Reduced Capacity but Spared Precision and Maintenance of Working Memory Representations in Schizophrenia

James M. Gold, PhD; Britta Hahn, PhD; Wei Wei Zhang, PhD; Benjamin M. Robinson, BA; Emily S. Kappenman, MA; Valerie M. Beck, BA; Steven J. Luck, PhD

Context: Working memory deficits are considered a core feature of schizophrenia. Several recent integrative articles have offered mechanistic computational and neurobiological models of the origins of this cognitive deficit.

Objective: To test predictions of these models using a new experimental paradigm from the basic science literature that makes it possible to determine whether patients with schizophrenia show (1) deficits in working memory storage capacity, (2) deficits in the precision of working memory representations, and (3) an amplification of these deficits as the retention interval increases.

Design: Case-control design. All subjects performed a color working memory test in which they were asked to recall 3 or 4 items after a 1- or 4-second delay. All subjects also received a standard measure of intelligence and the Matrics Consensus Cognitive Battery.

Setting: A tertiary care research outpatient clinic.

Patients: A total of 31 clinically stable patients with a DSM-IV diagnosis of schizophrenia or schizoaffective disorder and 26 healthy volunteers participated. The 2 groups were similar in age, sex, and ethnicity distribution.

Main Outcome Measures: (1) The number of items stored in working memory and (2) the precision of the working memory representations.

Results: Patients showed a clear reduction in the number of items stored in working memory. Patients did not differ from controls in the precision of their working memory representations. There was no evidence of delay-related amplification of impairment in either capacity or precision.

Conclusions: Patients do not show the type of imprecision or delay-dependent amplification of impairment that is predicted on the basis of current models of the neurobiology of schizophrenia. The models need to be revised to account for a pure reduction in the number of items that patients are able to store in working memory.

Arch Gen Psychiatry. 2010;67(6):570-577

Working Memory (WM) has been a major focus of recent schizophrenia research, driven by robust behavioral evidence of patient impairment and neuroimaging evidence suggesting abnormalities in neural activity during the performance of WM tasks. The clinical literature has been motivated by basic cognitive science models suggesting that WM is a critical building block of many higher cognitive functions. Further, there is extensive basic neuroscience literature suggesting that WM involves dopaminergic activity in the prefrontal cortex, and the known abnormalities in dopaminergic function in schizophrenia seem to be consistent with deficits in WM. More recently, findings from postmortem neuropathological studies of patients with schizophrenia as well as genetic findings have implicated abnormalities in the neural circuitry involved in WM.

Several investigators have recently proposed integrative theoretical accounts of the biological origins of cognitive impairment in schizophrenia. Each account involves an effort to translate the behavioral implications of basic biological findings. Lisman et al provide a circuit-based account of the implications of genetic findings involving the dopamine, glutamate, and \( \gamma \)-aminobutyric acid systems. They emphasize the cascading effect of reductions in inhibitory function needed to tune and focus cortical processing, with a particular focus on memory and sensory/perceptual processing. Durstewitz and Seaman explicitly address WM and propose that D1 hypofunction would result
in “highly unstable representations” leading to “an inability to hold and manipulate information.” Rolls et al17 address much of the same evidence from the standpoint of computational modeling, concluding that N-methyl-D-aspartate receptor hypofunction would result in a neural environment where the “stability of the attractor state is reduced, resulting in difficulty maintaining a short-term memory.” Furthermore, reductions in prefrontal dopamine function “could be measured as a decreased signal to noise ratio and impaired short-term memory performance.”17(p707)

While these accounts primarily address basic biological mechanisms, they lead to testable predictions about the types of cognitive impairment that would be expected in schizophrenia. Further, it is much easier to test these behavioral predictions than the predictions these models make about cellular activity. For example, Durstewitz and Seamans16 and Rolls et al17 imply that WM representations should be prone to accelerated decay owing to network instability. Further, Rolls et al, Durstewitz and Seamans, and perhaps Lisman et al17 suggest that WM representations in patients will have a poor signal to noise ratio, which should be evident behaviorally in the form of reduced memory precision. Here we ask whether these theoretically motivated claims, rooted in neurobiological evidence, accurately reflect the WM performance of patients with schizophrenia. To preview, we will argue that these theoretical accounts are largely at odds with the accumulated behavioral literature and we will present evidence from a new paradigm that provides direct evidence that visual WM representations are neither less precise nor more prone to decay in schizophrenia. Instead, patients exhibit a reduction in the number of items they can concurrently maintain in WM.

The overall pattern of WM findings in the schizophrenia literature does not provide much support for the idea that WM representations are less stable in patients, leading to faster decay. In a meta-analytic review of the WM literature, including 65 separate effect-size estimates with retention intervals that ranged from 1 to 30 seconds, Lee and Park2 concluded that the extent of patient impairment did not vary with length of delay interval. That is, the WM impairment in schizophrenia is just as pronounced at a 1-second delay as it is at longer delays, arguing against instability of the representations during the retention interval. However, relatively few studies have parametrically varied the retention interval, and these conclusions rely on comparisons across studies. Moreover, most studies used categorical response alternatives (eg, same vs different), which may have made it difficult to observe gradual declines in precision over time. Thus, it is possible that the methods used have not been optimal to document representational instability.

A few studies have provided evidence of reduced WM precision in patients with schizophrenia.18–20 However, the threshold estimation procedures in these studies can lead to biased threshold estimates when subjects occasionally fail to encode the stimuli, either owing to attention lapses or low WM capacity.25 Thus, the findings of these studies may reflect a higher rate of all-or-none failures of encoding rather than instability or impairment of the WM representations.

To provide a powerful test of WM instability in schizophrenia, a task must be able to directly measure the precision of WM representations, the number of representations stored in WM, and the decline in the number and/or precision of these representations with increasing delays. A new paradigm and analytic approach developed by Zhang and Luck25,26 can separately measure each of these aspects of WM performance. As illustrated in Figure 1A, participants are first shown a sample array of 3 to 4 different colors for 500 milliseconds. After a 1- or 4-second blank delay interval, one of the previous color locations is cued. Participants then indicate the color previously presented at the cued location by clicking on a color wheel displaying the entire range of possible colors.

If the cued item is present in WM, the recalled color should be close to the color of the originally presented item, with a bell-shaped distribution of errors (Figure 1B). If the cued item was not stored in WM, however, the response will be a random guess, leading to a flat distribution of errors. The observed data represent a mixture of these 2 types of trials but it is possible to separate this mixture, yielding 2 parameters that represent the 2 critical performance dimensions: (1) $P_n$ (probability in memory) represents the probability that the cued item was stored in WM and was available at time of test; (2) SD (standard deviation) represents the width of the bell curve, which is inversely related to the precision of the WM representation for trials in which it was actually present in memory. Thus, reductions in WM capacity should be evident in lower $P_n$ values, whereas reduced WM precision should be reflected in larger SD values. It should be noted that $P_n$ would also be reduced if subjects accidentally reported the color of one of the uncued items; the frequency of this type of error can be assessed by examining the distribution of responses around each of the uncued colors.

The inclusion of 2 delay intervals also makes it possible to determine whether WM representations are less stable in patients than in controls, which would yield a reduction in $P_n$ or an increase in SD over time. Most critically, a significant reduction in $P_n$ in the absence of a difference in SD would indicate that the capacity reduction in schizophrenia cannot be explained on the basis of impaired WM precision. We chose delay intervals of 1 and 4 seconds because healthy young adults begin to show a decline in performance sometime between 4 and 10 seconds.25 If WM representations are unstable in patients, they should exhibit a decline at an earlier delay than control subjects. We did not go beyond 4 seconds because longer delays may lead to an inability of patients to stay on task, artifactually producing the appearance of a WM decline.

In our view, recent theoretical accounts lead to strong predictions that patients should demonstrate reduced WM
precision (ie, an increased SD) and that the patient impairment of \( P_m \) and/or SD should be amplified as delay interval increases; each of these predictions is contradicted by the data presented here.

**METHODS**

**PARTICIPANTS**

Thirty-one patients who met DSM-IV criteria for schizophrenia (15 paranoid, 8 undifferentiated, 2 disorganized, 2 residual) or schizoaffective disorder (n=4) and 26 matched healthy control subjects participated in this study. Demographic information is summarized in the Table. Groups were virtually identical in age and parental education and did not differ in sex or ethnicity (\( P > .10 \) for both, \( \chi^2 \) test). However, patients had significantly fewer years of education than controls (\( P = .005 \), independent-samples \( t \) test).

The patients were clinically stable outpatients. At the time of testing, patients were mildly/moderately symptomatic with a mean (SD) total score of 37.7(8.0) on the Brief Psychiatric Rating Scale (range, 24-65), 36.2(14.4) on the Scale for the Assessment of Negative Symptoms (range, 14-72), and 2.5(2.5) on the Calgary Depression Scale (range, 0-12). All patients were receiving antipsychotic medication; 1 was treated with a first-generation antipsychotic; 29, second-generation antipsychotics; and 1, both. Eighteen patients received clozapine, either alone or in combination with other second-generation antipsychotics. Nineteen patients also received mood-stabilizing medications, and 9 received anxiolytic medication. Patients were taking stable medications for a minimum of 4 weeks prior to testing. Control participants were recruited from the community and had no current axis 1 or 2 diagnoses as established by the Structured Clinical Interview for DSM-IV Axis I Disorders, had no family history of psychosis, and were not taking any psychotropic medication. All participants provided informed consent for a protocol approved by the University of Maryland School of Medicine institutional review board.

**NEUROPSYCHOLOGICAL TESTING**

All participants completed the Wechsler Abbreviated Scale of Intelligence (WASI), the Wide Range Achievement Test Reading, the Wechsler Test of Adult Reading, and the Matrics Consensus Cognitive Battery. Patients tended to score lower on the WASI (\( P < .001 \), independent-samples \( t \) test), Wide Range Achievement Test Reading (\( P = .12 \)), Wechsler Test of Adult Reading (\( P = .09 \)), and Matrics Consensus Cognitive Battery (\( P < .001 \)) than healthy controls (Table).

**STIMULI AND TASK**

Stimuli were presented on a cathode ray tube monitor with a gray background (Figure 1). Each trial commenced with a fixa-
tion circle that remained visible throughout the trial. After 400 milliseconds, a sample array consisting of 3 or 4 colored squares was presented for 500 milliseconds. Each square subtended $2^\circ \times 2^\circ$ of visual angle and was presented at 1 of 8 possible positions on an invisible circle with a $4.5^\circ$ radius. A delay of either 1 or 4 seconds followed. The probe array was then presented, surrounded by a color wheel ($8.2^\circ$ radius; $2.2^\circ$ thick) consisting of 180 equally spaced equiluminant color values that covered the entire spectrum. $^{26}$ The sample array colors were randomly selected from this set with a minimum distance of $24^\circ$ between any 2 colors. The orientation of the color wheel varied randomly across trials. The probe array consisted of outlined squares at the sample locations. One of the outlined squares was thicker than the others, indicating the item to be recalled. Subjects reported the color remembered at this location by mouse-clicking on the appropriate location in the color wheel. The probe array and color wheel remained visible until a response was made. After the response, a feedback arrow indicated the correct location on the color wheel for 1000 milliseconds. After an intertrial interval of 600 milliseconds, the next trial began.

The 3- and 4-item versions of the task were tested in separate sessions on separate days in counterbalanced order. The 1- and 4-second delay intervals were equally likely and were randomly intermixed within each session. A total of 150 trials were presented at each delay in each session.

Each session began with 2 control tasks, 1 testing motor precision (20 trials) and 1 testing color perception precision (30 trials). To minimize memory requirements in these control tasks, the colored squares and color wheel were presented simultaneously and remained visible until a response was made. In the motor control task, 1 square was always white, and a thin white bar was presented at a random location on the color wheel. The task was to mouse-click on the white bar. In the sensory control task, 1 colored square was outlined, indicating that its color should be reported by clicking on the color wheel. After each response, an arrow indicated the correct location. Subjects were given no instructions regarding the use of verbal coding but were able to control the mouse just as well as control participants between patients’ task performance ($K$ and SD scores) and their WASI intelligence quotient (IQ) scores, total Matrics battery scores, and capacity estimate ($K$) from the change localization task.

STATISTICAL ANALYSIS

Raw data consisted of the degree of error on each trial, ie, the distance between the reported color and the original color value. Trials on which the probed item was not encoded into memory will yield a uniform distribution of error. In contrast, in trials on which the probed item was encoded, the recalled value will tend to be near the original color, and the error will follow a von Mises distribution (the circular analog of the Gaussian distribution). The 2 types of trials are mixed together in the data. As described by Zhang and Luck, $^{23,24}$ a maximum likelihood algorithm $^{39}$ was used to derive $P_e$, the probability that the probed item was present in memory, and SD, which is inversely related to the precision of the representation when the probed item was present in memory, and SD, which is inversely related to the precision of the representation when the probed item was present in memory.

The model provided an excellent fit to the data, accounting for 99% of the variance in both patient and control participants (adjusted $R^2$ for the pooled data of each group). As seen in Figure 2A, patients exhibited lower performance in the motor control task than controls, $P < .01$. This difference did not interact with SS ($P > .9$). The $K$ value was essentially at ceiling for both groups at both SSs in the sensory control condition. The $K$ value was essentially at ceiling for both experimental groups.

MOTOR AND SENSORY CONTROL TASKS

The mean error of responses in the motor control task (clicking on a thin white bar) was close to zero and did not differ between patients and controls ($t_{15}=0.01; P > .9$), indicating that patients were able to control the mouse just as well as control participants. In the sensory control task, the precision of color matching was lower for patients than control participants; this was confirmed by a main effect of group ($F_{1,15}=5.42; P < .02$). The $K$ value was essentially at ceiling for both groups at both SSs in the sensory control condition, indicating that both groups of subjects understood the task and could report the color of the cued item when it was visible at the time of report.

NUMBER OF ITEMS REPRESENTED IN WM ($K$)

The model provided an excellent fit to the data, accounting for 99% of the variance in both patient and control participants (adjusted $R^2$ for the pooled data of each group). As seen in Figure 2A, patients exhibited lower performance in the motor control task than controls, $P < .01$. This difference did not interact with SS ($P > .9$). The $K$ value was essentially at ceiling for both experimental groups.
memory capacity (K) than controls at both SS3 and SS4, with a similar between-group difference at the 1- and 4-second delay intervals. Overall, K was slightly higher at SS4 than at SS3, which probably reflects a ceiling on performance for some subjects at SS3. Values for K remained constant across the 1- and 4-second delays in the control group, as was previously observed with healthy college students, and there was also no sign of a decline in patients. These impressions were statistically supported by a main effect of group (F1,55 = 4.22; P < .05) and a main effect of SS (F1,55 = 7.42; P < .01) but no significant main effect (P > .5) or interaction (P > .2) involving delay. The between-group difference was somewhat larger at SS4 than SS3, such that controls displayed a steeper increase in K from SS3 to SS4, but the group x SS interaction fell short of significance (P = .13). The overall effect size for the between-group K difference was 0.56, very close to the meta-analytic results finding of 0.459 for visuospatial WM.

One possible explanation of the reduced K in patients is that they had difficulty binding the colors to their locations, causing them to report the color of one of the wrong items. We assessed this possibility by examining the distribution of responses relative to the unprobed colors, treating each unprobed item as if it were the probed item and estimating Pm and SD. We found that the distribution of responses around the unprobed items was essentially flat, and Pm for the unprobed items was near zero (<0.003) and did not differ significantly between groups (P > .3). Thus, the reduced K observed for patients in the main analysis was not a consequence of reporting the color of the wrong item.

### Precision of Stored Representations (SD)

As seen in Figure 2B, WM precision was very similar for patients and controls at both SSs and at each delay interval. Indeed, there was no main effect of group (F1,55 < 1; P > .5) on SD and no interaction involving group (P > .4). A main effect of SS (F1,55 = 6.58; P < .02) reflected somewhat lower memory precision at SS4 than SS3 in all participants. Most importantly, there was no main effect (P > .5) or interaction (P > .2) involving delay. Thus, the precision achieved at a 1-second delay was fully maintained over the 4-second interval, which matches findings from healthy college-aged subjects.

### PERFORMANCE CORRELATIONS

To determine whether estimates of WM capacity from the color wheel paradigm are similar to those observed with more conventional visual WM tasks, we examined the correlation between K estimates derived from the color wheel paradigm (averaged over delays and SSs) and from the change localization task. As expected, K values derived from the change localization task were significantly decreased in patients relative to controls (t55 = 3.51; P < .001). The K scores for the 2 tasks were strongly correlated in both controls (R = 0.63; P < .001) and patients (R = 0.65; P < .001). Thus, both tasks appear to be measuring a similar ability in both groups.

The correlations between K from the color wheel paradigm (averaged over delays and SSs) and WASI IQ scores and the overall T score from the Matrics battery are shown in Figure 3. The control participants displayed a remarkable correlation between K and total Matrics score (R = 0.89; P < .001) and a moderate correlation between K and WASI IQ scores (R = 0.51; P < .01). The SD correlated significantly with total Matrics score (R = -0.46; P < .02) but not WASI IQ (R = 0.28; P = .17). These correlations were attenuated in patients. The correlation of K with total Matrics scores was significant (R = 0.41; P < .03) but significantly weaker than in controls (Fisher z-transformation test for difference in correlation: z = 7.01; P < .001). The correlation of K with WASI IQ was R = 0.24 (P = .2) in patients. In both patients and controls, similar but smaller-magnitude correlations were observed with the Matrics WM domain score as for the overall T score. The SD was not significantly correlated with either the total Matrics score (R = -0.12; P > .5) or WASI IQ (R = -0.18; P > .3) in patients.

### COMMENT

These results provide several important insights into the nature of WM impairment in schizophrenia that constrain models linking cognitive deficits to the underlying neurobiological abnormalities. Patients show clear re-
ductions in the number of items that can be stored in WM but no evidence that their WM representations are less precise or less stable than those of healthy individuals. Although it is possible that schizophrenia leads to less stable and less precise representations of other types of stimuli, the present results demonstrate that WM capacity reduction can and does occur in the absence of impairments in WM precision.

We observed no evidence of delay-related magnification of patient WM impairment; the patient deficit was equally robust at the 1- and 4-second retention intervals. When combined with the lack of a reduction in precision, this absence of a magnification of impairment at a longer delay provides convincing evidence against the proposal that WM representations are unstable or inherently more noisy in schizophrenia. Moreover, this finding is consistent with earlier meta-analytic results showing a lack of delay dependency. However, the present results go beyond previous visual object WM studies by using a task that involves a fine-grained report, making it possible to separately measure the capacity and precision of WM.

It is possible, of course, that evidence of instability could be obtained at longer delay intervals. In healthy college-aged subjects, increasing the delay interval to 10 seconds resulted in a decline in capacity but no significant decline in precision, and it is possible that patients would show a decline in precision or a sharper decline in capacity at longer intervals. However, visual WM representations are typically used for periods ranging from a few hundred milliseconds to a few seconds in most real-world tasks. If schizophrenia involves a meaningful level
of WM instability that is important for other cognitive operations, then it should be evident by a 4-second delay interval. Further, the interpretation of impairments at long delay intervals may be complicated by the contribution of long-term memory systems or intermittent failures in goal maintenance.

Consistent with prior studies in healthy subjects, we observed a remarkably robust relationship between WM capacity and measures of general intellectual and cognitive ability. Indeed, the degree of covariation exceeded the levels typically documented in the literature. It will remain for future studies to determine if this is owing to the unusual measurement accuracy offered by the color wheel paradigm or an unusual group of healthy participants. Note, the lower level of correlation seen in the patient cohort is more typical of the magnitude of relationship between WM capacity and cognitive ability in healthy populations. However, it is intriguing that the relationship between WM capacity and general ability is different in patients than in healthy subjects. Perhaps, as capacity is pathologically decreased, different systems are engaged in a compensatory fashion.

Note that these data do not and cannot contradict the biological findings reviewed by Lisman, Durstewitz, and Rolls et al. Instead, they contradict the postulated links from biology to behavior. One of the challenges facing the field is the need to accurately translate the implications of findings across levels of evidence (from genes to cells, to systems, to behavior) so that progress in one area serves to drive progress in another. Such progress requires that models at more basic levels be constrained by an accurate understanding of the behavioral endpoints that characterize the illness. These are the targets that need to be “hit” by models and theories. In our view, the recent biological accounts discussed above are at odds with much of the behavioral literature, and clearly at odds with the data presented here.

Before accepting this assertion, it is important to consider the limitations of the present findings. Our patients were stably treated outpatients with chronic schizophrenia, many treated with clozapine. Thus, our results may not generalize to less treatment-resistant cohorts or to unmedicated patients with early illness. Also, as in most studies of visual WM in schizophrenia, the present study examined WM performance for relatively simple stimuli. Additional mechanisms may come into play for more complex objects, and the present study would have been unable to detect impairments in these mechanisms. However, the predictions of the biological models are clearest for simple stimuli, for which precision is well defined.

It is also possible that our findings might prove to be specific to WM for color or other ventral stream features. Indeed, the best evidence of impaired WM stability in schizophrenia comes from studies showing delay-dependent drift in spatial memory. Nonetheless, our data demonstrate that WM storage capacity can be impaired without degradation in WM precision in at least 1 very common WM task. Moreover, the biological models provide no reason to suspect that WM representations would be any more stable for nonspatial information than for spatial information. If further studies confirm that dorsal stream WM representations are unstable in patients with schizophrenia but ventral stream WM representations are not, then this will focus future theoretical efforts on the differences in circuitry between these types of representations.

It is also important to question the sensitivity of our methods. That is, might the experimental paradigm simply lack sensitivity to a change in precision? This is unlikely. Zhang and Luck showed that several experimental manipulations significantly affected the SD measure in healthy controls. In the present study, SD was significantly smaller in the perceptual matching condition than in the memory conditions, and SD was significantly larger at SS4 than SS3. Moreover, SD correlated significantly with measures of cognitive ability. Thus, the SD measure is sensitive to both experimental manipulations and individual differences. How, then, can we account for the observation of group differences in the perceptual control condition but similar WM precision? That is, how could WM precision be normal in the face of degraded sensory input? We suspect, but cannot prove, a very simple answer. In the control task, the sample array remained on the screen until a response was made. If controls (more so than patients) looked back and forth between the sample array and the color wheel in the perceptual control condition, this would have decreased the SD in this condition but not in the WM task, where encoding time was controlled. Unfortunately, we did not record response times or monitor eye movements, the evidence that is needed to confirm the proposed explanation. Note, however, that the purpose of the perceptual control condition was to aid in the interpretation of any differences in SD in the memory conditions; because patient and control SDs were nearly identical in the memory conditions (26.86 vs 26.13, averaged over conditions), the perceptual control condition was not needed for this purpose.

Left unanswered is the question of the origins of WM capacity limits in schizophrenia. Might capacity reduction result from slowed encoding in patients? We consider this unlikely because, in a previous study, we found nearly identical WM performance using 100- and 500-millisecond sample array exposures in both patients and healthy controls. Unfortunately, there is very little understanding of the origins of capacity limits in the basic cognitive neuroscience literature. Neuroimaging studies have implicated the posterior parietal cortex, including the intraparietal sulcus, as likely contributors to visual WM capacity limits, and dysfunction of these areas, or dysfunctional interaction of these areas and the prefrontal cortex, may be implicated in the patient deficit. Unfortunately, animal physiology studies have not required subjects to retain multiple items concurrently in WM; we therefore lack direct, detailed knowledge of the circuitry underlying WM capacity limitations. The field, therefore, has a great need of neurobiological models that can explain the nature of WM deficits in schizophrenia. However, these models must accurately capture the behavioral endpoint, which is characterized primarily by reductions in storage capacity and not by an instability of the WM representations.

Submitted for Publication: May 27, 2009; final revision received October 22, 2009; accepted November 16, 2009.