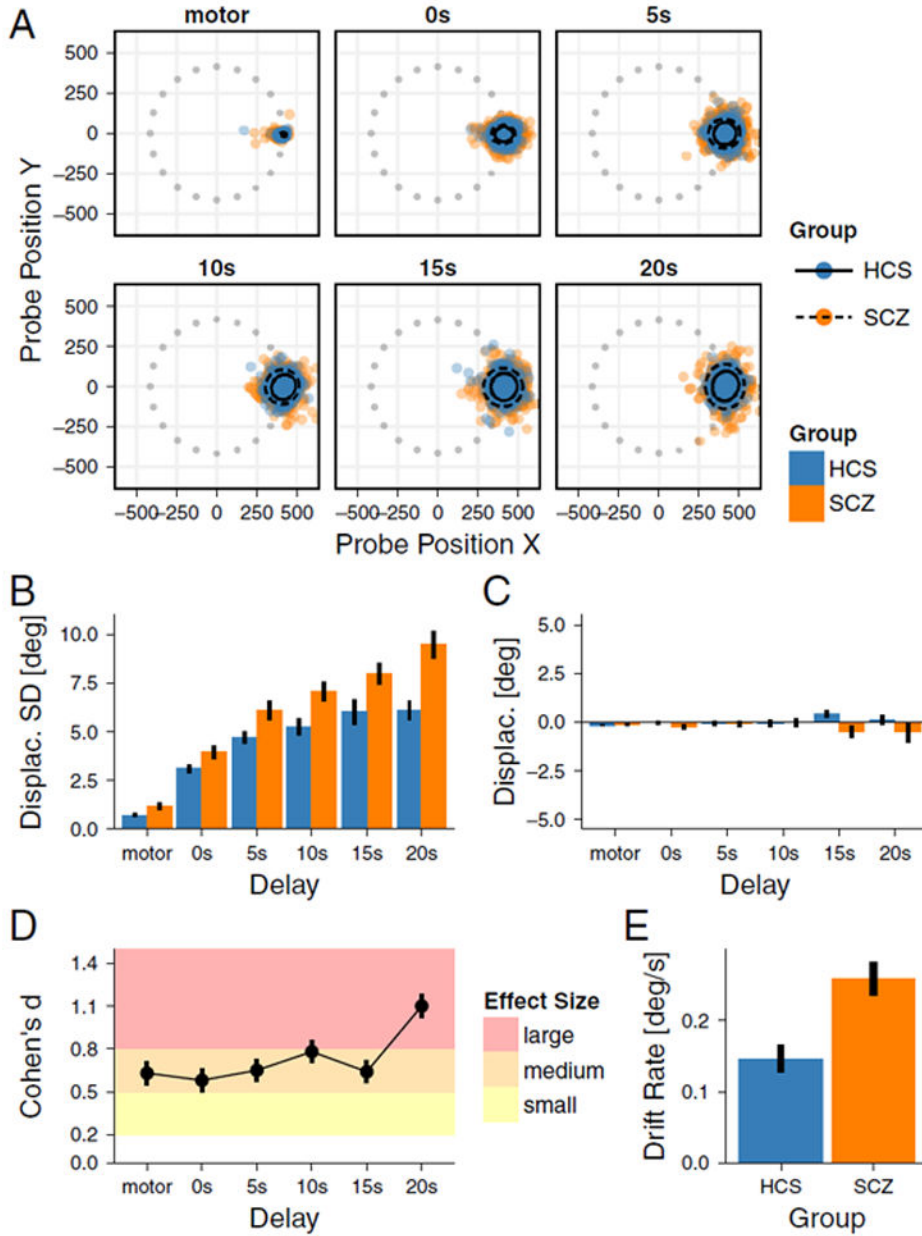


Fig. 3. Effects of Delay Duration on SWM Performance

Effects of delay duration on working memory drift. A. Each panel indicates positions of responses on the screen at different delays. Results are rotated such that every cue is presented at 0° angle to facilitate visual inspection. Gray dots indicate the pseudo-random positions of targets on the screen before rotation for analysis purposes. 95% confidence ellipses are shown around the mean response pattern for each group. Responses spread with increasing delay but remain centered on the cue position. B. Average standard deviation of response errors. Average variability is increased for the SCZ group compared to HCS and this difference increased with *Delay Duration*, indicating lower WM precision. C. As expected, average angular displacement of the responses is unaffected by either *Delay Duration* or *Diagnosis*, indicating no directional bias in the response pattern, as predicted by

the model. D. Effect sizes for between-group differences shown in B are in the medium to large range and remain relatively stable as delay increases, with the only notable increase at the longest delay. E. Average drift rate across the five delays shown in B is increased for the SCZ group compared to HCS. Error bars show ± 1 standard error of the mean.

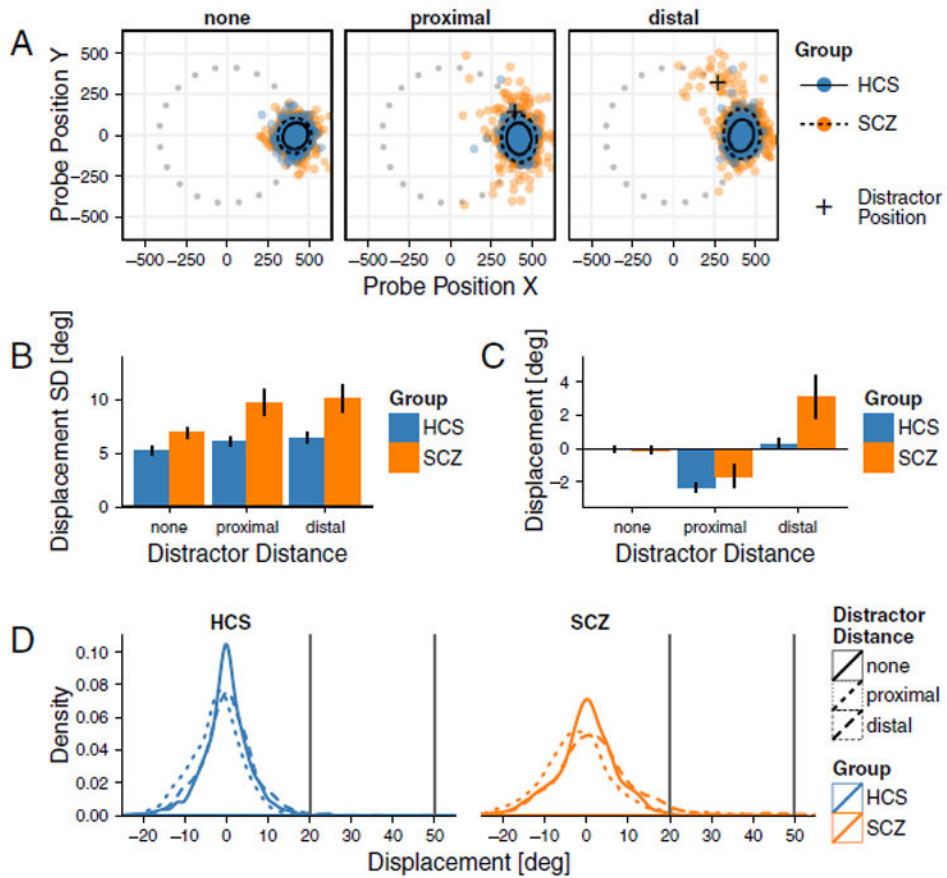


Fig. 4. Effects of Distractor Distance on Spatial Working Memory Performance

Effects of distractor position on working memory performance. **A.** Each panel indicates positions of responses on the screen rotated to angle 0 with 95% confidence ellipses and gray dots marking target positions before rotation, as in Fig. 3. Responses for the SCZ group show spreading in the presence of distractors, displaced away from cues toward distractors (crosses), in the distal distraction condition. **B.** Average standard deviation of response errors. Variability is increased in the SCZ group compared to HCS and in both groups in conditions with distractors. **C.** Distractors cause angular displacement of responses. While proximal distractors bias both groups to move away from the distractor, distal distractors only affect the SCZ group by biasing their responses toward the distractor location. Error bars show ± 1 standard error of the mean. **D.** Density plots for HCS (left) and SCZ (right) illustrating the shifts in responses. Critically, these density plots are not consistent with bi-modal patterns of responses, but rather indicate subtle general right distribution shift for SCZ following distal distractors.

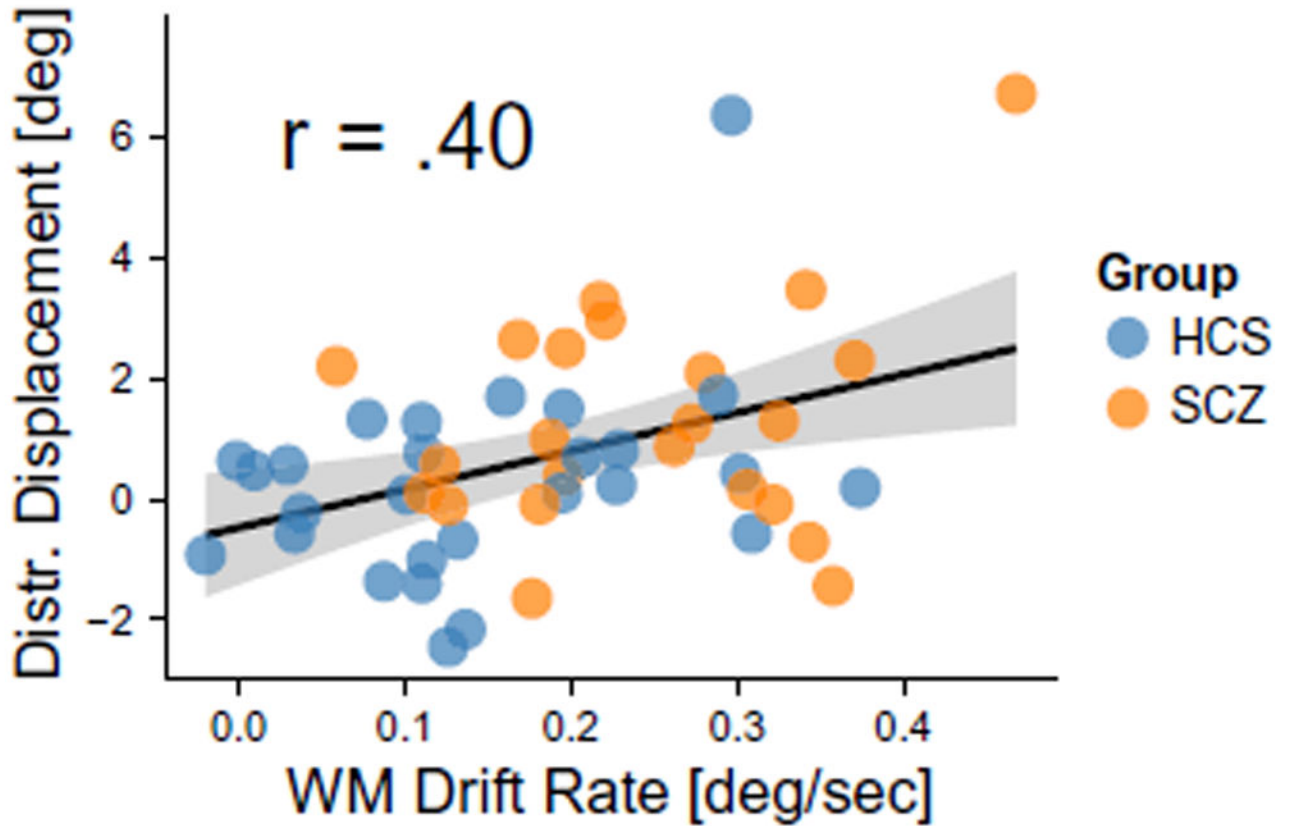


Fig. 5.

Positive relationship between working memory drift and distractibility. Linear regression was used to calculate the slope of working memory drift over time for each subject. The resulting regression slopes (β values) for each subject were correlated with angular displacement in the distal distractor condition. The black line shows the positive correlation between the two—the more WM drift with increasing delay, the greater the distractor bias—suggesting possible shared mechanisms (Fig. 2E) (Murray et al., 2014). Here we collapsed the analysis across both SCZ and HCS to maximize power given that spatial WM is a continuous measure and a construct highly consistent with a ‘dimensional’ perturbation (Cuthbert and Insel, 2013).

Table 1

Demographics. SAPS/SANS, Scale for the Assessment of Positive and Negative Symptoms (Andreasen, 1983a, b); CPZ, chlorpromazine equivalents, PANSS, Positive and Negative Symptom Scale (Kay et al., 1987). No subjects met criteria for current alcohol/drug use or dependence. CPZ equivalents were calculated using recently revised approaches (Andreasen et al., 2010) with 23/27 patients receiving medication. Of note, the SAPS and PANSS Positive scores of patients were highly correlated at $r = 0.68$, $t(23) = 4.396$, $p < 0.001$, 2-tailed, and the SANS and PANSS Negative Scores were correlated at $r = 0.78$, $t(23) = 5.962$, $p < 0.001$, 2-tailed, illustrating the high consistency of clinical ratings across instruments measuring similar clinical constructs.

Demographic characteristic	Controls (N = 28)		Patients (N = 27)		Between-group statistic	
	M	SD	M	SD	T value/Chi-Square	P Value (2-tailed)
Age (in years)	25.38	2.82	28.44	7.86	1.915	0.064
Gender (% male)	71	n/a	89	n/a	1.641	0.200
Paternal education (in years)	15.29	3.69	13.44	3.18	-1.956	0.056
Maternal education (in years)	15.14	3.02	13.92	2.78	-1.525	0.134
Paternal SES	31.79	11.64	24.78	14.18	-1.900	0.064
Maternal SES	26.25	12.07	23.13	13.29	-0.869	0.389
Subject's education (in years)	16.54	2.22	13.22	2.15	-5.619	0.000
Handedness (% right)	93	n/a	82	n/a	0.741	0.389
WRAT-3	51.75	5.30	47.32	5.72	-2.913	0.005
IQ Verbal	121.93	16.11	101.25	19.24	-4.162	0.000
IQ Non-verbal (Matrix)	115.36	12.39	99.40	14.53	-4.276	0.000
Medication (CPZ equivalents)	n/a	n/a	310.58	172.19	n/a	n/a
PANSS Positive	7.64	0.62	20.36	4.64	13.582	0.000
PANSS Negative	8.18	0.9	21.56	6.02	11.001	0.000
PANSS General	16.75	1.08	39.44	7.37	15.257	0.000
Mean SAPS Global Item Score	0.02	0.07	2.19	0.97	11.216	0.000
Mean SANS Global Item Score	0.19	0.25	2.85	0.86	14.917	0.000
Disorganization	0.25	0.44	6.48	2.58	11.902	0.000
Poverty	0.57	1.03	11.8	3.54	15.306	0.000
Reality Distortion	0.04	0.19	4.76	2.65	8.893	0.000