Schizophrenia Bulletin vol. 45 no. 4 pp. 804–812, 2019
doi:10.1093/schbul/sby134
Advance Access publication 26 September 2018

Working Memory Impairment Across Psychotic disorders

James M. Gold*[,1], Deanna M. Barch[2], Leah M. Feuerstahler[3], Cameron S. Carter[4], Angus W. MacDonald III[5], J. Daniel Ragland[4], Steven M. Silverstein[6,7], Milton E. Strauss[8], and Steven J. Luck[9]

[1]Department of Psychiatry, Maryland Psychiatric Research Center, University of Maryland School of Medicine, Baltimore, MD; [2]Department of Psychological and Brain Sciences, Washington University in St. Louis, St. Louis, MO; [3]Graduate School of Education, University of California at Berkeley, Berkeley, CA; [4]Department of Psychiatry and Behavioral Sciences, University of California at Davis, Davis, CA; [5]Department of Psychology, University of Minnesota, Minneapolis, MN; [6]Rutgers University Behavioral Health Care, Piscataway, NJ; [7]Robert Wood Johnson Medical School Department of Psychiatry, Rutgers University, Piscataway, NJ; [8]Department of Psychology, University of New Mexico, Albuquerque, NM; [9]Department of Psychology, Center for Mind and Brain, University of California at Davis, Davis, CA

*To whom correspondence should be addressed: tel: 410-402-7871, fax: 410-402-7198, e-mail: jgold@mprc.umaryland.edu

Background: **Working memory (WM) has been a central focus of cognitive neuroscience research because WM is a resource that is involved in many different cognitive operations. The goal of this study was to evaluate the clinical utility of WM paradigms developed in the basic cognitive neuroscience literature, including methods designed to estimate storage capacity without contamination by lapses of attention.** ***Methods:*** **A total of 61 people with schizophrenia, 49 with schizoaffective disorder, 47 with bipolar disorder with psychosis, and 59 healthy volunteers were recruited. Participants received multiple WM tasks, including two versions each of a multiple Change Detection paradigm, a visual Change Localization paradigm, and a Running Span task.** ***Results:*** **Healthy volunteers performed better than the combined patient group on the visual Change Localization and running span measures. The multiple Change Detection tasks provided mixed evidence about WM capacity reduction in the patient groups, but a mathematical model of performance suggested that the patient groups differed from controls in their rate of attention lapsing. The 3 patient groups performed similarly on the WM tasks. Capacity estimates from the Change Detection and Localization tasks showed significant correlations with functional capacity and functional outcome.** ***Conclusions:*** **The patient groups generally performed in a similarly impaired fashion across tasks, suggesting that WM impairment and attention lapsing are general features of psychotic disorders. Capacity estimates from the Change Localization and Detection tasks were related to functional capacity and outcome, suggesting that these methods may be useful in a clinical context.**

Key words: capacity limitations/schizophrenia/psychosis/cognitive impairment/working memory

Introduction

Working memory (WM), the temporary maintenance, and oftentimes, manipulation of information in the service of behavioral goals, has been a focus of basic and clinical neuroscience research.[1–4] WM is a resource required by multiple cognitive operations, and differences in WM capacity are related to differences in broad cognitive ability.[5,6] Consequently, WM deficits are of particular interest because they could underlie the broad cognitive impairment observed in psychotic disorders.

Clinically, WM has been assessed using tasks such as digit span, spatial span, and letter-number span. On these measures, people with schizophrenia (SZ) and schizoaffective disorder (SZAFF) are impaired (approximately 1 SD below the healthy control [HC] mean), whereas people with bipolar (BP) disorder typically score in-between SZ and HCs.[7,8] When the individuals with BP disorder have a history of psychosis, impairment levels are often similar to that seen in SZ[9–15]). These tasks provide coarse measures of WM and can be impacted by impairments in motivation, attention, and executive control as well as reduced storage capacity. Thus, different measurement approaches are needed to better isolate WM capacity. Here we present results on 3 approaches to measuring WM performance.

Variants of the visual Change Detection task are widely used to measure WM capacity in the cognitive neuroscience literature.[16] Participants first see an encoding array (typically 1–8 colored rectangles presented for 100–500 ms), followed by a delay (typically 1–4 s). Then a test array appears, and the participant indicates whether any items changed between the encoding and test arrays. All items are identical on 50% of trials, and one item has

changed in color on 50%. By examining accuracy across memory set sizes, it is possible to estimate an individual's WM storage capacity (termed *K*). Visual change tasks are of special interest in SZ because: (1) patients with SZ demonstrate robust reductions in *K*[17,18]; (2) both *K* and measures of executive control independently contribute to discriminating patients from controls[19]; (3) the K reductions observed in SZ do not reflect impairments of selective attention or distractibility,[20–22] suggesting that K is a relatively pure measure of WM capacity; and (4) based on fMRI evidence, reductions in K appear to primarily reflect posterior parietal cortex dysfunction rather than the broad prefrontal–parietal network that mediates more complex WM tasks such as the n-back.[23,24]

We sought to address several limitations of the canonical Change Detection paradigm. First, estimating *K* requires using arrays that exceed capacity, leading to chance performance levels. Guessing adds measurement noise, requiring large numbers of trials to obtain reliable capacity estimates. Second, the task may be experienced as being very difficult when arrays exceed capacity, potentially leading to reduced engagement in patients relative to controls. Third, performance may be impacted by lapses of attention (ie, mind-wandering resulting in a failure to sustain task engagement), and differences in the ability to maintain attention to the task will impact WM capacity estimates. In a previous study of visual sensory processing, we used "catch" trials (trials so easy that anyone paying attention should be able to perform at 100% correct) to estimate the lapse rate, and found that lapses explained the patient impairment.[25] Thus, quantifying attention lapses may be important for accurately measuring WM capacity.

To address these issues, we used a modified paradigm called Multiple Change Detection, in which participants are presented with a 5-item encoding array (which exceeds the WM capacity of nearly all adults).[2] After a short delay, they are shown a 5-item test array that has 0, 1, 2, or 5 items that differ from the encoding array, and they indicate whether or not any changes were detected. This task addresses the shortcomings noted earlier. First, by varying the number of changes, it is not necessary to use arrays that are clearly supra-capacity as is typical of many Change Detection paradigms. Second, because multiple items may change with a constant display size, the task appears to be less difficult. Third, the 5-change trials serve as "catch" trials and using this method and a mathematical modeling approach (see Broadway and Engle[26] and supplementary materials), it is possible to estimate WM capacity unconfounded by attention lapses

We also used a Change Localization task where there is a single change on the test array on every trial, and the participant is asked to click on the location of the changed item. Chance performance is lower on Change Localization than on standard Change Detection, increasing reliability of WM capacity estimates. People with SZ demonstrate robust deficits on Change Localization and task performance correlates with overall neuropsychological ability in both SZ and HCs.[18] However, Change Localization performance can be impacted by attention lapses and the task design precludes catch trials.

The third approach comes from the individual differences literature where different "complex span" tasks have demonstrated robust correlations with general cognitive ability.[5,6] In Running Span, people are given a sequence of 2–8 letters on each trial and when the sequence stops they must report the final 1–6 letters in order. Thus, the task involves "flushing" and updating items held in WM as the list length gets longer suggesting that overall task success requires more than simple storage capacity as in visual Change Detection.[5,27] We also included the Letter-Number Span test where people are presented with an alternating series of numbers and letters, ranging from 2 to 7 items and are asked to report the numbers in ascending order followed by the letters in alphabetical order. As with Running Span, success on this task requires the ability to manipulate stored information, not just simple storage capacity, and people with SZ show robust impairments on this task.[28–30] It should also be noted that Running span performance is less influenced by the automaticity of retrieving the alphabet than is Letter-Number Span.

We had several goals in this study. By comparing tasks that assess simple storage capacity vs tasks that involve more active updating and manipulation of information stored in WM we sought to better characterize the nature of WM impairment observed in psychotic disorders. This was of particular interest in BP disorder where no prior studies have used visual Change Detection measures. Second, we wanted to examine construct validity of the experimental tasks by examining their intercorrelations and correlations with Letter-Number Span. Third, we also wished to examine ecological validity by examining correlations with measures of functional capacity, community functioning, and symptom severity.

Methods

Participants

A total of 216 participants were recruited across 5 CNTRACS sites: Washington University in St. Louis, University of California—Davis, University of Maryland School of Medicine, Rutgers University, University of Minnesota—Twin Cities from the following groups: (1) HCs (N = 59); (2) *DSM-IV*[31] SZ (N = 61); (3) *DSM-IV* SZAFF (N = 49); and (4) *DSM-IV* BP disorder with lifetime psychosis (N = 47). The clinical samples were recruited from outpatient clinics and day programs as well as using flyers and online advertisements. Similar online and flyer methods were used to recruit HCs. All patients were considered to be clinically stable at the time of testing. All study procedures were approved by each site's institutional review board. Details on inclusion/exclusion criteria and screening procedures are in supplementary

materials. These data are drawn from a protocol that involved WM and reinforcement learning tasks.[32]

The groups had similar demographics although mean levels of personal education and Wechsler Tests of Adult Reading[33] scores were significantly higher in the HC group than in the 3 clinical groups (table 1). The SZ and SZAFF groups were on higher doses of olanzapine equivalent medication doses than the BP group, but there were no correlations between olanzapine dose and WM performance.[34]

Diagnosis and Clinical Assessment

A masters-level clinician conducted diagnostic assessments using the Structured Clinical Interview for *DSM-IV-TR1*,[35] the Brief Psychiatric Rating Scale (BPRS),[36,37] the Young Mania Rating Scale,[38] the Bipolar Depression Rating Scale,[39] and the Clinical Assessment Interview for Negative Symptoms (CAINS).[40] We used the participant and informant versions of the Specific Levels of Functioning Scale[41] to assess community functioning, and the University of California San Diego Performance-based Skills Assessment-Brief (UPSA-B)[42–44] to assess functional capacity. We also administered the Hopkins Verbal Learning Test and the Brief Assessment of Cognition in Schizophrenia Symbol Coding subtests from the MATRICS Consensus Cognitive Battery (MCCB).[28]

WM Tasks

Change Localization

Participants first saw a 5-item encoding array for 500 ms (figure 1) at a viewing distance of 70 cm. The items were equally spaced around an imaginary circle that was centered on the screen and spanned 1/3 of the screen's vertical height. The encoding array was followed by a 1000-ms delay and then the test array. Participants used the mouse to click on the one changed item. Accuracy was stressed, and responses were untimed. In one version of the task (figure 1A), the 5 items were colored rectangles, one of which changed color in the test array. In a second version (figure 1B), the items were colored shapes (eg, circle and star), and the changed item differed in both shape and color from the corresponding item in the encoding array. Each version had 60 trials with order counterbalanced across participants. The WM capacity score was calculated by multiplying the percentage of correct responses by the memory array size.[45]

Multiple Change Detection

This task used the same stimuli, presentation arrangement, and delay interval as described earlier. However, the test array contained 0, 1, 2, or 5 changed items (P = .25 for each) relative to the encoding array. Participants indicated change or no change using left and right index fingers on a keyboard. A total of 240 items were presented in each version, 60 of each trial type.

We applied a mathematical modeling procedure to the accuracy scores for each trial type providing an estimate of WM capacity (K), the probability that the participant was paying attention on a given trial (A), and a guessing bias parameter (G) (see Feuerstahler *et al.*[26] and supplementary materials).

Table 1. Demographic and Clinical Characteristics of the Sample

	HC (N = 59)		SZ (N = 61)		SZAFF (N = 49)		BP (N = 47)		Group Differences
	M	SD	*M*	SD	*M*	SD	*M*	SD	
Age	35.7	10.9	37.5	11.6	38.9	11.5	35.7	10.0	NS
Gender (% Female)	47.5		37.7		44.9		53.2		NS
Race (% Black)	28.8		41.0		24.5		25.5		NS
Personal education	15.1	2.2	13.3	2.3	13.8	2.9	13.9	2.5	HC > SZ, SZAFF, BP
Parental SES	45.9	13.4	44.5	14.3	45.3	15.4	47.6	16.0	NS
WTAR	39.6	8.5	31.9	9.3	36.9	9.0	34.5	11.5	HC > SZ, BP; SZAFF
BPRS positive	—		7.2	4.1	8.2	3.3	4.2	1.9	SZ, SZAFF > BP
BPRS negative	—		7.7	3.1	7.4	2.2	6.0	2.3	SZ, SZAFF > BP
BPRS disorganization	—		5.0	1.7	4.7	1.2	4.6	1.1	NS
BPRS depression	—		7.5	3.3	11.2	4.6	9.9	4.7	SZAFF, BP > SZ
BPRS mania	—		6.6	2.2	7.0	2.6	6.7	2.5	NS
YMRS	—		8.2	6.0	11.6	6.9	7.4	7.1	SZAFF > SZ, BP
BDRS	—		9.3	5.2	14.7	7.6	12.2	8.3	SZAFF > SZ > BP
CAINS ANH/AMOT	—		10.6	5.9	11.3	5.6	7.6	4.8	SZ, SZAFF > BP
CAINS blunting	—		3.7	3.3	2.4	2.0	1.3	2.1	SZ > SZAFF > BP
SLOF self-report	—		4.3	0.5	4.2	0.4	4.3	0.5	NS
SLOF informant	—		4.3	0.5	4.2	0.6	4.3	0.4	NS

Note: WTAR, Wechsler Test of Adult Reading; BPRS, Brief Psychiatric Rating Scale; YMRS, Young Mania Rating Scale; BDRS, Bipolar Depression Rating Scale; CAINS, Clinical Assessment Interview for Negative Symptoms; SLOF, Specific Levels of Function Scale; SZ, Schizophrenia; SZAFF, schizoaffective disorder; BP, bipolar disorder; HC, healthy controls; SES, socioeconomic status; ANH/AMOT, anhedonia/amotivation.

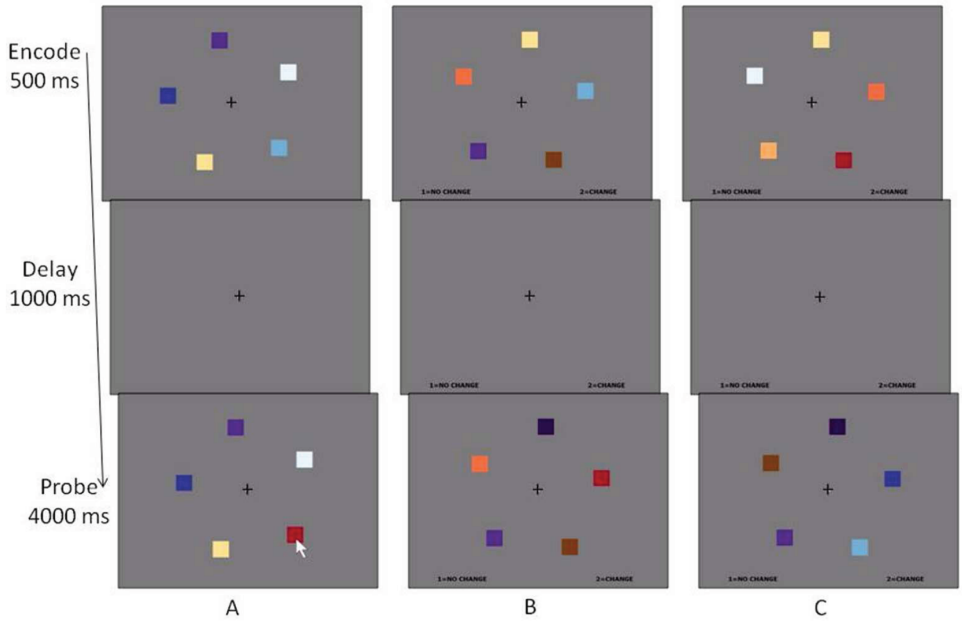

Fig. 1. Illustration (not to scale) of the Change Localization (A) and Multiple Change Detection tasks (B, C). In Change Localization a single item (circled) always differed between the encoding array and the test array. Panel B illustrates a trial where 2 items change from the encoding array to the test array. Panel C illustrates a "catch trial" from the Multiple Change Detection task where all 5 items changed on the test array.

Running Span

Participants were instructed to report the final 2, 3, 4, 5, or 6 letters presented in a list that varied from 2 to 8 letters. Instructions (Report last 2, etc.) were shown before each list presentation. In the original version, each of the 5 report sizes was tested with 3 consecutive lists, presented in random order. Then all five report sizes were retested in random order. Stimuli were presented at central fixation for 500 ms, with 200 ms inter-stimulus intervals. Total 30 trials were administered, 6 at each report size. The dependent measure was the total number of items recalled in correct order. We also administered an "adaptive version" to determine if we could reliably estimate span with fewer supra-capacity trials. In this version, participants start with a block of 4 trials at report size 1. If they responded correctly to at least 2 of the trials, they advanced to the next report size. This continued until they failed to get 2 of 4 trials correct for a given report size or until report size 6 was completed (maximum of 24 trials). The dependent measure was the total number of items recalled in correct order (see supplementary materials for adaptive span results).

Data Analysis

To streamline presentation, most analyses are conducted on the standard Running Span task and the colored square versions of the Change Localization and Multiple Change Detection tasks as these can be compared with others in the literature. These are also the versions that we recommend to other researchers. Analyses of the other task versions are provided in supplementary materials. The main dependent measures were all examined with analysis of variance (ANOVA) models that examined the main effect of psychotic illness (HC vs patients [collapsed across diagnosis]), recruitment site, and the interaction of diagnosis by site. For the capacity and attention parameters from the Change Detection task, the ANOVA models were fit using weighted least squares regression, taking 1-root-mean-square error as the weight as a way to control for model fit. When the main effect of diagnostic group was significant, we then used least significant difference tests to compare the groups to one another. We implemented false discovery rate (FDR) correction within the comparison of HC to each patient group (which we expected to be significant) and in comparing patient groups to each other (which we did not expect to be significant). We had power ranging from 78% to 86% to detect medium effect size differences between groups. Pearson's correlations were used to examine intertask and clinical correlates. We used both task versions for a factor analyses to identify the factor structure of the WM tasks.

Results

Change Localization (colored squares)

As seen in table 2, HC had higher WM capacity (K) than the total patient group ($F(1,215) = 10.794$, $P = .001$). In post hoc tests, the HC differed from each of the individual diagnostic groups (all P's $< .05$), all of which passed FDR correction. None of the patient groups differed significantly from one another (all P's $> .67$).

Multiple Change Detection (colored squares)

HCs had higher estimated WM capacity (*K*) than the total patient group ($F(1, 214) = 4.822$, $P = .029$). Post hoc tests indicated that HC > BP ($P = .007$, passes FDR), with no significant differences between HC and SZ ($P = .054$), or between HC and SZAFF ($P = .336$). The estimated rate of attentiveness was greater in HC than in the total patient group ($F(1,214) = 8.7$, $P = .004$). In post hoc tests, HC had fewer lapses than all patient groups (all *P*'s < .05 and passing FDR). No significant differences were observed between the SZ, SZAFF, and BP groups on the WM capacity (all *P*'s > 0.20), or attention measure.

Running Span

The HC group had higher running span performance than the total patient group (F (1, 215), = 28.137, $P < .001$). In post hoc tests, HC performed better than each diagnostic group (all *P*'s < .02, passing FDR), and BP > SZ ($P = .017$), though this last comparison did not pass FDR correction.

Relationships Among the WM Measures

All of the WM measures, including the Attention parameters, correlated with each other (*P*'s < .05) in both groups (supplementary tables S1 and S2). To address if they were assessing one or more constructs, we conducted an exploratory factor analysis (including both versions of each WM measure and Letter-Number Span to bolster ability to detect factors) using maximum likelihood extraction and oblimin factor rotation. We first conducted a parallel analysis, which suggested that 3 factors should be retained (table 3). The first factor included the capacity measures from the Change Detection/Localization tasks. The second factor included the two Running Span measures and Letter-Number Span, whereas the third factor included the Attention parameters from the Multiple Change detection tasks. Thus, the factor analysis provides evidence for 3 distinct constructs: visual WM capacity, complex WM span, and attentiveness.

In the 3-factor solution, the Attention parameter from the Change Detection Shape measure was a Heywood case, meaning that the estimated unique variance for this measure was negative. Heywood cases occur when there are a small number of measures that belong to a factor—as is the case here with only 2 measures of attention. Moreover, our attention measures depend on the untested assumption that informed and uniformed guessing behavior is identical. Thus, better measures of attention need to be established to definitively establish attention as a distinct factor (although, note that the 2 attention measures are correlated, 72 in the combined sample). Finally, we ran the factor analysis again, removing the 2 attention measures. Two factors were retained by the parallel analysis identified by the visual WM capacity tasks and the complex WM span tasks.

Correlations With Clinical Measures and Community Function

We examined correlations with the UPSA-B as a measure of functional capacity and with the SLOF to assess functional outcome. As seen in table 4, all of the WM measures correlated with UPSA-B performance (the other MCCB subtests are shown for comparative purposes). The capacity measures from both Multiple Change Detection and Change Localization significantly correlated with both self- and informant-report SLOF ratings. Neither Running Span nor the Attention parameter correlated with functional outcome. Further, using the methods for comparing correlated correlation coefficients,[46] the correlation between Multiple Change Detection and informant-reported ratings of community function was significantly stronger than the correlation between the attention parameter and informant-rated function.

Table 2. Cognitive Task Performance Across Groups

	Means and SD				Effect Sizes (Cohen's *D*) Compared to Controls		
	HC	BP	SZAFF	SZ	BP	SZAFF	SZ
CD SQ K	2.79 (0.84)	2.44 (0.92)	2.65 (0.86)	2.50 (0.84)	0.40	0.16	0.35
CD SQ A	0.93 (0.17)	0.87 (0.18)	0.84 (0.17)	0.87 (0.16)	0.34	0.53	0.36
CD SQ G	0.13 (0.11)	0.19 (0.19)	0.21 (0.15)	0.20 (0.17)	−0.40	−0.62	−0.49
CL SQ K	2.40 (0.77)	2.07 (0.76)	2.22 (0.77)	2.01 (0.90)	0.43	0.23	0.47
Running Span	75.63 (19.97)	65 (20.78)	59.90 (17.67)	5716 (18.38)	0.52	0.83	0.96
BACS Symbol Coding T	51.32 (10.68)	41.71 (10.94)	39.94 (11.32)	40.92 (12.70)	0.89	1.04	0.89
HVLT T	48.76 (9.08)	44.24 (8.18)	42.47 (8.41)	40.66 (9.65)	0.52	0.72	0.86
LN Span T	50.78 (9.93)	44.87 (8.93)	44.74 (9.32)	41.00 (9.88)	0.62	0.63	0.99

Note: CD, Multiple Change Detection Task; SQ, colored square stimuli; K, Working Memory capacity estimate; A, Attention parameter; G, Guess parameter; CL, Change Localization Task; T, T score from the MATRICS Consensus Cognitive Battery; LN Span, Letter-Number Span Test. Abbreviations are explained in the first footnote to Table 1.

Table 3. Oblimin Rotated Factor Pattern Matrices

Factor	3-Factor Solution			2-Factor Solution	
	I	II	III	I	II
	Visual WM	Complex span	Attention	Visual WM	Complex span
CD SH K	**0.74**	0.10	−0.03	**0.75**	0.06
CD SQ K	**0.84**	−0.08	−0.01	**0.83**	−0.10
CL SH K	**0.63**	0.07	0.07	**0.68**	0.05
CL SQ K	**0.58**	0.10	0.18	**0.68**	0.09
Run Span	−0.06	**0.97**	0.04	−0.05	**0.97**
ARun Span	0.04	**0.80**	−0.03	0.02	**0.80**
LNS Total	0.19	**0.66**	−0.04	0.18	**0.65**
CD SH A	-0.04	0.01	**1.01**		
CD SQ A	0.22	−0.03	**0.64**		
Factor Correlations		II	III		II
	I	0.46	0.45	I	0.49
	II		0.30		

Note: SH, shape stimuli; Run Span, total correct from Running Span; ARun Span, adaptive version of Running Span; LNS Total, Total correct from the Letter Number Span Task. Abbreviations are explained in the first footnote to Table 1. Variables in bold all load P <.0001.

Table 4. WM and MCCB Correlations with Functional Capacity and Outcome

Correlations			
	UPSA B Total	SLOF Self	SLOF Inf
Running Span	0.27**	0.05	0.12
CD SQ K	0.27**	0.20*	0.34**
CD SQ A	0.27**	0.14	0.15
CL SQ K	0.39**	0.16*	0.31**
LNS T	0.38**	0.03	0.15
HVLT T	0.41**	0.03	0.15
BACS symbol Coding T	0.36**	0.11	0.20*

Note: UPSA B N's = 155–157, SLOF Self N's = 155–157, SLOF Informant N's = 92–93; BACS, Brief Assessment of Cognition in Schizophrenia. Abbreviations are explained in the first footnote to Table 1.
**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

However, the magnitude of the other correlations did not differ significantly. When we examined the relationship of the factor scores described earlier, we found that the WM capacity factor correlated with all 3 functional measures, whereas the complex span and attention factors only correlated with the UPSA-B. We obtain essentially the same results if we simply z score average the measures for each factor. A full table of symptom correlations is shown in supplementary table S3. We observed significant but modest correlations between the CAINS motivation and pleasure score and all the WM capacity measures (r's −.21 to −.26, all P < .01) but not the Attention parameter. There was one significant (uncorrected) correlation with BPRS positive symptoms (Running Span: r = −0.23, P =.01), a single correlation with BPRS Mania (Attention parameter: r = −.17, P = .029), and no significant correlations with BPRS disorganization factor ratings (all P's > −.14).

Discussion

These data provide strong evidence that WM deficits are shared across all 3 psychotic disorders evaluated in this study. These results are noteworthy because BP often perform substantially better than SZ on many clinical neuropsychological tests,[8] suggesting that WM capacity may be an unusually sensitive indicator of cognitive dysfunction in BP disorder with psychosis. The SZ group had somewhat more severe deficits than the BP group on more complex span tasks, scoring significantly more poorly than BP on Running Span, though the BP group also demonstrated impairments relative to HC. Thus, differences across diagnosis were subtle whereas the differences for all 3 diagnoses vs controls were generally robust.

The factor analysis results suggest that the simple visual capacity measures define one construct, whereas the more complex verbal span measures define a second, and the Attention parameters defines a third, findings similar to those of Shipstead *et al*.[47] In both Running Span and Letter-Number Span, individuals must act on stored information, adding complexity. In contrast, the Multiple Change Detection/Localization task measures how many items a person can successfully encode and store, changes appear to be detected automatically, and no further operations or delays are involved.[48] The additional complexity involved on Running Span and Letter-Number Span likely explains the larger effect sizes in between group contrasts than were seen on the Change Localization and

Multiple Change Detection tasks. A measure that combines multiple impaired cognitive operations often shows more robust group differences than a purer measure, leading to a trade-off between group discriminability and precision in construct assessment.

Although all 3 patient groups were impaired compared with HC on the Change Localization tasks, the level of impairment seen when comparing HC and SZ ($d = 0.52$) was substantially smaller than the d of 1.11 previously reported on a similar task by Johnson *et al.*[18] (notably, the present Change Localization K score is substantially lower in both groups than in the previous study). The main procedural difference is that this study used 5-item arrays whereas the previous study used 4-item arrays. This may have led to differences in strategy, reducing estimated capacity and reducing between-group differences. Although 5-item arrays might appear to be advantageous in avoiding ceiling effects, we have not found ceiling effects using 4-item arrays, and the smaller array size may be superior in terms of task engagement.

The Multiple Change Detection task yielded the most surprising findings. First, although estimated WM capacity was significantly impaired in the patient group as a whole relative to HC, only the BP group was significantly impaired relative to HC, with a marginally significant deficit in SZ relative to HC, and no significant impairment in SZAFF. (As seen in supplementary materials, there was not an overall HC vs combined diagnostic group difference with the colored shape stimuli). There were no significant differences among the 3 patient groups, so the data are not suggesting that impaired WM capacity is specific to BP. HC differed from the patient groups most robustly on the modeled Attention parameter. The findings on K and A parameters are provocative in suggesting that estimates of WM capacity reduction in patients reflect the impact of both attentional lapses and reduced capacity.

Together, the present data suggest that psychosis patients have reduced WM capacity in addition to poorer attentional engagement. Both of these impairments likely contribute to the broad cognitive deficit observed in psychotic disorders. This mathematical modeling framework presents an important challenge and benefit to the field: we need to develop experimental methods and analytic approaches that facilitate measurement of specific cognitive functions, free from the influence of attention lapses that are more likely to occur in clinical groups than in HCs.

Importantly, the capacity measures from Change Localization and Multiple Change Detection showed significant correlations with functional capacity and outcome. Although the other WM measures showed correlations with 1 or 2 of these measures, none showed the same consistent pattern as the Change Detection/Localization measures. Further, only the visual WM factor score was associated with all 3 measures of function, as the complex span and attention measures were only associated with functional capacity. These links to function may make the visual Change Detection measures more clinically useful tools than complex span measures. Our study has limitations including that all of our patients were medicated (although we observed no correlations with antipsychotic dose, an analysis confounded by non-random assignment). Further, our study lacked power to detect small between patient group differences. In addition, without evidence from experimental manipulations of encoding, maintenance, and retrieval processes, we cannot be certain that the pathway to observed overall impairments seen here is identical across disorders.

Conclusion

In summary, WM impairment is shared across psychotic disorders, with the extent of impairment correlated with negative symptom severity, functional capacity, and level of community function. In addition, the modeling results suggest that attention lapsing will be important to consider in future clinical cognitive research.

Supplementary Material

Supplementary data are available at *Schizophrenia Bulletin* online.

Funding

Funding for this study was provided by National Institutes of Mental Health (NIMH) (ROI1s MH084840 (to D.M.B.), MH084826 (to C.S.C.), MH084821 (to J.M.G.), MH084861 (to A.W.M.) and MH084828 (to S.M.S.).

Acknowledgments

Gold has received grants from the National Institutes of Health (NIH), receives royalty payments from Brief Assessment of Cognition in Schizophrenia, and has acted as a consultant to Amgen, AstraZeneca, GlaxoSmithKline, Hoffman LaRoche, Merck, Pfizer, and Solvay. Barch has received grants from the Brain and Behavior Research Foundation and the NIH, and is a consultant for Pfizer, Amgen, Upsher-Smith, and Takeda on studies related to the treatment of negative symptoms in schizophrenia. Carter has received research grants from the NIH, the Brain and Behavior Foundation, the Burroughs Wellcome foundation, GlaxoSmithKline, and the Robert Wood Johnson Foundation and has been an external consultant for Lilly, Merck, Pfizer, Roche, and Servier. Luck has received research grants from the NIH, the National Science Foundation (NSF) and the Harry Frank Guggenheim Foundation. MacDonald has received research grants from the NIH and the Brain and Behavior

Research Foundation. Ragland has received research grants from the NIH, the Brain and Behavior Research Foundation, the EJLB Foundation, and the Robert Wood Johnson Foundation. Silverstein has received research grants from the NIMH, the Brain ad Behavior Research Foundation, the van Ameringen Foundation, the Jacob and Valeria Langaloth Foundation, the New England Research Institutes, the New York State Office of Mental Health, the New Jersey Division of Mental Health and Addiction Services, Janssen Pharmaceutica, AstraZeneca, and Pfizer. Strauss and Feuerstahler have no conflicts to report.

References

1. Baddeley A. Working memory: theories, models, and controversies. *Annu Rev Psychol.* 2012;63:1–29.
2. Luck SJ, Vogel EK. Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn Sci.* 2013;17:391–400.
3. Eriksson J, Vogel EK, Lansner A, Bergström F, Nyberg L. Neurocognitive architecture of working memory. *Neuron.* 2015;88:33–46.
4. Park S, Gooding DC. Working memory impairment as an endophenotypic marker of a schizophrenia diathesis. *Schizophr Res Cogn.* 2014;1:127–136.
5. Unsworth N, Redick TS, Heitz RP, Broadway JM, Engle RW. Complex working memory span tasks and higher-order cognition: a latent-variable analysis of the relationship between processing and storage. *Memory.* 2009;17:635–654.
6. Engle RW, Tuholski SW, Laughlin JE, Conway AR. Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J Exp Psychol Gen.* 1999;128:309–331.
7. Schaefer J, Giangrande E, Weinberger DR, Dickinson D. The global cognitive impairment in schizophrenia: consistent over decades and around the world. *Schizophr Res.* 2013;150:42–50.
8. Hill SK, Reilly JL, Keefe RS, et al. Neuropsychological impairments in schizophrenia and psychotic bipolar disorder: findings from the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) study. *Am J Psychiatry.* 2013;170:1275–1284.
9. Bora E. Neurocognitive features in clinical subgroups of bipolar disorder: a meta-analysis. *J Affect Disord.* 2018;229:125–134.
10. Frydecka D, Eissa AM, Hewedi DH, et al. Impairments of working memory in schizophrenia and bipolar disorder: the effect of history of psychotic symptoms and different aspects of cognitive task demands. *Front Behav Neurosci.* 2014;8:416.
11. Jiménez-López E, Aparicio AI, Sánchez-Morla EM, Rodriguez-Jimenez R, Vieta E, Santos JL. Neurocognition in patients with psychotic and non-psychotic bipolar I disorder. A comparative study with individuals with schizophrenia. *J Affect Disord.* 2017;222:169–176.
12. Nenadic I, Langbein K, Dietzek M, Forberg A, Smesny S, Sauer H. Cognitive function in euthymic bipolar disorder (BP I) patients with a history of psychotic symptoms vs. schizophrenia. *Psychiatry Res.* 2015;230:65–69.
13. Sperry SH, O'Connor LK, Öngür D, Cohen BM, Keshavan MS, Lewandowski KE. Measuring cognition in bipolar disorder with psychosis using the MATRICS consensus cognitive battery. *J Int Neuropsychol Soc.* 2015;21:468–472.
14. Glahn DC, Bearden CE, Cakir S, et al. Differential working memory impairment in bipolar disorder and schizophrenia: effects of lifetime history of psychosis. *Bipolar Disord.* 2006;8:117–123.
15. Pirkola T, Tuulio-Henriksson A, Glahn D, et al. Spatial working memory function in twins with schizophrenia and bipolar disorder. *Biol Psychiatry.* 2005;58:930–936.
16. Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions. *Nature.* 1997;390:279–281.
17. Gold JM, Wilk CM, McMahon RP, Buchanan RW, Luck SJ. Working memory for visual features and conjunctions in schizophrenia. *J Abnorm Psychol.* 2003;112:61–71.
18. Johnson MK, McMahon RP, Robinson BM, et al. The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. *Neuropsychology.* 2013;27:220–229.
19. Gold JM, Robinson B, Leonard CJ, et al. Selective attention, working memory, and executive function as potential independent sources of cognitive dysfunction in schizophrenia. *Schizophr Bull.* Epub ahead of print. doi:10.1093/schbul/sbx155
20. Gold JM, Fuller RL, Robinson BM, McMahon RP, Braun EL, Luck SJ. Intact attentional control of working memory encoding in schizophrenia. *J Abnorm Psychol.* 2006;115:658–673.
21. Erickson M, Hahn B, Leonard C, Robinson B, Luck S, Gold J. Enhanced vulnerability to distraction does not account for working memory capacity reduction in people with schizophrenia. *Schizophr Res Cogn.* 2014;1:149–154.
22. Erickson MA, Hahn B, Leonard CJ, et al. Impaired working memory capacity is not caused by failures of selective attention in schizophrenia. *Schizophr Bull.* 2015;41:366–373.
23. Hahn B, Harvey AN, Gold JM, Ross TJ, Stein EA. Load-dependent hyperdeactivation of the default mode network in people with schizophrenia. *Schizophr Res.* 2017;185:190–196.
24. Minzenberg MJ, Laird AR, Thelen S, Carter CS, Glahn DC. Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. *Arch Gen Psychiatry.* 2009;66:811–822.
25. Barch DM, Carter CS, Dakin SC, et al. The clinical translation of a measure of gain control: the contrast-contrast effect task. *Schizophr Bull.* 2012;38:135–143.
26. Feuerstahler, LM, Luck SJ, MacDonald III A, Waller NG. A note on the identification of change detection task models to measure storage capacity and attention in visual working memory. *Behav Res Methods*. 2018. In press. doi:10.3758/s13428-018-1082-z
27. Broadway JM, Engle RW. Validating running memory span: measurement of working memory capacity and links with fluid intelligence. *Behav Res Methods.* 2010;42:563–570.
28. Nuechterlein KH, Green MF, Kern RS, et al. The MATRICS consensus cognitive battery, part 1: test selection, reliability, and validity. *Am J Psychiatry.* 2008;165:203–213.
29. Kern RS, Gold JM, Dickinson D, et al. The MCCB impairment profile for schizophrenia outpatients: results from the MATRICS psychometric and standardization study. *Schizophr Res.* 2011;126:124–131.
30. August SM, Kiwanuka JN, McMahon RP, Gold JM. The MATRICS Consensus Cognitive Battery (MCCB): clinical and cognitive correlates. *Schizophr Res.* 2012;134:76–82.

31. American-Psychiatric-Association (Ed.). *Diagnostic and Statistical Manual of Mental Disorders*. 4th Text Revision ed. Washington, DC: American Psychiatric Association; 2000.
32. Barch DM, Carter CS, Gold JM, et al. Explicit and implicit reinforcement learning across the psychosis spectrum. *J Abnorm Psychol.* 2017;126:694–711.
33. Wechsler D. *Wechsler Test of Adult Reading*. San Antonia, TX: The Psychological Corporation; 2001.
34. Gardner DM, Murphy AL, O'Donnell H, Centorrino F, Baldessarini RJ. International consensus study of antipsychotic dosing. *Am J Psychiatry.* 2010;167:686–693.
35. First MB, Spitzer RL, Miriam G, Williams JBW. *Structured clinical interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition (SCID-I/NP)*. New York: Biometrics Research, New York State Psychiatric Institute; 2002.
36. Overall JE, Gorham DR. The brief psychiatric rating scale. *Psychol Rep.* 1962;10:799.
37. Ventura J, Green MF, Shaner A, Liberman RP. Training and quality assurance on the Brief Psychiatric Rating Scale: the "drift busters". *Int J Meth Psychiatry Res.* 1993;3:221–226.
38. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry.* 1978;133:429–435.
39. Berk M, Malhi GS, Cahill C, et al. The Bipolar Depression Rating Scale (BDRS): its development, validation and utility. *Bipolar Disord.* 2007;9:571–579.
40. Kring AM, Gur RE, Blanchard JJ, Horan WP, Reise SP. The Clinical Assessment Interview for Negative Symptoms (CAINS): final development and validation. *Am J Psychiatry.* 2013;170:165–172.
41. Schneider LC, Struening EL. SLOF: a behavioral rating scale for assessing the mentally ill. *Soc Work Res Abstr.* 1983;19:9–21.
42. Harvey PD, Velligan DI, Bellack AS. Performance-based measures of functional skills: usefulness in clinical treatment studies. *Schizophr Bull.* 2007;33:1138–1148.
43. Patterson TL, Goldman S, McKibbin CL, Hughs T, Jeste DV. UCSD performance-based skills assessment: development of a new measure of everyday functioning for severely mentally ill adults. *Schizophr Bull.* 2001;27:235–245.
44. Twamley EW, Doshi RR, Nayak GV, et al. Generalized cognitive impairments, ability to perform everyday tasks, and level of independence in community living situations of older patients with psychosis. *Am J Psychiatry.* 2002;159:2013–2020.
45. Kyllingsbaek S, Bundesen C. Changing change detection: improving the reliability of measures of visual short-term memory capacity. *Psychon Bull Rev.* 2009;16:1000–1010.
46. Meng X, Rosenthal R, Rubin DB. Comparing correlated correlation coefficients. *Psychol Bull.* 1992;111:172–175.
47. Shipstead Z, Harrison TL, Engle RW. Working memory capacity and the scope and control of attention. *Atten Percept Psychophys.* 2015;77:1863–1880.
48. Hyun JS, Woodman GF, Vogel EK, Hollingworth A, Luck SJ. The comparison of visual working memory representations with perceptual inputs. *J Exp Psychol Hum Percept Perform.* 2009;35:1140–1160.