Running title: Capacity and precision of visual working memory in schizophrenia

Atypically larger variability of resource allocation accounts for visual working memory deficits in schizophrenia

Yijie Zhao[1], Xuemei Ran[1], Li Zhang[1], Ruyuan Zhang[2*], Yixuan Ku[1,3*]

[1]The Shanghai Key Lab of Brain Functional Genomics, Shanghai Changning-ECNU Mental Health Center, School of Psychology and Cognitive Science, East China Normal University, Shanghai, 200062 China

[2]Center for Magnetic Resonance Research, Department of Neuroscience, University of Minnesota, Minneapolis, MN 55455 USA

[3]NYU-ECNU Institute of Brain and Cognitive Science, NYU Shanghai and Collaborative Innovation Center for Brain Science, Shanghai, 200062 China

* co-senior author

Correspondence:

Ruyuan Zhang ruyuanzhang@gmail.com

Yixuan Ku yk1616@nyu.edu

Authors' contributions. RZ, YK &YZ developed research idea and study concepts; RZ, YK & YZ design the research; XR and LZ collected the data; RZ performed the data analyses and modeling; RZ & YZ wrote the manuscript.

Acknowledgement. We thank Zheng Ma, Ting Qian and Haojiang Ying for their invaluable comments on the manuscript. This work was supported by the National Social Science Foundation of China (17ZDA323), the National Key Fundamental Research Program of China (2013CB329501), the Major Program of Science and

29 Technology Commission Shanghai Municipal (17JC1404100), the Fundamental
30 Research Funds for the Central Universities (2018ECNU-QKT015), and the NYU-
31 ECNU Institute of Brain and Cognitive Science at NYU (YK).

32
33

ABSTRACT

Atypical performance of people with schizophrenia (PSZ) in visual working memory (VWM) has long been attributed to decreased capacity compared with healthy control subjects (HCS). This notion, however, largely ignores the effects of other VWM components, such as precision, on behavioral performance. Here, we measured the performance of 60 PSZ and 61 HCS in a classical delay-estimation task and disentangled the contribution of various VWM components by thoroughly comparing several influential computational models of VWM. Surprisingly, none of the models suggest group differences in memory capacity and in memory resources across set size levels–two diagnostic features of VWM. Notably, we find that the model assuming variable precision (VP) across items and trials is the best model to explain the performance of both groups, indicating the two groups employ the qualitatively same internal process in VWM. According to the VP model, PSZ only exhibit abnormally larger variability of allocating memory resources rather than resources *per se*. These results directly challenge the widely accepted decreased-capacity theory and establish the first time to our knowledge that the elevated resource allocation variability is the major determinant of the atypical VWM behavior in schizophrenia. Finally, we show that individual differences in the resource allocation variability predict variation of symptom severity in PSZ, highlighting its functional relevance to schizophrenic pathology. Taken together, our findings provide a novel account for the VWM deficits in schizophrenia and shed a new light on the utility of generative computational models to characterize mechanisms of mental deficits in clinical neuroscience.

Keywords:

Schizophrenia, Visual working memory, Memory precision, Generative model

SIGNIFICANCE STATEMENT

Working memory deficits in schizophrenia have long been thought to arise from abnormally decreased capacity. We revisit this topic by leveraging the most recent advances in research on visual working memory. Combining a classical visual working memory paradigm and generative computational modeling, we challenge the conventional decreased-capacity theory and demonstrate that the atypically larger variability of distributing memory resources across items and trials accounts for behavioral deficits observed in schizophrenia. The current study offers a new perspective for future studies aiming for characterizing the diagnostic pathology of schizophrenia.

INTRODUCTION

Schizophrenia is a severe mental disorder accompanied by a range of dysfunctions in perceptual and cognitive behavior, among which working memory deficit is considered as a core feature (1–4). Working memory refers to the ability to temporally store and manipulate information in order to guide appropriate behavior, and it has been shown to link with a broad range of other brain functions, including perception, attention, problem-solving and executive control (5–8). Dysfunctions in working memory therefore might cascade into multiple mental processes, causing a wide spectrum of negative consequences (2, 3, 9).

A well-established finding in lab-based experiments is that people with schizophrenia (PSZ) exhibit worse performance than healthy control subjects (HCS) in visual working memory (VWM) tasks (2). This phenomenon has long been attributed to decreased VWM *capacity* in PSZ (2, 10, 11), a theory supported by previous studies using various VWM tasks, including the 'span' tasks (e.g., digit span, spatial span, verbal span) (12, 13), the N-back task (14–16), the delayed-response task (17–19), the change detection task (20–24) and the delay-estimation task (25–27). Despite the considerable differences across tasks, almost all previous studies converged to the same conclusion that decreased-capacity is the major cause for the VWM deficits in PSZ.

In another line of research, people have increasingly recognized memory *precision* as a pivotal determinant of VWM performance (28). Precision indicates the amount of memory resources assigned to individual items–a larger amount of resources produce higher memory precision. At the neural level, low perceptual precision might arise from either the intrinsic noise in neural responses (29–31) or the fluctuations of internal cognitive factors (e.g., arousal, attention) (31, 32). Atypically increased variability in both behavioral and neural responses has been discovered in patients with mental diseases such as autism spectrum disorder (33, 34), dyslexia (35), ADHD (36). These theoretical and empirical studies raise the possibility that PSZ and HCS might differ in memory precision rather than capacity. That is, these two groups might be able to remember an equal number of items (i.e., comparable capacity) but PSZ generally process and maintain the items in a less

precise manner. Yet, few studies have attempted to directly estimate memory capacity and precision at the same time in schizophrenic patients (25, 26). Using the same delayed-estimation task and the same computational model, Gold et al (25) found decreased capacity and intact precision in PSZ but Xie&Zhang (26) obtained the opposite conclusion. Those studies, however, employed a phenomenological model that categorizes response errors into a mixture of two distributions, and did not consider the trial-by-trial sensory uncertainty and encoding-decoding mechanisms underlying VWM, thereby offering limited insight into the mechanistic differences between PSZ and HCS.

Disentangling the effects of different VWM components (e.g., capacity, precision) on performance requires generative computational models that explicitly describe how these components give rise to a behavioral response. van den Berg *et al.* (37) proposed a variable precision (VP) model (Fig. 1), suggesting that memory resources assigned to individual items are not only continuous but also variable across items and trials. Variable resources cause variable processing precision and thus affect the trial-by-trial response error. Furthermore, the VP model explicitly distinguishes the precision of encoding items from the precision of exerting a behavior choice (e.g., motor or decision noise), an aspect that most previous models ignored. Tested on a large-scale benchmark database of VWM, the VP model has been shown to outperform other conventional models (38).

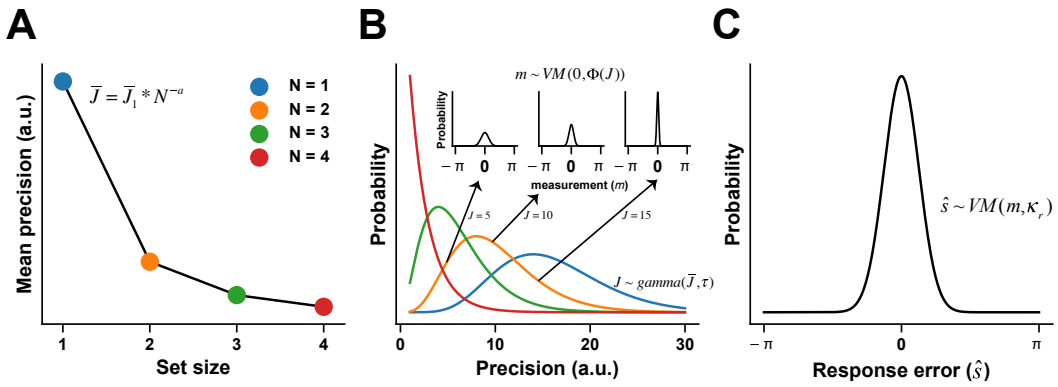
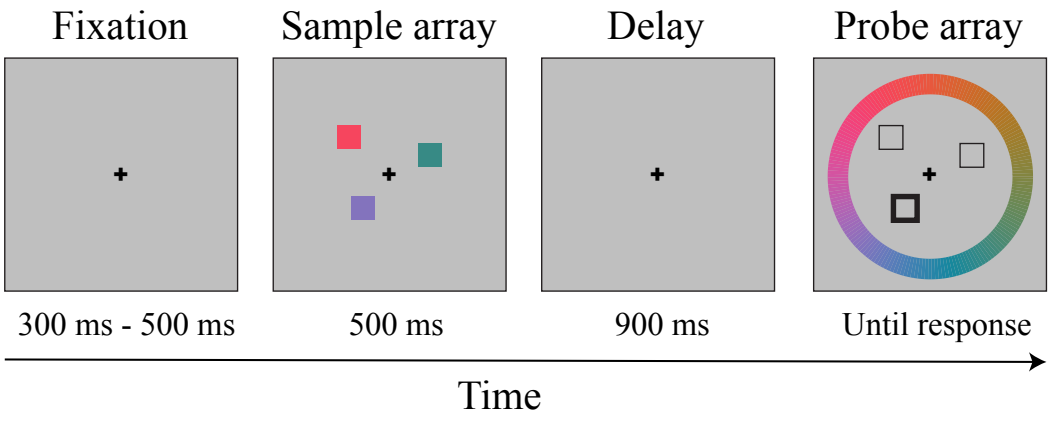


Figure 1. Variable precision model of VWM. ***A***. Resource decay function. The VP model assumes that the mean resource ($\bar{J}$) for processing a single item declines as a power function of set size N, a trend characterized by two free parameters–initial resources ($\bar{J}_1$) and decaying exponent (a). ***B***. The resources across items or trials

follow a gamma distribution with the mean resource ($\bar{J}$) determined by the resource decay function (panel A) and the resource allocation variability (τ). Larger amounts of resources (J) indicate higher precision and therefore generate narrower von Mises distributions (three small axes indicating the precision equals to 5, 10 and 15 respectively) of stimulus measurement (m). The widths of the von Mises distributions indicate the degree of trial-by-trial sensory uncertainty. ***C***. The eventual behavioral choice given the internal stimulus measurement (m) is also uncertain, following a von Mises distribution with the choice variability (κ_r) (39). In the VP model, initial resources ($\bar{J}_1$), decaying exponent (a), resource allocation variability (τ) and choice variability (κ_r) are four free parameters to estimate (see details in SI and van den Berg *et al*. (37)). All numbers here are only for illustration purposes and not quantitatively related to the model fitting results in this paper.

Leveraging recent advances in basic science, we aimed to use the VP model to understand the computational underpinnings of VWM deficits in schizophrenia. We sought to systematically disentangle the impact of memory capacity and resources, as well as other factors (i.e., variability in allocating resources and variability in choice) in PSZ. In this study, the performance of PSZ and demographically matched HCS was measured in a standard VWM delayed-estimation task (Fig. 2). Using a standard task allows us to compare our results to previous ones (25, 40–43). We believe that a well-controlled task and thorough computational modeling will shed a new light on the mechanisms of perceptual deficits associated with schizophrenia.



159 Figure 2. The color delay-estimation task. This figure depicts an
160 example trial (set size = 3) of the color delay-estimation task.
161 Subjects are instructed to first memorize the colors of all squares on
162 the screen and after a 900 ms-delay choose the color of the probed
163 square (the one in the lower visual field in this example) on a color
164 wheel. Set size (i.e., number of squares in the sample array) varies
165 trial by trial. Response error is the difference between the reported
166 color and the real color of the probe in the standard color space.
167
168

RESULTS

We measured the performance of 60 PSZ and 61 HCS in a standard color delay-estimation task (Fig. 2). In the task, subjects were instructed to first memorize a set of colored squares. After a short delay, they chose the color of the probed square on a color wheel that represents the standard color space. The difference between the reported color and the true color of the target is considered as the response error.

Variable precision model accounts for VWM behavior in both HCS and PSZ

The VP model proposes that memory resources are continuous and the amount of resource assigned to individual items varies across items and trials. Previous studies have suggested the VP model, compared with other conventional models, as the best account of VWM so far in normal adults (37, 38). However, PSZ might use a qualitatively different observer model other than the VP model to perform the task. We therefore first compared the VP model against the other five computational models that summarize major existing theories of VWM (see model details in Supplementary Appendix note 1). The first one is the item-limit (IL) model. The IL model assumes no uncertainty in the sensory encoding stage and that each subject has a fixed memory capacity and a fixed response variability across set size levels (44). The second one is the mixture (MIX) model, similar to the IL model but assuming response variability is set-size dependent. Note that the MIX model has been used to study schizophrenia (25, 26), thus it enables us to directly compare our results with previous studies. Also, the MIX model is originally called “slots-plus-averaging model” in (45) but we denote it as the MIX model here to differentiate it from our third model described as follows. Compared with the MIX model, our slots-plus-averaging (SA) model (45) further elaborates the idea that memory resources manifest as discrete chunks and these chunks can be flexibly assigned to multiple items. The fourth one is the equal-precision (EP) model, which is similar to the VP model but assumes that the memory resources are equally distributed across items (46, 47). Note that the VP model does not include the capacity component thus one might argue that we cannot draw any direct conclusion about memory capacity by merely analyzing the VP model. To circumvent this, we also

constructed a variable-precision-with-capacity (VPcap) model that not only acknowledges the variable precision mechanisms and but also explicitly estimates the capacity of individual subjects.

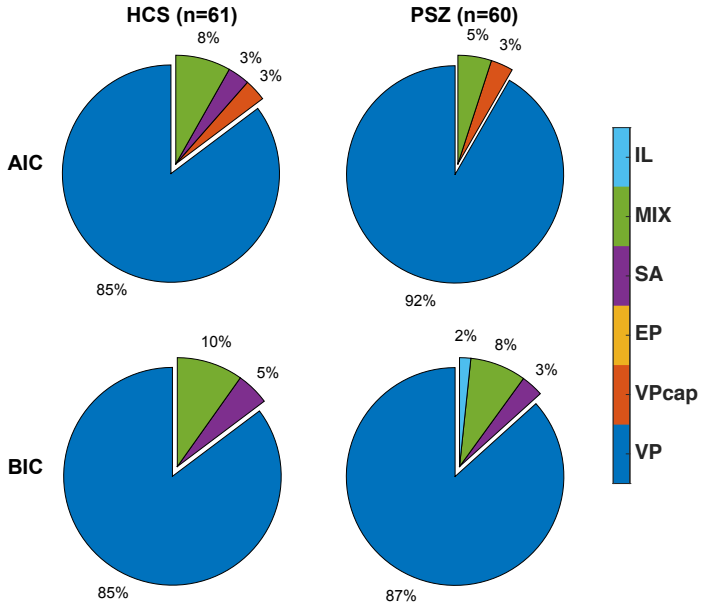


Figure 3. Model comparison results. The VP model is tested against other five computational models. The pie charts illustrate the proportion of subjects for whom each model is their best-fitting model. The VP model is the best-fitting model for over 85% of subjects in both groups and under both AIC and BIC, indicating both groups share a qualitatively similar internal process of VWM.

We compared all six models using the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) (48, 49). It turned out that, among all models, the VP model was the best-fitting model for over 85% of subjects in the HCS group under both metrics, replicating previous results in normal subjects (37, 38). Most importantly, the VP model was also the best-fitting model for over 87% of subjects in the PSZ group, indicating that both groups used a qualitatively same observer model to perform the task.

It is worth highlighting two comparisons here. First, the VP model outperformed the MIX model in both groups. This result directly questions the reliability of the previous study based on the MIX model (25). Second, we found that the VP model was better than the VPcap model, even though the later one

incorporates the additional capacity parameter. This result suggests that adding the capacity parameter seems unnecessary from the modeling perspective. This result is also in line with the literature showing that a fixed capacity might not exist in VWM (50, 51). Although systematically examining the existence of a fixed capacity is beyond the scope of this paper, this result invites a rethink of whether memory capacity should be considered as a key factor that limits VWM behavior in PSZ.

Larger resource allocation variability in PSZ

Analyses above have established that HCS and PSZ employ the qualitatively same observer model to complete the VWM task. Their behavioral differences thus can only be attributed to the parameter differences on some components of VWM. We next compared the two groups' fitted parameters of the VP model. Results showed that the two groups had comparable resource decay functions (Fig. 4A, initial resources, $t(119) = 0.689$, $p = 0.492$, $d = 0.125$; decaying exponent, $t(119) = 1.065$, $p = 0.289$, $d = 0.194$), indicating a similar trend of diminished memory resources as set size increases. PSZ, however, had larger variability in allocating resources (Fig. 4B, resource allocation variability, $t(119) = 4.03$, $p = 9.87 \times 10^{-5}$, $d = 0.733$). This suggests that, although the two groups have on average the same amount of memory resources across different set size levels, PSZ allocate the resources across items or trials in a more heterogeneous manner, with some items in some trials receiving considerably larger amounts and vice versa in other cases. This is theoretically suboptimal with respect to completing the task since the probe was randomly chosen among all presented items with an equal probability. The optimal strategy therefore should be to assign an equal amount of resources to every item and in every trial to tackle the unpredictable target. Furthermore, our VP model explicitly distinguishes the variability of processing items and the variability in exerting a behavior choice (e.g., motor or decision noise). We found no significant group difference in the choice variability (Fig. 4C, $t(119) = 1.7034$, $p = 0.091$, $d = 0.31$), diminishing the possibility that the atypical performance of PSZ arises from larger variability at the choice stage.

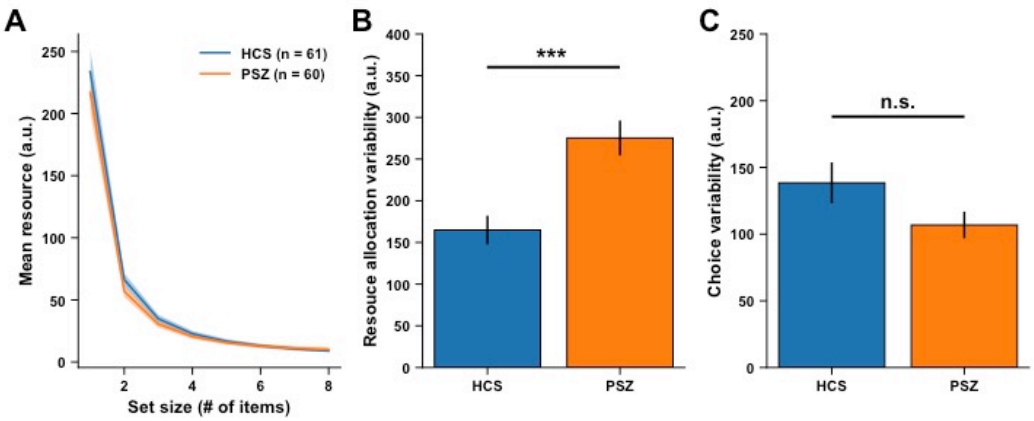


Figure 4. Fitted parameters of the VP model. No significant group differences are noted between two groups in resource decay functions (panel A), and choice variability (panel C). PSZ have larger resource allocation variability than HCS (panel B). The individual resource decay functions are computed by $\bar{J}_1 = \bar{J} * N^{-a}$, where N is the set size, $\bar{J}_1$ and a are the estimated initial resources and the decaying exponent of one subject. The solid lines represent the averaged resource decay functions across subjects. The shaded areas in panel A, and all error bars in panel B and C represent ±SEM across subjects. Significance symbol conventions are ***: $p < 0.001$; n.s.: non-significant.

Suboptimal models reveal no capacity difference between HCS and PSZ

Although the VP model was the most appropriate model for both groups, we further examined other suboptimal models. We believe this is valuable for several reasons. First, the VP model does not have the concept of capacity thus we cannot completely rule out the influence of capacity. One might argue that resource allocation variability and limited capacity might jointly manifest in PSZ thus a hybrid model that aggregates two factors might yield a better explanation. Second, the fitted parameters of all models might be derived from specific model settings or possible idiosyncratic model fitting processes.

We argue that the VPcap model is such a hybrid model that accommodates both the variable precision mechanism and a fixed capacity. However, this model is worse than the VP model. We also evaluated the fitted parameters in the VPcap model since it inherits the variable precision mechanism. The results from the VPcap model largely replicated the results of the VP model. Again, we found a

significantly larger resource allocation variability in PSZ ($t(119) = 3.891$, $p = 1.65 \times 10^{-4}$, $d = 0.707$, see full statistical results in Supplementary Appendix note 2), This result suggests that the effect of resource allocation variability is quite robust even though we altered the model structure.

We further examined the estimated capacity of all subjects in the IL, the MIX, the SA and the VPcap model. Most importantly, none of the four models suggest decreased capacity in PSZ (see full stats in Supplementary Appendix note 3). Particularly, no significant group difference in capacity was noted in the MIX model, which contradicts the results of the previous study that used the same task and the same MIX model (25).

In sum, we found robustly larger resource allocation variability in PSZ in both the VP and the VPcap models, strongly supporting its key role in limiting performance in PSZ. Also, we found no evidence for decreased capacity in PSZ in all models that include the capacity parameter. These results directly challenge the widely accepted decreased-capacity account, and highlight the role of resource allocation variability in VWM deficits of PSZ.

Resource allocation variability predicts the severity of schizophrenic symptoms

We next turned to investigate whether the results from the VP model are predictive of clinically measured symptoms. A set of correlational analyses was carried out to link the estimated resource allocation variability from the VP model to the schizophrenia symptomatology (BPRS, SANS, and SAPS).

We noticed that the estimated resource allocation variability of individual subjects correlates with their BPRS scores (Fig. 5A, $r = 0.259$, $p = 0.045$) and the SANS scores (Fig. 5B, $r = 0.302$, $p = 0.019$) in PSZ. No significant correlation was noted on the SAPS scores (Fig. 5C, $r = -0.121$, $p = 0.358$). These results suggest that resource allocation variability not only is the key factor from a computational perspective but also can quantitatively predict the severity of clinically measured symptoms.

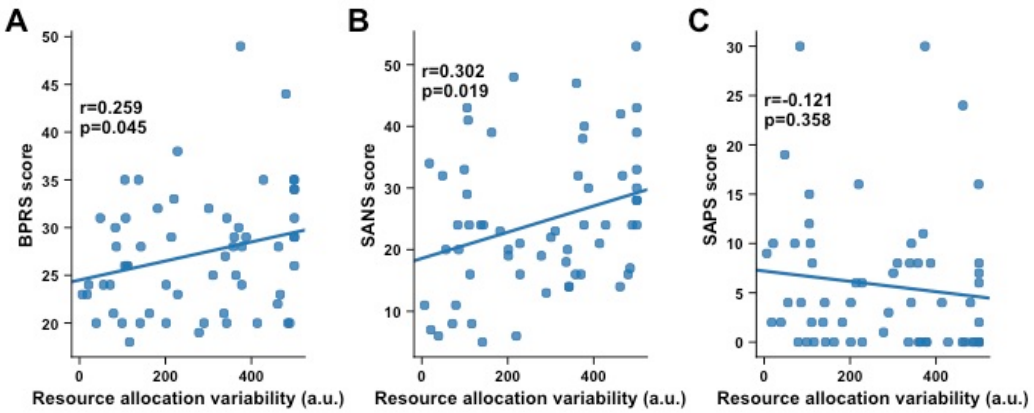


Figure 5. Individual differences in resource allocation variability predict the scores in symptom assessments. Estimated resource allocation variability values in the PSZ group significantly correlates with their scores on BPRS (panel A) and SANS (negative symptoms, panel B) but not on SAPS (positive symptoms, panel C).

DISCUSSION

The mechanisms of VWM deficits in schizophrenia have been a matter of debate in the past few years. One widely accepted view proposes decreased capacity as the major cause of the deficits in PSZ. This conclusion, however, is merely based on estimated behavioral performance or suboptimal VWM models. In the present study, we re-examined this conclusion by comparing the performance of PSZ and HCS using the computational models that summarize several major theories of VWM. We first established that the VP model was the best model to characterize performance in both groups, indicating the presence of a qualitative similar internal process in both groups. We then further evaluated different components in the VP model as well as other suboptimal models, with special focuses on memory capacity and the declining trend of mean precision as a function of set size. Surprisingly, we found that PSZ and HCS differ in none of these two diagnostic features of VWM. Most importantly, we showed that PSZ had larger variability in allocating memory resources. Furthermore, individual differences in resource allocation variability predict variation of patients' symptom severity, highlighting the clinical functionality of this factor. Taken together, our results challenge the long-standing decreased-capacity explanation for the VWM deficits in schizophrenia and propose for the first time that resource allocation variability might be the key determinant that limits their performance.

A large body of literature has documented that PSZ perform poorly in various forms of working memory tasks (2, 3, 52, 53). Most studies reached the same conclusion: memory capacity is decreased in schizophrenia. However, a careful examination of the literature reveals that the definition of capacity varies across studies. Many studies directly equated worse performance with decreased capacity without quantitatively demonstrating how capacity modulates performance. For example, memory capacity is usually defined as the number of digits that can be recalled in the longest strand in digit span tasks (12). Also, in N-back tasks capacity is defined as the number of backs corresponding to a certain accuracy level (14–16). Moreover, the calculation of capacity resembles the d-prime metric in change detection tasks (22–24, 44, 54). The majority of these metrics are more like

behavioral thresholds related to capacity rather than the quantitative estimates of capacity. Although these metrics indeed suggest worse performance in PSZ, we cannot easily equate the worse performance with decreased capacity given the presence of other components such as memory resource or choice variability. This is partly because we lack appropriate generative models for the majority of the tasks thus it is unclear how these components jointly determine performance. The VP model can describe the generative process of the delay-estimation task and the change-detection task (37), and allow us to disassociate the effect of capacity from other VWM components.

Only a few studies have quantitatively estimated capacity and precision in schizophrenia but they demonstrated inconsistent findings. Gold et al (25) employed the same delay-estimation task and estimated individual's capacity and precision using the MIX model (see introduction above and details in Supplementary Appendix note 1). The model assumes that response errors arise from a mixture distribution combining a von Mises distribution whose variance reflects memory precision, and a uniform distribution accounting for the random guessing if the set size exceeds capacity. Results obtained from the MIX model echoed the decreased-capacity theory. The MIX model, however, does not take into account two important factors. First, the model assumes an equal precision across items held in memory. Second, the model does not separate the variability for processing stimuli (i.e., sensory uncertainty) and the variability in exertion of a choice (i.e., choice uncertainty), an issue most previous VWM models also ignored. Such distinction is important since it highlights different types of uncertainty in encoding and decoding stages of VWM. Mathematically, these two types of uncertainty can be distinguished by manipulating set size since the encoding variability depends on set size but the choice variability does not. The issues of the MIX model have been symmetrically addressed in a recent work (55). Most importantly, we showed in this study that the MIX model is not only conceptually inappropriate but also quantitatively less accurate than the VP model. In addition, we did not find any group difference in capacity even examining the fitted parameters.

Compared with capacity and precision–two diagnostic features of VWM, resource allocation variability is a new concept. It describes the heterogeneity of allocating resources across multiple items and trials. It has been recently shown that such variability might not only manifest in VWM and but also act as a ubiquitous mechanism when processing multiple objects (56). We speculate that resource allocation variability reflects the extent of attentional control when the brain processes multiple objects. Two aspects of available evidence support this argument. First, it has been shown that attention and WM are two core components of executive control and tightly linked with each other (57, 58). Second, schizophrenia is known to have deficits in top-down attentional modulation (53, 57). Particularly, recent studies discovered the phenomenon of spatial hyperfocusing in schizophrenia patients (19, 59–61). If schizophrenia patients overly attend one item and ignore others in the memory encoding stage, unbalanced resource allocation will likely occur.

What are the neural mechanisms of this resource allocation variability? Recent neurophysiological studies proposed that the neural representation of a stimulus may follow a doubly stochastic process (62, 63), which suggests that the variability in encoding precision is a consequence of trial-to-trial and item-to-item fluctuations in attentional gain (32, 37, 64). A recent study combined functional magnetic resonance imaging and the VP model, showing that the superior intraparietal sulcus (IPS) is the cortical locus that controls the resource allocation (65). Interestingly, schizophrenic patients have been known to have IPS deficits (66). Note that besides top-down factors, we cannot rule out the contribution of bottom-up neural noise in perceptual and cognitive processing (62, 63), as found in several other mental diseases (33–36).

The current results also reveal links between resource allocation variability and patients' negative symptoms, but not positive symptoms (Fig. 5). These findings are consistent with several experimental and meta-analysis studies claiming dissociable mechanisms underlying the cluster of negative symptoms versus that of positive symptoms (67–70). More broadly, a growing collection of evidence suggests that visual perceptual deficits in schizophrenic patients are more

likely to link to negative rather than positive symptom severity (71–75). Negative symptoms in turn might produce improvised social functioning. Humans depend heavily on VWM to interact with multiple agents and complete social tasks. Deficits in distributing processing resources over multiple agents therefore might cause disadvantages in social cognition.

In conclusion, our study proposes a new explanation that the resource allocation variability accounts for the atypical VWM performance in schizophrenia. This view differs from the decreased-capacity theory and provides a new direction for future studies that attempt to promote diagnosis and rehabilitation for schizophrenic patients.

MATERIALS AND METHODS

Ethics Statement. All experimental protocols were approved by the institutional review board at the East China Normal University. All research was performed in accordance with relevant guidelines and regulations. Informed written consent was obtained from all participants.

Subjects. 61 HCS and 60 PSZ participated in the study. PSZ were clinically stable inpatients (N = 33) and outpatients (N = 27) who met DSM-IV criteria (76) for schizophrenia. All patients were receiving antipsychotic medication (2 first-generation, 43 second-generation, 15 both). Symptom severity was evaluated by the Brief Psychiatric Rating Scale (BPRS) (77), the Scale for the Assessment of Negative (SANS) and Positive Symptoms (SAPS) (78, 79). HCS were recruited by advertisement. All HCS had no current diagnosis of axis 1 or 2 disorders as well as no family history of psychosis nor substance abuse or dependence. All subjects are right-handed with normal sight and color perception.

The two groups were matched in age ($t(119) = 1.58$, $p = 0.118$, $d = 0.284$), gender ($t(119) = 1.20$, $p = 0.234$, $d = 0.218$) and education level of parents ($t(119) = 0.257$, $p = 0.798$, $d = 0.047$). Inevitably, the PSZ had fewer years of education than the HCS ($t(119) = 5.51$, $p = 2.09 \times 10^{-7}$, $d = 1.00$). The detailed demographic information is summarized in the Table 1.

463 **Table 1. Demographics and clinical information of people with schizophrenia**
464 **(PSZ) and healthy control subjects (HCS)**

	PSZ (N = 60)		HCS (N = 61)	
	Mean	SD	Mean	SD
age	35.67	6.58	33.82	9.90
range	23-48	n/a	21-57	n/a
Female/male	31/29	n/a	29/32	n/a
Inpatient/outpatient	33/27	n/a	n/a	n/a
Subject's education (in years)	12.03	2.24	15.13	3.70
Paternal education (in years) [a]	9.89	2.53	9.76	2.95
Maternal education (in years)	9.62	2.91	9.29	3.63
BPRS	27.25	6.27	n/a	n/a
SAPS	5.77	7.02	n/a	n/a
SANS	24.43	11.45	n/a	n/a

465 [a] Average of mother's and father's years of education

466 BPRS: Brief Psychiatric Rating Scale (77); SAPS: Scale for the Assessment of
467 Positive Symptoms (79); SANA: Scale for the Assessment of Negative Symptoms
468 (78).

469

470 **Stimuli and Task.** Subjects sat 50 cm away from an LCD monitor. All stimuli were
471 generated by Matlab 8.1 and Psychtoolbox 3 (80, 81), and then presented on the
472 LCD monitor.

473 In the color delay-estimation VWM task, each trial began with a fixation cross
474 presented at center-of-gaze for a duration randomly chosen from a sequence of 300,
475 350, 400, 450 and 500 ms. Subjects shall keep their fixation on the cross throughout
476 the whole experiment. A set of colored squares (set size = 1 or 3) was shown within
477 an invisible area with 4^{o} radius. The sample array lasted 500 ms. Each square was
478 $1.5^{o} \times 1.5^{o}$. Their colors were randomly selected from the 180 colors that are
479 equally distributed along the wheel representing the CIE L*a*b color space. The
480 color wheel was centered at (L = 70, a = 20, b = 38) with a radius of 60 in the color
481 space (40). The sample array then disappeared and was followed by a 900 ms blank

period for memory retention. After the delay, an equal number of outlined squares were shown at the same location of each sample array item, with one of them bolded as the probe. At the meantime, a randomly rotated color wheel was shown. The color wheel was 2.1^{o} thick and centered on the monitor with the inner and the outer radius as 7.8^{o} and 9.8^{o} respectively. Subjects were asked to choose the remembered color of the probe by clicking a color on the color wheel using a computer mouse. Subjects shall choose the color as precisely as possible and response time was not constrained. Every subject completed 2 blocks for the set size 1 and 3, respectively. The order of the two blocks was counterbalanced across subjects. Each block had 80 trials.

494 **REFERENCES**

495

496 1. Gold JM, Randolph C, Carpenter C (1997) Auditory working memory and
497 Wisconsin Card Sorting Test Performance in Schizophrenia. *Arch Gen*
498 *Psychiatry* 54:159–165.
499 2. Lee J, Park S (2005) Working Memory Impairments in Schizophrenia: A
500 Meta-Analysis. *J Abnorm Psychol* 114:599–611.
501 3. Forbes NF, Carrick LA, McIntosh AM, Lawrie SM (2009) Working memory
502 in schizophrenia: a meta-analysis. *Psychol Med* 39:889–905.
503 4. Goldman-Rakic PS (1994) Working memory dysfunction in schizophrenia. *J*
504 *Neuropsychiatry Clin Neurosci* 6:348–357.
505 5. Mayer RE, Moreno R (1998) A split-attention effect in multimedia learning:
506 Evidence for dual processing systems in working memory. *J Educ Psychol*
507 90:312–320.
508 6. Postle BR (2006) Working memory as an emergent property of the mind and
509 brain. *Neuroscience* 139:23–38.
510 7. Nee DE, et al. (2013) A meta-Analysis of executive components of working
511 memory. *Cereb Cortex* 23:264–282.
512 8. Luck SJ, Vogel EK (2013) Visual working memory capacity: From
513 psychophysics and neurobiology to individual differences. *Trends Cogn Sci*
514 17:391–400.
515 9. Coleman MJ, Krastoshevsky O, Tu X, Mendell NR, Levy DL (2012) The
516 effects of perceptual encoding on the magnitude of object working memory
517 impairment in schizophrenia. *Schizophr Res* 139:60–65.
518 10. Leonard CJ, et al. (2013) Testing sensory and cognitive explanations of the
519 antisaccade deficit in schizophrenia. *J Abnorm Psychol* 122:1111–1120.
520 11. Johnson MK, et al. (2013) The relationship between working memory
521 capacity and broad measures of cognitive ability in healthy adults and people
522 with schizophrenia. *Neuropsychology* 27:220–229.
523 12. Conklin HM (2000) Verbal Working Memory Impairment in Schizophrenia
524 Patients and Their First-Degree Relatives: Evidence From the Digit Span
525 Task. *Am J Psychiatry* 157:275–277.
526 13. Chey J, Lee J, Kim Y-S, Kwon S-M, Shin Y-M (2002) Spatial working
527 memory span, delayed response and executive function in schizophrenia.
528 *Psychiatry Res* 110:259–271.
529 14. Callicott J (1998) Functional Magnetic Resonance Imaging Brain Mapping in
530 Psychiatry: Methodological Issues Illustrated in a Study of Working Memory
531 in Schizophrenia. *Neuropsychopharmacology* 18:186–196.
532 15. Barch DM, Csernansky JG, Conturo T, Snyder AZ (2002) Working and
533 long-term memory deficits in schizophrenia: Is there a common prefrontal
534 mechanism? *J Abnorm Psychol* 111:478–494.
535 16. Jansma JM, Ramsey NF, Van Der Wee NJA, Kahn RS (2004) Working
536 memory capacity in schizophrenia: A parametric fMRI study. *Schizophr Res*
537 68:159–171.
538 17. Park S, Holzman PS (1992) Schizophrenics Show Spatial Working Memory
539 Deficits. *Arch Gen Psychiatry* 49:975–982.

18. Keedy SK, Ebens CL, Keshavan MS, Sweeney JA (2006) Functional magnetic resonance imaging studies of eye movements in first episode schizophrenia: Smooth pursuit, visually guided saccades and the oculomotor delayed response task. *Psychiatry Res* 146:199–211.
19. Sawaki R, et al. (2017) Hyperfocusing of attention on goal-related information in schizophrenia: Evidence from electrophysiology. *J Abnorm Psychol* 126:106–116.
20. Gold JM, Wilk CM, McMahon RP, Buchanan RW, Luck SJ (2003) Working memory for visual features and conjunctions in schizophrenia. *J Abnorm Psychol* 112:61–71.
21. Lencz T, et al. (2003) Impairments in perceptual competency and maintenance on a visual delayed match-to-sample test in first-episode schizophrenia. *Arch Gen Psychiatry* 60:238–243.
22. Erickson MA, et al. (2015) Impaired Working Memory Capacity Is Not Caused by Failures of Selective Attention in Schizophrenia. *Schizophr Bull* 41:366–373.
23. Erickson MA, et al. (2014) Enhanced vulnerability to distraction does not account for working memory capacity reduction in people with schizophrenia. *Schizophr Res Cogn* 1:149–154.
24. Leonard CJ, et al. (2013) Toward the Neural Mechanisms of Reduced Working Memory Capacity in Schizophrenia. *Cereb Cortex* 23:1582–1592.
25. Gold JM, et al. (2010) Reduced Capacity but Spared Precision and Maintenance of Working Memory Representations in Schizophrenia. *Arch Gen Psychiatry* 67:570–577.
26. Xie W, et al. (2018) Schizotypy is associated with reduced mnemonic precision in visual working memory. *Schizophr Res* 193:91–97.
27. Starc M, et al. (2017) Schizophrenia is associated with a pattern of spatial working memory deficits consistent with cortical disinhibition. *Schizophr Res* 181:107–116.
28. Ma WJ, Husain M, Bays PM (2014) Changing concepts of working memory. *Nat Neurosci* 17:347–356.
29. Bialek W (1987) Physical Limits to Sensation and Perception. *Annu Rev Biophys Biophys Chem* 16:455–478.
30. Rolls ET, Deco G (2010) *The Noisy BrainStochastic Dynamics as a Principle of Brain Function* (Oxford University Press).
31. Faisal AA, Selen LPJ, Wolpert DM (2008) Noise in the nervous system. *Nat Rev Neurosci* 9:292–303.
32. Goris RLT, Movshon JA, Simoncelli EP (2014) Partitioning neuronal variability. *Nat Neurosci* 17:858–865.
33. Dinstein I, et al. (2012) Unreliable Evoked Responses in Autism. *Neuron* 75:981–991.
34. Park WJ, Schauder KB, Zhang R, Bennetto L, Tadin D (2017) High internal noise and poor external noise filtering characterize perception in autism spectrum disorder. *Sci Rep* 7:17584.
35. Sperling AJ, Lu Z-L, Manis FR, Seidenberg MS (2005) Deficits in perceptual noise exclusion in developmental dyslexia. *Nat Neurosci* 8:862–

863.

36. Bubl E, et al. (2015) Elevated Background Noise in Adult Attention Deficit Hyperactivity Disorder Is Associated with Inattention. *PLoS One* 10:e0118271.
37. van den Berg R, Shin H, Chou W-C, George R, Ma WJ (2012) Variability in encoding precision accounts for visual short-term memory limitations. *Proc Natl Acad Sci* 109:8780–8785.
38. van den Berg R, Awh E, Ma WJ (2014) Factorial comparison of working memory models. *Psychol Rev* 121:124–149.
39. Osborne LC, Lisberger SG, Bialek W (2005) A sensory source for motor variation. *Nature* 437:412–416.
40. Zhang WW, Luck SJ (2008) Discrete fixed-resolution representations in visual working memory. *Nature* 453:233–235.
41. Zhang WW, Luck SJ (2009) Sudden Death and Gradual Decay in Visual Working Memory: Research Report. *Psychol Sci* 20:423–428.
42. Zhang WW, Luck SJ (2011) The number and quality of representations in working memory. *Psychol Sci* 22:1434–1441.
43. Foster JJ, Bsales EM, Jaffe RJ, Awh E (2017) Alpha-Band Activity Reveals Spontaneous Representations of Spatial Position in Visual Working Memory. *Curr Biol* 27:3216–3223.
44. Pashler H (1988) Familiarity and visual change detection. *Percept Psychophys* 44:369–378.
45. Zhang WW, Luck SJ (2008) Discrete Fixed-Resolution Representations in Visual Working Memory Weiwei. *Nature* 453:233–235.
46. Palmer J (1990) Attentional limits on the perception and memory of visual information. *J Exp Psychol Hum Percept Perform* 16:332–350.
47. Shaw M (1980) Identifying attentional and decision-making components in information processing. *Attention and Performance*, ed Nickerson R (Erlbaum, Hillsdale, NJ), pp 227–296.
48. Wit E, van den Heuvel E, Romeijn JW (2012) 'All models are wrong. ': An introduction to model uncertainty. *Stat Neerl* 66:217–236.
49. Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer-Verlag, New York, NY). 2nd Ed.
50. Bays PM, Husain M (2008) Dynamic Shifts of Limited Working Memory Resources in Human Vision. *Science* 321:851–854.
51. Bays PM, Catalao RFG, Husain M (2009) The precision of visual working memory is set by allocation of a shared resource. *J Vis* 9:7.1-11.
52. Shakow D (1972) The Worcester State Hospital research on schizophrenia (1927-1946). *J Abnorm Psychol* 80:67–110.
53. Gold JM, Hahn B, Strauss GP, Waltz JA (2009) Turning it upside down: Areas of preserved cognitive function in schizophrenia. *Neuropsychol Rev* 19:294–311.
54. Cowan N (2001) The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav Brain Sci* 24:87–114.
55. Ma WJ (2018) Problematic usage of the Zhang and Luck mixture model.

bioRxiv. doi:10.1101/268961.

56. Shen S, Ma WJ (2018) Variable precision in visual perception. *bioRxiv*. doi:10.1101/153650.
57. Gold JM, et al. (2017) Selective Attention, Working Memory, and Executive Function as Potential Independent Sources of Cognitive Dysfunction in Schizophrenia. *Schizophr Bull* 12:1–8.
58. Christophel TB, Iamshchinina P, Yan C, Allefeld C, Haynes JD (2018) Cortical specialization for attended versus unattended working memory. *Nat Neurosci* 21:494–496.
59. Fuller RL, et al. (2006) Impaired control of visual attention in schizophrenia. *J Abnorm Psychol* 115:266–275.
60. Kreither J, et al. (2017) Electrophysiological Evidence for Hyperfocusing of Spatial Attention in Schizophrenia. *J Neurosci* 37:3813–3823.
61. Luck SJ, et al. (2014) Hyperfocusing in schizophrenia: Evidence from interactions between working memory and eye movements. *J Abnorm Psychol* 123:783–795.
62. Churchland AK, et al. (2011) Variance as a Signature of Neural Computations during Decision Making. *Neuron* 69:818–831.
63. Churchland MM, et al. (2010) Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci* 13:369–378.
64. Nienborg H, Cumming BG (2009) Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature* 459:89–92.
65. Galeano Weber EM, Peters B, Hahn T, Bledowski C, Fiebach CJ (2016) Superior Intraparietal Sulcus Controls the Variability of Visual Working Memory Precision. *J Neurosci* 36:5623–5635.
66. Zhou S-Y, et al. (2007) Parietal lobe volume deficits in schizophrenia spectrum disorders. *Schizophr Res* 89:35–48.
67. de Gracia Dominguez M, Viechtbauer W, Simons CJP, van Os J, Krabbendam L (2009) Are Psychotic Psychopathology and Neurocognition Orthogonal? A Systematic Review of Their Associations. *Psychol Bull* 135:157–171.
68. Cameron AM, et al. (2002) Working memory correlates of three symptom clusters in schizophrenia. *Psychiatry Res* 110:49–61.
69. Carter C, et al. (1996) Spatial working memory deficits and their relationship to negative symptoms in unmedicated schizophrenia patients. *Biol Psychiatry* 40:930–932.
70. Park S, Sauter BH, Rentsch M, Hell D (2002) Visual object working memory function and clinical symptoms in schizophrenia. *Schizophr Res* 59:261–268.
71. Cadenhead KS, et al. (1997) Information processing deficits of schizophrenia patients: Relationship to clinical ratings, gender and medication status. *Schizophr Res* 28:51–62.
72. Butler PD, Javitt DC (2005) Early-stage visual processing deficits in schizophrenia. *Curr Opin Psychiatry* 18:151–157.
73. Kéri S, Kiss I, Kelemen O, Benedek G, Janka Z (2005) Anomalous visual experiences, negative symptoms, perceptual organization and the magnocellular pathway in schizophrenia: A shared construct? *Psychol Med*

35:1445–1455.
74. Slaghuis WL (2004) Spatio-temporal luminance contrast sensitivity and visual backward masking in schizophrenia. *Exp brain Res* 156:196–211.
75. Slaghuis WL, Bishop AM (2001) Luminance flicker sensitivity in positive- and negative-symptom schizophrenia. *Exp Brain Res* 138:88–99.
76. American Psychiatric Association (1994) *Diagnostic and statistical manual of mental disorders (4th ed.)* (Washington, DC).
77. Overall JE, Gorham DR (1962) The Brief Psychiatric Rating Scale. *Psychol Rep* 10:799–812.
78. Andreasen NC (1983) *The Scale for the Assessment of Negative Symptoms (SANS)* (Iowa City, IA).
79. Andreasen NC (1984) *The Scale for the Assessment of Positive Symptoms (SAPS)* (Iowa City, IA).
80. Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436.
81. Pelli DG (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spat Vis* 10:437–442.

695

Supplementary Information for

Atypically larger variability of resource allocation accounts for visual working memory deficits in schizophrenia

Yijie Zhao, Xuemei Ran, Li Zhang, Ruyuan Zhang, Yixuan Ku

Ruyuan Zhang
Email: ruyuanzhang@gmail.com

Yixuan Ku
Email: yk1616@nyu.edu

This PDF file includes:

Supplementary Note 1: Computational modeling of VWM

VP model. The variable precision (VP) model has been shown as the state-of-the-art computational model of VWM. Details of the VP model have been documented in several previous studies (1, 2) and the model codes are publicly available (http://www.cns.nyu.edu/malab/resources.html).

The VP model assumes a resource decaying function describing the decreasing trend of mean memory resource ($\overline{J}$) assigned to individual items as the set size (N) increases (3, 4):

$$\overline{J} = \overline{J_1} * N^{-a} \ , \tag{S1}$$

where $\overline{J_1}$ is the initial resources when only 1 item ($N = 1$) should be memorized and a is the decaying exponent. The key component of the VP model is that the memory resources J across items and trials follow a Gamma distribution with the mean $\overline{J}$ and the scale parameter τ:

$$J \sim Gamma(\overline{J}, \tau) \ , \tag{S2}$$

Intuitively, a larger τ indicates a more uneven distribution of memory resources across items or trials, with some items in some trials receiving a larger amount of resource while others receive comparative fewer. Note that a larger amount of memory resource produces a higher precision. Thus, we do not explicitly distinguish resource and precision and denote them as J. Defining precision as Fisher information (5), precision J can be linked to the variance of the von Mises distribution of sensory measurement:

$$J = k \frac{I_1(\kappa)}{I_0(\kappa)}, \tag{S3}$$

where I_0 and I_1 are modified Bessel functions of the first kind of order 0 and 1 respectively, with the concentration parameter κ. Eq. S3 specifies a one-on-one mapping between precision J and variance κ. We can rewrite their relationship as:

$$\kappa = \Phi(J) \ , \tag{S4}$$

where Φ is the mapping function. The distribution of sensory measurement (m) given the input stimulus (s) can be written as:

$$p(m \mid s) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(m-s)} \equiv VM(m; s, \kappa), \tag{S5}$$

We further assume that the reported color ($\hat{s}$) by participants also follows a von Mises distribution:

$$p(\hat{s} \mid m) = \frac{1}{2\pi I_0(\kappa_r)} e^{\kappa_r \cos(\hat{s}-m)} \equiv VM(\hat{s}; m, \kappa_r), \tag{S6}$$

where k_r represents the variability at the choice stage.

Taken together, this VP model has four free parameters: $\bar{J}_1$, a, τ and k_r.

Given the four free parameters and stimulus color s in a trial, we can derive the probability of the observed response in a trial by marginalizing over sensory measurement m and variable precision J:

$$\begin{aligned} p(\hat{s} \mid s; \bar{J}, \tau) &= \int p(\hat{s} \mid s; J)\, p(J \mid \bar{J}; \tau)\, dJ \\ &= \int VM(\hat{s}; s, \Phi(J)) Gamma(J; \bar{J}, \tau) dJ \\ &= \iint VM(\hat{s}; m, k_r) VM(m; s, \Phi(J)) Gamma(J; \bar{J}, \tau) dJ dm \\ &= \int \frac{I_0\left(\sqrt{\Phi(J)^2 + k_r^2 + 2\Phi(J) k_r \cos(\hat{s}-s)}\right)}{2\pi I_0(k_r) I_0(\Phi(J))} Gamma(J; \bar{J}, \tau) dJ \end{aligned}, \tag{S7}$$

Note that in Eq. S7, sensory measurement (m) can be analytically eliminated. Since precision J is a random variable across items and trials, we sampled it 10000 times from the Gamma distribution with mean $\bar{J}$ and scale parameter τ. Note that (1) confirmed that 500 samples are enough in the model fitting. We then used all the samples to calculate response probability in each trial.

Variable-precision-with-capacity model. The variable-precision-with-capacity (VPcap) model inherits all parameters and the structure of the VP model above, except that an additional capacity parameter (K) is introduced to estimate memory capacity of individuals. If the set size N is smaller than capacity K, the VPcap model is identical to the VP model. If the set size N exceeds the capacity K, the model assumes that the probe is stored in the VWM with the probability K/N, and out of memory with the probability 1- K/N. In the latter case, a participant randomly guesses a color. The response probability therefore can be written as:

$$p(\hat{s}\,|\,s)=\begin{cases}\frac{K}{N}p(\hat{s}\,|\,s;\bar{J},\tau)+(1-\frac{K}{N})\frac{1}{2\pi}, & K<=N\\ p(\hat{s}\,|\,s;\bar{J},\tau), & K>N\end{cases}, \tag{S8}$$

Where $p(\hat{s}\,|\,s;\bar{J},\tau)$ is defined in Eq. S7. In essence, the VPcap model is a mixture model of the VP model and a random guessing process when the set size exceeds the participant's capacity. The VPcap model has five parameters, four as the same in the VP model and the additional capacity parameter (K).

Item-limit model. The item-limit (IL) model assumes no uncertainty in the sensory encoding stage such that the internal sensory measurement m is equal to the input stimulus s. But there exists choice variability from measurement m to the reported color ($\hat{s}$). Such choice variability does not vary across set size levels. The IL model also assumes a fixed capacity K. The response probability is:

$$p(\hat{s}\,|\,s)\equiv p(\hat{s}\,|\,m)=\begin{cases}\frac{K}{N}VM(\hat{s}\,|\,s,\kappa_r)+(1-\frac{K}{N})\frac{1}{2\pi}, & K<=N\\ VM(\hat{s}\,|\,s,\kappa_r), & K>N\end{cases}, \tag{S9}$$

The IL model has two free parameters: choice variability (κ_r)and capacity (K).

Slots-plus-averaging model. The slots-plus-averaging (SA) model was originally proposed in ref. (6) and further elaborated in ref. (1). Unlike the IL model, the SA model acknowledges the presence of noise in sensory encoding stage, however the memory resource is discrete chunks and a single or multiple chunks can be assigned to one item. For one item, the SA model assumes the Eq. S4 still holds as the relationship between the resource assigned to that item and the width of the Von Mises distribution:

$$\kappa=\Phi(SJ_s)\ , \tag{S10}$$

where S is the number of chunks and J_s is the resource of one chunk. The SA model also assumes a capacity K. When $N > K$, an item should receive either 0 or 1 chunk. Then the allocation should be similar to the IL model. When $N \leq K$, some items receive either one

or more chunks. Assuming that the resource chunks should be assigned as equally as possible across items, the S can be calculated as:

$$S=\begin{cases}\left\lfloor\frac{K}{N}\right\rfloor, & \text{with probability } 1-\frac{K \bmod N}{N}\\ \left\lfloor\frac{K}{N}\right\rfloor+1, & \text{with probability } \frac{K \bmod N}{N}\end{cases}, \tag{S11}$$

where $\lfloor x \rfloor$ represents the *floor* function in Matlab. The response probability in the SA model can be written as:

$$p(\hat{s}\,|\,s)=\begin{cases}\frac{K}{N}VM(\hat{s}\,|\,s,\Phi(J_s))+(1-\frac{K}{N})\frac{1}{2\pi}, & K<N\\ (1-\frac{K \bmod N}{N})VM(\hat{s}\,|\,s,\Phi(\left\lfloor\frac{K}{N}\right\rfloor J_s))+\frac{K \bmod N}{N})VM(\hat{s}\,|\,s,\Phi(\left\lfloor\frac{K}{N}+1\right\rfloor J_s)), & K>=N\end{cases}, \tag{S12}$$

The SA model has two free parameters: unit resource J_s and capacity K.

Mixture model. The mixture model (MIX) has been used in previous clinical research (7). Similar to the IL model, the MIX model only assumes the uncertainty from stimulus s to the reported color ($\hat{s}$) and a fixed capacity K. The difference is that the uncertainty (κ) reflects both sensory noise and choice variability, and thus the uncertainty is set-size dependent (each set size has one κ). The response probability can be written as:

$$p(\hat{s}\,|\,s)=\begin{cases}\frac{K}{N}VM(\hat{s}\,|\,s,\kappa_{1/3})+(1-\frac{K}{N})\frac{1}{2\pi}, & K<=N\\ VM(\hat{s}\,|\,s,\kappa_{1/3}), & K>N\end{cases}, \tag{S13}$$

where κ_1 and κ_3 denote the uncertainty for set size 1 and 3, respectively. The MIX model has three parameters: uncertainty levels κ_1, κ_3 and capacity K.

Equal-precision model. The equal precision (EP) model is very similar to the VP model, except that an equal amount of resources is assigned to every item and in any trial. Namely, the Eq. S2 does not apply to the EP model. In the EP model, the resource assigned to one item declines as a power function (as Eq. S1). Then the resource at each

set size level can be converted to the width of the Von Mises distribution using (Eq. S4). The response probability is given by:

$$p(\hat{s} \mid s; J_1, a, \kappa_r) = \frac{I_0(\sqrt{\Phi(J_1 N^{-a})^2 + k_r^2 + 2\Phi(J_1 N^{-a}) k_r \cos(\hat{s} - s)})}{2\pi I_0(\kappa_r) I_0(\Phi(J_1 N^{-a}))}, \qquad \text{(S14)}$$

where J_1 is the resource when set size is 1 (initial resources). The EP model has three free parameters: initial resources (J_1), decaying exponent (a), and choice variability (κ_r).

Model fitting. The BADS optimization toolbox in MATLAB (8) was used to search the best-fit parameters that maximize the likelihood of responses. BADS has been shown to outperform other default nonlinear optimization algorithms in MATLAB, especially in the problems where gradients on loss function are not available or hard to compute (8). We fit all models separately to each participant. To avoid local minima, we repeated the optimization process with 20 different initial seeds that are equally spaced within a lower and an upper bound. Parameters bounds were set to be very broad to avoid bias. The parameters with the maximum likelihood value were used as the best-fit parameters for one subject.

Supplementary Note 2: Model comparisons

We compared the performance of all models fitted in this study, including the VP model, the VPcap model, the EP model, the SA model, the MIX model and the IL model. Model comparisons were performed for both groups using both Akaike information criterion (AIC) and Bayesian information criterion (BIC) (9, 10) metrics (Fig. S1). All models were compared for each subject and we derived the best model for each subject. Results showed that the VP model outperformed all other models over 85% of subjects in both groups and with respect to both AIC and BIC (Fig. S2). Particularly, the VP model is the best-fitting model in 52 out of 61 (85%) HCS and in 55 out of 60 PSZ (92%) under the AIC. Using the BIC, the VP model is the best-fitting model in 52 out of 61 HCS (85%) and 52 out of 60 (87%) PSZ. These results strongly support the idea that the VP model assuming no fixed capacity better explains the VWM behavior. This result also questions

the conventional theory whether capacity really acts as a key determinant of limiting VWM performance in PSZ.

Supplementary Note 3: Results of other suboptimal models

Fitted parameters of the VPcap model. The VPcap model is a variant of the VP model and explicitly incorporates the capacity parameter. Estimated parameters in the VPcap model largely replicated the results of the VP model (Fig. S2). Again, PSZ have larger resource allocation variability than HCS (Fig. S2B, $t(119) = 3.891$, $p = 1.65 \times 10^{-4}$, $d = 0.707$) and the two groups did not significantly differ in the resource decay function (Fig. S2A, initial resources, $t(119) = 0.012$, $p = 0.990$, $d = 0.002$; decaying exponent, $t(119) = 1.142$, $p = 0.256$, $d = 0.208$). We observed a significant larger choice variability in HCS (Fig. S2C, choice variability, $t(119) = 2.365$, $p = 0.02$, $d = 0.43$). Most importantly, the estimated capacity values of two groups were statistically comparable (Fig. S2D, $t(119) = 0.459$, $p = 0.647$, $d = 0.083$).

Comparing capacity of the two groups in suboptimal models. We further investigated the estimated capacity of all subjects in the IL, the SA, the MIX and the VPcap model, the four models having the capacity parameter. We found no significant group difference in capacity measured by all four models (Fig. S3, IL model, $t(119) = 1.554$, $p = 0.123$, $d = 0.283$; SA model, $t(119) = 1.03$, $p = 0.306$, $d = 0.187$; MIX model, $t(119) = 0.273$, $p = 0.786$, $d = 0.050$; VPcap model, $t(119) = 0.459$, $p = 0.647$, $d = 0.083$). These results directly reject the decreased-capacity account of PSZ. Notably, one previous study employed the same task and fit the same MIX model and found smaller capacity of PSZ compared with HCS (7). But we could not replicate this finding here.

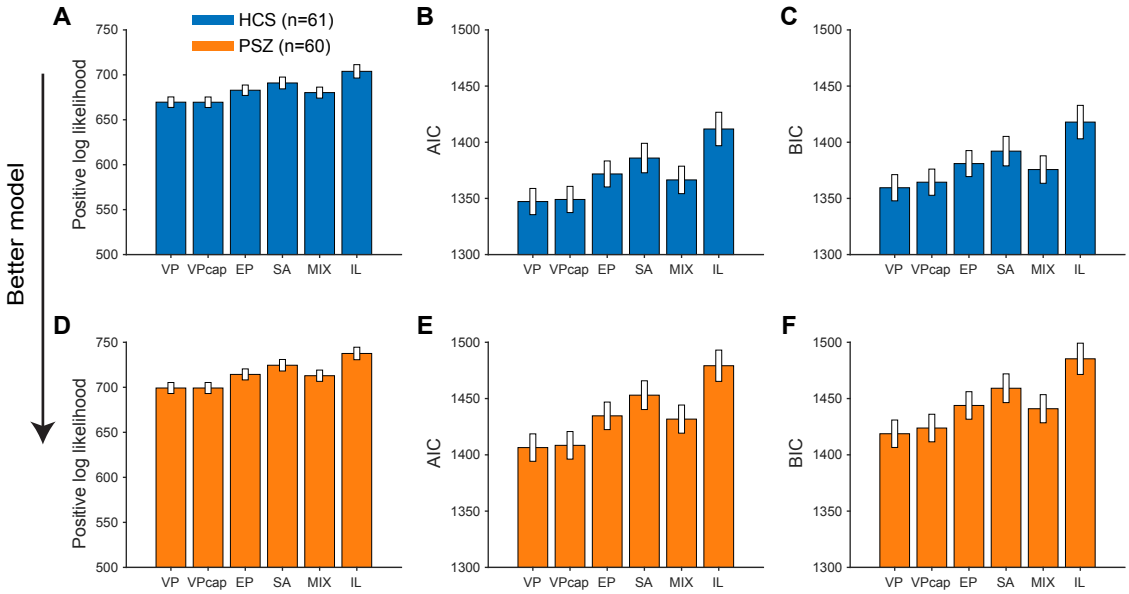


Figure S1. Positive log-likelihood (panels A, D), AIC (panels B, E) and BIC (panels C, F) values for all models. Note that here we display the positive log-likelihood values to help visually compare models since maximum negative log-likelihood values are equivalent to minimum positive log-likelihood values. Thus, in all panels a lower y-axis value indicates a better model. The upper (panels A-C) and lower (panels D-F) rows depict the model comparison results for HCS and PSZ respectively. The best-fitting model is the VP model for both groups and using both AIC and BIC metrics (also see the Fig. 3 in the main text).

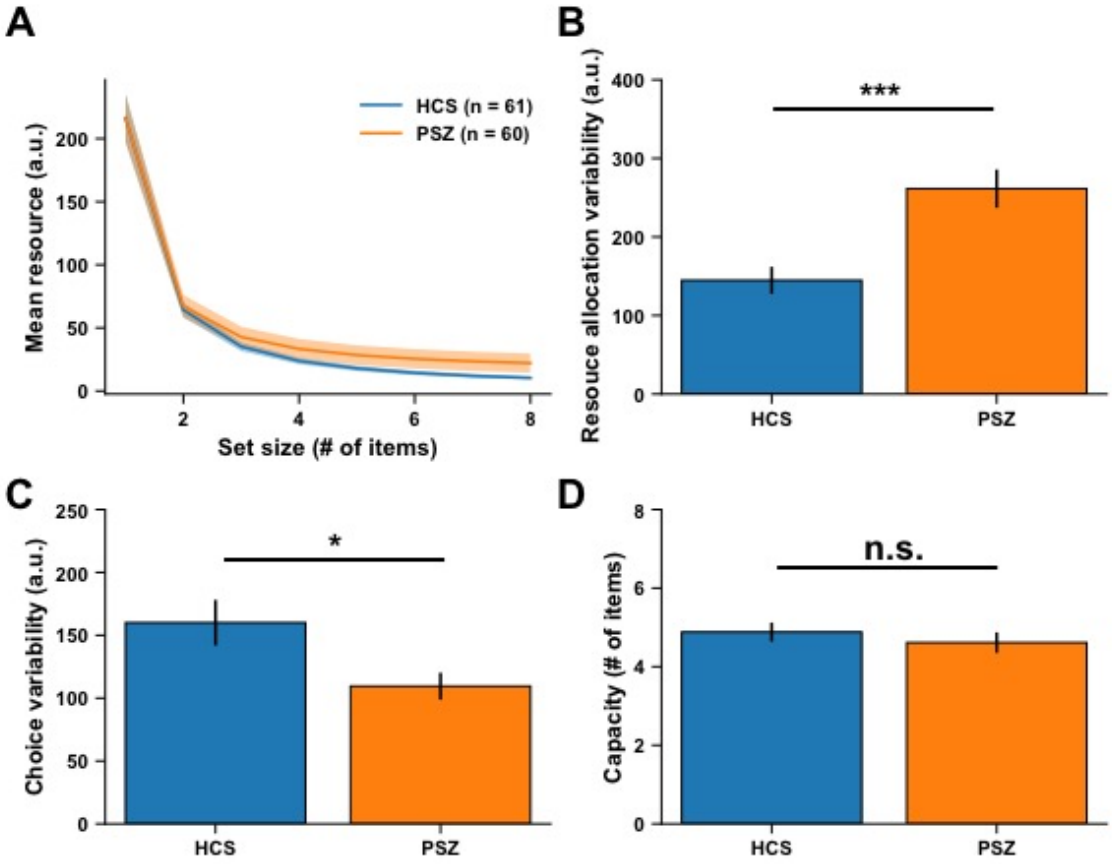


Figure S2. Fitted parameters (panel A: resource decay functions; panel B: resource allocation variability; panel C: choice variability; panel D: capacity) of the VPcap model. The results replicate the results in the Fig. 4. Furthermore, this model estimates capacity in individual subjects and the result show that the two groups have comparable capacity (panel D). All error bars are ± SEM across subjects. Other figure captions are the same as in the Fig. 4 in the main text. Significance symbol conventions are *:$p < 0.05$; ***: $p < 0.001$; n.s.: non-significant.

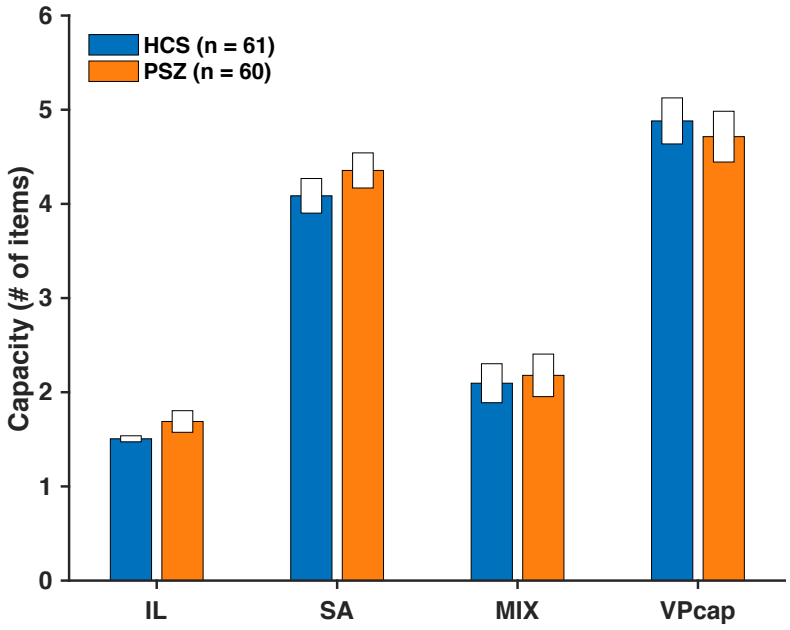


Figure S3. Capacity of the two groups measured by four suboptimal models. None of the four models reveal the significant group differences in capacity. These results directly challenge the conventional decreased-capacity account of PSZ. All error bars are ± SEM across subjects.

References

1. van den Berg R, Shin H, Chou W-C, George R, Ma WJ (2012) Variability in encoding precision accounts for visual short-term memory limitations. *Proc Natl Acad Sci* 109:8780–8785.
2. van den Berg R, Awh E, Ma WJ (2014) Factorial comparison of working memory models. *Psychol Rev* 121:124–149.
3. Bays PM, Husain M (2008) Dynamic Shifts of Limited Working Memory Resources in Human Vision. *Science* 321:851–854.
4. Bays PM, Catalao RFG, Husain M (2009) The precision of visual working memory is set by allocation of a shared resource. *J Vis* 9:7.1-11.
5. Ma WJ, Beck JM, Latham PE, Pouget A (2006) Bayesian inference with probabilistic population codes. *Nat Neurosci* 9:1432–1438.
6. Zhang WW, Luck SJ (2008) Discrete fixed-resolution representations in visual working memory. *Nature* 453:233–235.
7. Gold JM, et al. (2010) Reduced Capacity but Spared Precision and Maintenance of Working Memory Representations in Schizophrenia. *Arch Gen Psychiatry* 67:570–577.
8. Acerbi L, Ma WJ (2017) Practical Bayesian Optimization for Model Fitting with Bayesian Adaptive Direct Search. *Advances in Neural Information Processing Systems 30*, pp 1836–1846.
9. Wit E, van den Heuvel E, Romeijn JW (2012) 'All models are wrong. ': An introduction to model uncertainty. *Stat Neerl* 66:217–236.
10. Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer-Verlag, New York, NY). 2nd Ed.