The Relationship Between Working Memory Capacity and Broad Measures of Cognitive Ability in Healthy Adults and People With Schizophrenia

Melissa K. Johnson
Baltimore VA Mental Health Care System, Baltimore, Maryland

Robert P. McMahon, Benjamin M. Robinson, Alexander N. Harvey, and Britta Hahn
University of Maryland School of Medicine

Carly J. Leonard and Steven J. Luck
University of California, Davis

James M. Gold
University of Maryland School of Medicine

Objective: Working memory (WM) capacity, typically measured with cognitively complex span tasks, is correlated with higher order cognitive abilities in healthy adults. The goals of this study were to determine whether a more focused measure of visual WM storage capacity would show similar higher order ability correlations in healthy adults and in people with schizophrenia (PSZ), thereby demonstrating the importance of simple storage capacity; determine whether the illness alters the pattern of correlations across cognitive domains; and evaluate whether between-groups differences in WM capacity could account for the generalized cognitive impairment in PSZ. Method: Ninety-nine PSZ and 77 healthy controls (HCs) completed a visual WM change-localization task, the Wechsler Abbreviated Scale of Intelligence (WASI), and the MATRICS Consensus Cognitive Battery (MCCB). Results: PSZ performed more poorly than HCs on all cognitive measures. The between-groups effect size for WM capacity was large ($d = 1.11$). WM robustly correlated with WASI and MCCB performance, with no significant differences in the magnitude or pattern of correlations across groups. When the groups were pooled, WM capacity correlated at $r = .68$ with MCCB composite score and at $r = .56$ with WASI estimated Full Scale IQ. WM capacity accounted for approximately 40% of the between-groups variance across the WASI and MCCB. Conclusions: A simple measure of WM storage capacity is robustly associated with the higher order cognitive abilities assessed by the WASI and MCCB in HCs and PSZ. WM capacity reduction may be a critical determinant of the general cognitive impairment in PSZ.

Keywords: working memory, schizophrenia, intelligence

Supplemental materials: http://dx.doi.org/10.1037/a0032060.supp

There is a large body of evidence that individual differences in working memory (WM) storage capacity are predictive of performance on many broad measures of complex cognition in healthy people, including scholastic aptitude, fluid intelligence, and executive function (Cowan et al., 2005; Fukuda, Vogel, Mayr, & Awh, 2010; Miyake, Friedman, Rettinger, Priti, & Hegarty, 2001). Much of this evidence comes from correlations between criterion cognitive measures and complex WM span measures that assess the ability of participants to maintain information in WM while performing distracting cognitive operations (Daneman & Carpenter, 1980; Kail, 2001; Kane, Hambrick, & Conway, 2005). However, complex span measures involve multiple cognitive processes beyond simple WM capacity, including task switching and suppression of proactive interference, and it is not clear which of the cognitive demands posed by complex span measures accounts for the relationships with higher order cognitive performance.

Despite the importance of task switching and inhibitory processes, some evidence suggests that the WM storage capacity per se may account for some of the variance in criterion measures. For example, studies have shown associations between higher order abilities and simple span measures, which depend to a much greater degree on pure WM storage capacity (Colom, Shih, Flores-Mendoza, & Quiroga, 2006; Miyake et al., 2001; Unsworth & Engle, 2007). These associations were of similar magnitude as with the more complex span tasks described above.

For decades, researchers have assumed that WM span is correlated with broader measures of cognition because complex cognitive tasks require the active representation of multiple pieces of
Working Memory and Broad Cognitive Ability

Information in a short-term storage buffer (Anderson, 1993; Baddeley, 1986; Meyer & Kieras, 1997). However, traditional WM span tasks do not appear to provide a pure measure of this ability because they are heavily influenced by executive control processes (Cowan, 2001; Saults & Cowan, 2007). Visual change detection tasks appear to provide a purer measure of the amount of information present in a short-term storage system. In these tasks (see Figure 1), observers are shown a sample array containing a few simple stimuli (e.g., four colored squares) for a short period (e.g., 100 ms). After a brief delay (e.g., 1,000 ms), a test array appears that is either identical to the sample array or differs in some way (e.g., one of the items might have changed colors). Participants then report whether or not they detected a change. A change-localization task can also be used in which a change occurs on every trial and the participants report the location of the change (Gold et al., 2006; Kyllingsbæk & Bundesen, 2009). A simple equation can be used to estimate a participant’s storage capacity from performance in these tasks (Cowan et al., 2005; Pashler, 1988). Healthy college students typically have a capacity of three to five items, although substantial individual differences have been observed even in this population (Vogel & Awh, 2008).

Several pieces of evidence indicate that the measure of storage capacity obtained from this task provides a relatively pure index of the amount of information that can be actively maintained in a short-term buffer. First, both functional MRI and event-related potential studies have shown that the maintenance of information in memory in this task is accompanied by sustained neural activity during the delay interval, indicating that the information is actively maintained; as the number of items to be stored increases, the sustained neural activity increases, reaching an asymptote at the individual observer’s behaviorally measured capacity level (Todd & Marois, 2004; Vogel & Machizawa, 2004). Second, performance is strongly related to the visual properties of the stimuli (Alvarez & Cavanagh, 2004; Zhang & Luck, 2008) but is not strongly influenced by links to long-term memory (Olson & Jiang, 2004; Pashler, 1988). Third, whereas other span measures are strongly influenced by proactive interference from previous trials, no evidence of proactive interference is observed in the change-detection task when typical parameters are used (Lin & Luck, 2012). Fourth, the memory representations that underlie change detection can be created in as little as 50 ms (Vogel, Woodman, & Luck, 2006), which is presumably too fast for the creation of long-term memory representations. Fifth, the creation of these representations immediately impacts other ongoing cognitive processes (Carlisle, Arita, Pardo, & Woodman, 2011; Hollingworth, Richard, & Luck, 2008), indicating that they constitute a working memory.

Like the measures of capacity provided by other span tasks, the relatively pure measure of WM storage capacity that can be obtained in change-detection tasks is strongly correlated with broader measures of cognitive ability (Cowan et al., 2005; Fukuda et al., 2010; Gold et al., 2010). Demonstrating such effects with simple measures of visual change detection is particularly attractive from an interpretive point of view because individual differences in WM capacity derived from this type of task have been linked to a relatively circumscribed neural system (Edin et al., 2009; McNab & Klingberg, 2008; Todd & Marois, 2005), thereby implicating the functional integrity of this system to much broader aspects of cognition.

In addition to varying across healthy individuals, WM capacity is also impaired in individuals with psychiatric disorders. In particular, dozens of studies have shown that people with schizophrenia (PSZ) reliably show impairments in WM performance (see meta-analyses by Lee & Park, 2005; Piskulic, Olver, Norman, & Maruff, 2007). Many other studies have shown substantial impairments in broader measures of cognitive function in PSZ, and these impairments are among the best predictors of functional outcome in the disease (Green, 1996; Heinrichs & Zakzanis, 1998). However, relatively little research has attempted to link the WM impairment to the broader general cognitive deficit of schizophrenia. Some research on PSZ has demonstrated that complex span measures such as letter–number span and listening span are related to complex problem-solving measures (such as the Wisconsin Card Sorting Task and Tower of Hanoi), as well as verbal free recall (Chan, Wang, Cao, & Chen, 2010; Gold, Carpenter, Randolph, Goldberg, & Weinberger, 1997; Perry et al., 2001; Stone, Gabrieli, Stebbins, & Sullivan, 1998), but this research did not examine broad composite measures of cognitive ability. Moreover, it is not known whether the observed correlations reflect the storage capacity component or the executive component of the complex span tasks. Thus, despite overwhelming evidence that both WM capacity and broad cognitive abilities are impaired in PSZ, there has

Figure 1. Top panel: Timeline of the change-localization task. Bottom panel: distribution of K scores in people with schizophrenia (PSZ) and healthy controls (HC). The value under each bar reflects the lower limit of the corresponding K bin, with a bin width of 0.5.
been little research asking whether and how the WM deficit can explain the broad cognitive impairment.

Our previous research (Gold et al., 2010; Leonard et al., 2012) has provided some preliminary evidence that visual WM storage capacity is correlated with IQ and overall performance on the MATRICS Consensus Cognitive Battery (MCCB; Nuechterlein & Green, 2006), a battery of cognitive tasks across a wide range of domains, including processing speed, attention, learning, and problem solving. Large reductions in WM storage capacity were found in PSZ, and strong correlations were observed between WM storage capacity and the MCCB, but this research was limited to relatively small sample sizes.

In the present study, we examined reductions in visual WM capacity and their correlations with MCCB and Wechsler Abbreviated Scale of Intelligence (WASI) IQ in a more comprehensive fashion in a much larger group of participants. We sought to address four key issues. First, we asked how severely visual WM capacity is reduced in PSZ. The specific issue of capacity reduction was not addressed in the previous meta-analyses (Lee & Park, 2005; Piskulic et al., 2007), which integrated results from studies using many different procedures and performance metrics that may be measuring different underlying abilities and that have varying levels of reliability, whereas we used a single task that is known to isolate visual WM storage capacity and has high reliability. Second, we asked whether the highly focused measure of storage capacity that is obtained from the change-detection/localization task is correlated with measures of broad cognitive function. Previous research in healthy populations has shown that complex WM span is moderately correlated with broad measures of cognitive function (Unsworth & Engle, 2007), but much of this may reflect the executive components of the complex span tasks (Lustig, May, & Hasher, 2001). The relatively pure measure of storage capacity provided by the change-detection/localization task is correlated with narrow measures of global fluid intelligence in college students (Fukuda et al., 2010), but no studies have examined correlations with broader measures of cognitive function in more diverse samples of adults.

A third key question was whether the relationship between visual WM capacity and broader measures of cognition is altered by schizophrenia. On the one hand, this relationship might be quite different in PSZ because the mechanisms that produce reduced capacity in PSZ appear to be different from those that cause variation in capacity among healthy individuals. Specifically, individual differences among healthy young adults mainly reflect differences in the ability to use selective attention to control the contents of WM (McNab & Klingberg, 2008; Vogel, McCollough, & Machizawa, 2005), but the pattern of behavioral performance and brain activity in PSZ is not like the pattern observed in low-capacity healthy individuals (Gold et al., 2006; Leonard et al., 2012). Consequently, correlations between WM capacity and broader measures of cognitive ability could be either amplified or attenuated in PSZ. The few studies that have examined this issue—using very different methods—have produced contradictory findings (Gold et al., 1997; Perry et al., 2001; Silver, Feldman, Bilker, & Gur, 2003). However, if WM capacity is associated with the ability to perform a broad array of cognitive tasks, then reduced capacity should be correlated with impaired cognitive performance no matter what originally caused the reduced capacity. This is a key question for the development of treatments for impaired cognition in schizophrenia: If reduced WM capacity in PSZ is strongly predictive of impaired cognitive performance, then it is reasonable to develop treatments aimed at increasing WM capacity.

A fourth and related question addressed by the present study was whether deficits in WM capacity among PSZ can account for a substantial proportion of their impairment in overall cognitive ability. If so, this provides additional incentive to develop treatments designed to increase WM storage capacity. Of course, our correlational methods cannot readily determine whether deficits in WM capacity actually cause the impairments in broader cognitive ability. Nonetheless, if a substantial proportion of the broad cognitive impairment is statistically associated with the degree of WM reduction, then this will be an important first step toward understanding the causes of impaired cognition in schizophrenia, especially given that our measure of WM capacity is closely linked to a circumscribed and well-studied set of neural mechanisms.

Method

Participants

Ninety-nine people who met Diagnostic and Statistical Manual of Mental Disorders (4th ed., text rev.; American Psychiatric Association, 2000) criteria for schizophrenia (42 paranoid, 27 undifferentiated, three disorganized, nine residual, one catatonic) or schizoaffective disorder (n = 17) and 77 matched healthy controls (HCs) participated in this study. Diagnosis was established using a consensus best estimate approach combining the results of a Structured Clinical Interview for DSM–IV (First, Spitzer, Gibbon, & Williams, 1997) with medical records and collateral informants when available. Demographic information is summarized in the Table 1. There was a significant difference in years of education between groups, t(174) = 5.85, p < .001; which presumably reflects the effects of schizophrenia on educational attainment rather than a premorbid demographic difference. Consistent with this, the groups did not differ in parental education, t(168) = 0.371, p = .711. They also did not differ in gender distribution, χ²(1, N = 176) = 0.002, p = .969; race, χ²(4, N = 176) = 5.39, p = .25; ethnicity, χ²(2, N = 176) = 1.56, p = .45; or age, t(174) = 0.08, p = .94.

The PSZ were clinically stable outpatients. At the time of testing, PSZ were mildly/moderately symptomatic as measured by the Scale for the Assessment of Negative Symptoms and the Brief Psychiatric Rating Scale (see Table 1; Andreasen, 1983; Overall & Gorham, 1962). All PSZ were receiving antipsychotic medication; 17 were treated with a first-generation antipsychotic; four with a combination of a first-generation and a second-generation antipsychotic; 29 with a single, second-generation antipsychotic other than clozapine; 33 received clozapine; 11 received clozapine with an additional antipsychotic; and five received two different second-generation antipsychotics other than clozapine. Additional psychotropic medication use was as follows: 17 received a mood stabilizer, 17 received an anticholinergic, 47 received antidepressants, and 29 antianxiety medication. PSZ were taking stable medications for a minimum of 4 weeks prior to testing. Control participants were recruited from the community, primarily through random digit dialing in local zip codes and by word of mouth among the friends of participants recruited in this fashion, and had no current Axis I or II diagnoses as established by the Structured...
soning/problem solving, and social cognition, which are combined fixation cross was shown, participants were presented with a test angle, for 100 ms. After a 900-ms delay during which only the 

The task was programmed in E-Prime, and stimuli were presented on a CRT monitor at a nominal viewing distance of 70 cm. As 

WM Change-Localization Stimuli and Task

All participants completed 60 trials of a change-localization task, an experimental paradigm that measures visual WM capacity. The task was programmed in E-Prime, and stimuli were presented on a CRT monitor at a nominal viewing distance of 70 cm. As illustrated in Figure 1, participants were presented with a sample array of four colored squares, each measuring 0.7° × 0.7° of visual angle, for 100 ms. After a 900-ms delay during which only the fixation cross was shown, participants were presented with a test array that was identical to the sample array except that one color had changed to a value that had not been present in the sample array. The colors were selected at random without replacement from a set of six colors with the following red, green, blue values: red (255, 0, 0), green (0, 255, 0), yellow (255, 255, 0), magenta (255, 0, 255), cyan (0, 255, 255), and bright blue (80, 60, 255).

From trial to trial, the colored squares were randomly placed around an imaginary circle with a radius of 3°, with the constraints that one item appeared in each of the four screen quadrants, and each item was separated by a minimum of 30° on the circle from the next item. Participants used a mouse to click on the location of the item in the test array that they believed had changed color from the sample array. They were encouraged to respond accurately, with no speed pressure, and a response was required on every trial. The test array was visible until the response was made, and trials were separated by a 2,000-ms intertrial interval. A total of 60 trials was tested in each subject, which typically required approximately 8–10 min of testing time. The E-Prime script is available at http://mindbrain.ucdavis.edu/labs/luck-lab/change-localization.

It should be noted that this is an extremely simple task: Participants view two arrays of highly discriminable items, separated by a brief delay, and make an unspeeded response to indicate the location of the one item that differed between arrays. There is no time pressure or task switching. Moreover, it is very easy for subjects to understand the instructions, and no background knowledge (e.g., vocabulary or math facts) is needed to perform the task. This makes it a purer measure of WM storage capacity than complex span tasks. In addition, the fact that chance is 25% rather than 50% for this task minimizes guessing effects and increases measurement reliability (Kyllingsbæk & Bundesen, 2009).

This task yields highly reliable results. We have retested a total of 45 participants (14 HC and 31 PSZ) twice with a mean retest interval of 1.5 years (SD = 1.1). The mean K was 2.68 (SD = 0.56) at Time 1 and was 2.63 (SD = .60) at Time 2. The Pearson correlation between testing occasions was .77, p < .001, and the intraclass correlation coefficient was .76.

Table 1

Demographic and Psychiatric Characteristics of Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Schizophrenia (n = 99)</th>
<th>Healthy controls (n = 77)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (SD) age (years)</td>
<td>39.4 (10.5)</td>
<td>39.6 (10.4)</td>
</tr>
<tr>
<td>Age range (years)</td>
<td>18–58</td>
<td>18–54</td>
</tr>
<tr>
<td>Mean (SD) participant education (years)</td>
<td>12.79 (2.3)</td>
<td>14.71 (2.0)</td>
</tr>
<tr>
<td>Mean (SD) parental education (years)*</td>
<td>13.5 (3.0)</td>
<td>13.6 (2.5)</td>
</tr>
<tr>
<td>Gender, % male</td>
<td>61</td>
<td>61</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Native American</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Asian</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>African American</td>
<td>36</td>
<td>32</td>
</tr>
<tr>
<td>Mixed</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Caucasian</td>
<td>54</td>
<td>44</td>
</tr>
<tr>
<td>Mean total symptom scores</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPRS anxiety/depression</td>
<td>7.72 (2.65)</td>
<td></td>
</tr>
<tr>
<td>BPRS angeria</td>
<td>6.88 (2.46)</td>
<td></td>
</tr>
<tr>
<td>BPRS negative symptoms</td>
<td>5.8 (2.39)</td>
<td></td>
</tr>
<tr>
<td>BPRS psychosis</td>
<td>9.26 (4.12)</td>
<td></td>
</tr>
<tr>
<td>BPRS activation</td>
<td>4.14 (1.42)</td>
<td></td>
</tr>
<tr>
<td>BPRS hostile/suspiciousness</td>
<td>5.36 (2.26)</td>
<td></td>
</tr>
<tr>
<td>BPRS total</td>
<td>32.57 (6.97)</td>
<td></td>
</tr>
<tr>
<td>SANS affective flattening/blunting</td>
<td>7.75 (5.71)</td>
<td></td>
</tr>
<tr>
<td>SANS alogia</td>
<td>1.88 (1.83)</td>
<td></td>
</tr>
<tr>
<td>SANS avolition/apathy</td>
<td>7.95 (4.53)</td>
<td></td>
</tr>
<tr>
<td>SANS anhedonia/assocility</td>
<td>8.02 (3.97)</td>
<td></td>
</tr>
<tr>
<td>SANS total</td>
<td>25.58 (11.51)</td>
<td></td>
</tr>
</tbody>
</table>

Note. BPRS = Brief Psychiatric Rating Scale; SANS = Schedule for the Assessment of Negative Symptoms.

* Parental education data were unavailable for six participants.

Clinical Interview for DSM–IV Axis I Disorders (First et al., 1997), 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 Measures

All participants completed the WASI (Wechsler, 1999) and the MCCB (Nuechterlein & Green, 2006). The MCCB yields measures of seven cognitive domains: processing speed, working memory, verbal learning, visual learning, attention/vigilance, reasoning/problem solving, and social cognition, which are combined in the calculation of an overall composite T score.

WM Change-Localization Stimuli and Task

All participants completed 60 trials of a change-localization task, an experimental paradigm that measures visual WM capacity. The task was programmed in E-Prime, and stimuli were presented on a CRT monitor at a nominal viewing distance of 70 cm. As illustrated in Figure 1, participants were presented with a sample array of four colored squares, each measuring 0.7° × 0.7° of visual angle, for 100 ms. After a 900-ms delay during which only the fixation cross was shown, participants were presented with a test array that was identical to the sample array except that one color had changed to a value that had not been present in the sample array. The colors were selected at random without replacement from a set of six colors with the following red, green, blue values: red (255, 0, 0), green (0, 255, 0), yellow (255, 255, 0), magenta (255, 0, 255), cyan (0, 255, 255), and bright blue (80, 60, 255).

From trial to trial, the colored squares were randomly placed around an imaginary circle with a radius of 3°, with the constraints that one item appeared in each of the four screen quadrants, and each item was separated by a minimum of 30° on the circle from the next item. Participants used a mouse to click on the location of the item in the test array that they believed had changed color from the sample array. They were encouraged to respond accurately, with no speed pressure, and a response was required on every trial. The test array was visible until the response was made, and trials were separated by a 2,000-ms intertrial interval. A total of 60 trials was tested in each subject, which typically required approximately 8–10 min of testing time. The E-Prime script is available at http://mindbrain.ucdavis.edu/labs/luck-lab/change-localization.

It should be noted that this is an extremely simple task: Participants view two arrays of highly discriminable items, separated by a brief delay, and make an unspeeded response to indicate the location of the one item that differed between arrays. There is no time pressure or task switching. Moreover, it is very easy for subjects to understand the instructions, and no background knowledge (e.g., vocabulary or math facts) is needed to perform the task. This makes it a purer measure of WM storage capacity than complex span tasks. In addition, the fact that chance is 25% rather than 50% for this task minimizes guessing effects and increases measurement reliability (Kyllingsbæk & Bundesen, 2009).

This task yields highly reliable results. We have retested a total of 45 participants (14 HC and 31 PSZ) twice with a mean retest interval of 1.5 years (SD = 1.1). The mean K was 2.68 (SD = 0.56) at Time 1 and was 2.63 (SD = .60) at Time 2. The Pearson correlation between testing occasions was .77, p < .001, and the intraclass correlation coefficient was .76.

Statistical Analysis

WM storage capacity was quantified using a variant of the Pashler/Cowan K equation, where K represented how many items in the memory array; see Gold et al., 2006; Kyllingsbæk & Bundesen, 2009).

Effect sizes (Cohen’s d) for group differences on K, the WASI, and the MCCB were calculated by d = (difference of group means)/pooled SD, where the pooled SD was calculated as $SD_{pooled} = \sqrt{(\frac{n_1-1}{n_1}SD_1^2 + \frac{n_2-1}{n_2}SD_2^2)}(\frac{n_1+n_2-2}{n_1+n_2})^{1/2}$. Partial Spearman correlations between WM capacity (K) and the MCCB and WASI domain scores were computed, adjusted for age, separately for each group. Spearman correlations were used because of nonnormality of distributions and to reduce the impact of outlier values. Age was included because of its well-documented relationship to WM (Jost, Bryck, Vogel, & Mayr, 2011; Myers, Emery, White, & Hale, 2003). Fisher’s Z-transformation—$Z(R) = 0.5ln[(1 + R)/(1 – R)]$—was used to compare the magnitude of the differences in age-adjusted Spearman correlations between PSZ
and HCs (Kleinbaum, Kupper, & Muller, 1988). As described below, none of these differences was statistically significant; therefore, we (a) recomputed the correlations after pooling the data from HCs and PSZ; (b) applied the method of Meng, Rosenthal, and Rubin (1992) to test for all pairwise differences in magnitude of correlations of MCCB individual cognitive domain \( T \) scores (i.e., working memory, processing speed, etc.) with \( K \); and (c) used Benjamini and Hochberg’s (1995) false discovery rate method of multiple testing to assess the significance of the \( p \) values for these 41 pairwise correlations. In addition, we fitted linear regression models separately in HCs and PSZ to predict WASI and MCCB scores from \( K \), using the models of the general form Domain Score = Intercept + \( K \). Separate pairs of models (for HCs and PSZ) were fitted for each WASI and MCCB domain score, as well as the MCCB composite score. For each domain or composite score, we then tested the significance of the difference of the estimated slopes for \( K \) (\( \beta_{HC} \) and \( \beta_{PSZ} \)) in the two groups, using the formula \( Z = (\beta_{HC} - \beta_{PSZ})/SE(\beta_{HC})^2 + SE(\beta_{PSZ})^2)^{1/2} \), rejecting equality of the two slopes if \( |Z| > 1.96 \) (Kleinbaum et al., 1988). No HC – PSZ differences of slopes were statistically significant.

No HC – PSZ differences of intercepts were statistically significant. We chose to focus our report on the Spearman results as these are more easily understood.

To assess the extent to which the magnitude of HC – PSZ differences in broader cognition can be explained by differences in \( K \), we compared the magnitude of the group differences for each MCCB cognitive domain, the WASI Vocabulary and Matrix Reasoning subtests, as well as the MCCB overall composite score and WASI estimated IQ score after adjustment for age alone and then after adjustment for both age and \( K \), using analysis of covariance (ANCOVA) models: (a) cognitive \( T \) score = age + group, and (b) cognitive \( T \) score = age + \( K \) + group. Inclusion of age in these models was motivated by the observation that, whereas MCCB and WASI \( T \) scores are normed for age, change-localization \( K \) scores decrease with age.

### Results

Table 2 shows mean scores for \( K \) and each of the domain-specific and composite measures of cognitive ability. PSZ performed significantly lower than HCs on the MCCB composite score and all seven domain scores as well as on the WASI estimated IQ and composite scores (\( ps < .001 \)).

As seen in Figure 1, the overall distribution of \( K \) scores was substantially shifted toward lower values in PSZ relative to HCs, even though it appears that the use of four-item arrays created an artificial performance ceiling for HCs. Nearly half of the HCs (47%) had a \( K \) score of 3 or higher, whereas only 12% of PSZ achieved scores of 3 or higher. In contrast, nearly a third of PSZ (31.3%) had \( K \) scores of 2 or lower, whereas only 2.6% of the HC group had \( K \) scores in that range. The difference between groups on \( K \) was robust: \( t(174) = 7.61, p < .001 \), Cohen’s \( d = 1.11 \) (Cohen, 1992).

The Spearman correlations, adjusting for age, between \( K \) and the MCCB and WASI scores are shown in Table 3, along with the results of statistical tests of whether the correlations differed across groups. Although the correlations were consistently higher among PSZ than HCs (which could reflect the greater range of values among PSZ), none of these correlations differed significantly between the two groups. Accordingly, the two samples were pooled to compute age-adjusted partial Spearman correlations in the total sample.

The correlations calculated on the combined sample were higher than in either group alone, likely because the range of values was greater in the pooled group than in either diagnostic group considered separately (see the within-group correlation matrix in online Supplemental Table 1 and the combined group matrix in online Supplemental Table 2). Correlations between \( K \) and the individual MCCB and WASI domains in the pooled sample ranged from a low of \( r = .40 \) for Social Cognition to a high of \( r = .67 \) for Processing Speed, with eight of 11 correlations exceeding Cohen’s \( r \approx .50 \) criterion for a large effect size. Interestingly, although \( K \) is a measure of WM, its correlation with the MCCB Working Memory domain was no higher than the correlations observed with some of the other domains. This likely reflects the coarseness of the MCCB domain measures, which are highly intercorrelated (August, Kiwanuka, McMahon, & Gold, 2012).

After adjusting for multiple comparisons, the Spearman correlation between \( K \) and processing speed (\( r = .67 \)) was significantly higher than the correlation of any other MCCB domain with \( K \), and

<table>
<thead>
<tr>
<th>Test</th>
<th>Schizophrenia</th>
<th>Healthy controls</th>
<th>Cohen’s ( d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change-localization K score</td>
<td>2.34 (0.6)</td>
<td>2.93 (0.43)</td>
<td>1.11</td>
</tr>
<tr>
<td>WASI estimated IQ</td>
<td>97.08 (14.39)</td>
<td>113.86 (10.77)</td>
<td>1.30</td>
</tr>
<tr>
<td>WASI Matrix Reasoning ( T ) score</td>
<td>49.32 (9.94)</td>
<td>58.64 (6.47)</td>
<td>1.08</td>
</tr>
<tr>
<td>WASI Vocabulary ( T ) score</td>
<td>46.28 (11.18)</td>
<td>56.52 (8.41)</td>
<td>1.02</td>
</tr>
<tr>
<td>MCCB composite ( T ) score</td>
<td>30.71 (13.51)</td>
<td>52.55 (11.72)</td>
<td>1.71</td>
</tr>
<tr>
<td>MCCB Processing Speed</td>
<td>34.85 (11.75)</td>
<td>54.38 (10.84)</td>
<td>1.72</td>
</tr>
<tr>
<td>MCCB Attention Vigilance</td>
<td>39.43 (10.84)</td>
<td>52.25 (9.37)</td>
<td>1.25</td>
</tr>
<tr>
<td>MCCB Working Memory</td>
<td>38.67 (10.68)</td>
<td>53.14 (10.70)</td>
<td>1.35</td>
</tr>
<tr>
<td>MCCB Verbal Learning</td>
<td>38.81 (10.42)</td>
<td>51.42 (11.23)</td>
<td>1.17</td>
</tr>
<tr>
<td>MCCB Visual Learning</td>
<td>34.1 (13.53)</td>
<td>46.52 (11.24)</td>
<td>0.99</td>
</tr>
<tr>
<td>MCCB Reasoning/Problem Solving</td>
<td>42.46 (10.02)</td>
<td>52.95 (10.29)</td>
<td>1.03</td>
</tr>
<tr>
<td>MCCB Social Cognition</td>
<td>38.76 (13.23)</td>
<td>51.68 (10.77)</td>
<td>1.06</td>
</tr>
</tbody>
</table>

Note. WASI = Wechsler Abbreviated Scale of Intelligence; MCCB = MATRICS Consensus Cognitive Battery. The WASI estimated IQ score is reported as a standard score with a mean of 100 and SD of 15, and the individual subtests are reported as \( T \) scores with a mean of 50 and SD of 10. All of the MCCB scores are reported as \( T \) scores.
This document is copyrighted by the American Psychological Association or one of its allied publishers.

higher than the WASI estimated IQ, Matrix Reasoning, and Vocabulary correlations. The correlation between K and Social Cognition was significantly lower than the correlations between K and Processing Speed, Reasoning/Problem Solving, and Working Memory. The MCCB overall composite score had a significantly higher correlation (r = .68) with K than any individual MCCB domain except Processing Speed (which had a correlation of r = .90 with the MCCB composite score), and higher than the correlations between K and the WASI estimated IQ, Matrix Reasoning, and Vocabulary scores. Note that the interpretation of these differences among correlations may be impacted by differences in reliability among the domain and composite scores and therefore should be regarded with caution.

Given that WM capacity (assessed by change-localization K score) was correlated with measures of multiple cognitive domains, we addressed the question of whether impaired patient performance in these domains could be explained by reduced K. As described in the statistical methods, we ran separate ANCOVA models for each cognitive domain with domain score as the dependent variable, group as the independent variable, and age as a covariate. We next added the K score as a covariate to determine the amount of between-groups variance that could be accounted for by K. K was a significant covariate at p < .001 in all models except when Social Cognition was the independent variable, where p = .18. Diagnosis was significant at the p = .001 level in all models, with and without adding K; however, including K in the model resulted in substantial reductions in between-groups differences across all domains.

Figure 2 illustrates these results, with the age-adjusted group difference shown in light gray and the age- and K-adjusted difference score shown in dark gray. On average, the between-groups difference across all scores was reduced by approximately 40% when K was added as a covariate. That is, the difference in K between groups accounted for approximately 40% of the between-groups differences across the MCCB and the WASI. This is a large proportion of the variance, but K did not account for the entirety of the reduced MCCB and WASI performance in PSZ. A similar conclusion can be reached by examining the regression lines in

Table 3

<table>
<thead>
<tr>
<th>MCCB and WASI domain</th>
<th>$R_{\text{HC}}$ ($n = 77$)</th>
<th>$R_{\text{PSZ}}$ ($n = 99$)</th>
<th>$R_{\text{Pooled}}$ ($N = 176$)</th>
<th>Test for HCs vs. PSZ difference in correlation (T) $p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WASI estimated IQ</td>
<td>.26*</td>
<td>.44***</td>
<td>.56***</td>
<td>−1.36 .172</td>
</tr>
<tr>
<td>WASI Matrix Reasoning T score</td>
<td>.28*</td>
<td>.40***</td>
<td>.51***</td>
<td>−0.92 .359</td>
</tr>
<tr>
<td>WASI Vocabulary T score</td>
<td>.24</td>
<td>.35***</td>
<td>.48***</td>
<td>−0.81 .415</td>
</tr>
<tr>
<td>MCCB composite</td>
<td>.45***</td>
<td>.57***</td>
<td>.68***</td>
<td>−1.01 .313</td>
</tr>
<tr>
<td>MCCB Attention/Vigilance</td>
<td>.31***</td>
<td>.51***</td>
<td>.60***</td>
<td>−1.60 .110</td>
</tr>
<tr>
<td>MCCB Processing Speed</td>
<td>.46***</td>
<td>.48***</td>
<td>.67***</td>
<td>−0.16 .876</td>
</tr>
<tr>
<td>MCCB Reason/Problem Solving</td>
<td>.43***</td>
<td>.41***</td>
<td>.57***</td>
<td>0.21 .835</td>
</tr>
<tr>
<td>MCCB Social Cognition</td>
<td>.17</td>
<td>.23*</td>
<td>.40***</td>
<td>−0.39 .696</td>
</tr>
<tr>
<td>MCCB Verbal Learning</td>
<td>.25*</td>
<td>.41***</td>
<td>.55***</td>
<td>−1.12 .261</td>
</tr>
<tr>
<td>MCCB Visual Learning</td>
<td>.25</td>
<td>.37***</td>
<td>.47***</td>
<td>−0.84 .402</td>
</tr>
<tr>
<td>MCCB Working Memory</td>
<td>.28*</td>
<td>.49***</td>
<td>.58***</td>
<td>−1.63 .103</td>
</tr>
</tbody>
</table>

Note. $R_{\text{HC}}, R_{\text{PSZ}},$ and $R_{\text{Pooled}}$ = Spearman coefficients between the change-location K score and MCCB or WASI domain T scores in healthy control subjects (HCs), people with schizophrenia (PSZ), and all participants combined (pooled), respectively. WASI = Wechsler Abbreviated Scale of Intelligence; MCCB = MATRICS Consensus Cognitive Battery; $T =$ test statistic for difference in correlation coefficients based on normal approximation to the Fisher’s Z-transformation of $R_{\text{HC}}$ and $R_{\text{PSZ}}$. $^* p < .05$. $^{**} p < .01$.

Figure 3: The slopes of the regression lines are virtually identical for PSZ and HCs, but the line indicating the expected value of MCCB scores for PSZ at each value of K is lower than the line indicating the expected values for HCs, indicating that PSZ have reduced MCCB performance compared with HCs beyond what would be predicted from their lower K scores. For example, at K = 2.6, the mean value of K in the pooled sample, the predicted between-groups difference ≥ .6E for the MCCB composite score is 13.6 ± 1.9, t(172) = 7.10, $p < .001$. Thus, although we cannot be certain of the direction of causation, the decreased WM capacity in PSZ can potentially account for 40% of the broad cognitive deficit measured by the MCCB.

Discussion

These results provide several important insights into the relationship between WM capacity and broad measures of cognitive ability in the healthy population and in PSZ. First, PSZ demonstrated a substantial reduction in visual WM capacity (Cohen’s $d = 1.11$) using a simple visual change-localization task that is believed to provide a relatively pure measure of the amount of information that can be actively maintained in a short-term buffer. Thus, it is not necessary for a WM task to involve memory manipulation, task switching, or distractor suppression to yield a very robust deficit in schizophrenia. These more complex processes may also be an area of deficit in PSZ, but the current results clearly demonstrate a robust impairment in WM storage capacity.

---

1 We observed modest, but significant, correlations between haloperidol equivalents and MCCB composite score and WASI estimated IQ scores ($rs = -.26$ and -.23, respectively).
The extent of impairment in our sample was substantially larger than documented in a prior meta-analysis of WM in schizophrenia that examined a wide range of tasks but did not include change-detection measures (Lee & Park, 2005). However, our results are consistent with a meta-analysis of spatial WM performance that reported an overall effect size of 1.00 (Piskulic et al., 2007). Furthermore, the size of the K reduction was comparable to the effect size for the MCCB Working Memory domain score in the present study, which was in turn comparable to the effect size observed for the MCCB Working Memory domain score in a previous large study (Keefe et al., 2011). Effect sizes are limited by the reliability of the measure, and one possible reason for the discrepancy between these large effect sizes and the smaller average effect size in the Lee and Park (2005) meta-analysis is that their meta-analysis included studies using a wide variety of tasks, many of which may have had lower reliability than the highly reliable K and MCCB scores (which were designed to maximize reliability). An additional possibility is that K is a purer measure of WM capacity than the measures used in some of the studies included in the meta-analysis, which may reduce variance related to factors that do not differ as much between PSZ and HCs.

A second key finding was that WM capacity estimates derived from this simple task were strongly correlated with broader measures of cognitive ability, which provides evidence that robust correlations with higher cognitive abilities can be found for a relatively pure measure of WM storage capacity and does not require the inclusion of more complex executive processes such as the manipulation of representations, task switching, or distractor suppression. The relationship between K and higher cognitive abilities was very broad and robust, with correlations ranging from \( r = .47 \) to \( .68 \) in the pooled sample, suggesting that WM capacity

---

**Figure 2.** Age-adjusted and age plus K-adjusted Wechsler Abbreviated Scale of Intelligence (WASI) and the MATRICS Consensus Cognitive Battery (MCCB) marginal mean group differences. All differences are calculated from \( T \) scores except for WASI estimated IQ score, which is the difference of standard scores, rescaled to have an SD of 10. Adjusted differences are estimated with analysis of covariance models. All \( p \) values for adjusted between group differences were \( p < .001 \).

**Figure 3.** MATRICS Consensus Cognitive Battery (MCCB) composite score versus K by group, with plotted linear regressions (not adjusted for age) for healthy controls (HCs; composite score = 9.93 + 14.55 \( \times \) K) and people with schizophrenia (PSZ; composite score = \(-0.05 + 13.16 \times K\)). Differences between the slopes of the two regression lines are not statistically significant, \( r(172) = 0.42, p = .68 \).
is a critical cognitive capacity that is powerfully associated with performance across the types of cognitive functions demanded in the MCCB and WASI. This is consistent with prior findings in healthy populations (Fukuda et al., 2010), but this issue has not been addressed in the schizophrenia literature beyond two studies from our group with relatively modest sample sizes (Gold et al., 2010; Leonard et al., 2012).

The correlations between K and the MCCB and WASI did not vary significantly as a function of group, although they were consistently higher in the PSZ, likely reflecting the greater range of scores in this group compared with the HC group. This may be somewhat surprising because one typically thinks of WM deficits in schizophrenia as a consequence of the disease process, which might be expected to alter the relationship between WM and other aspects of the cognitive architecture. Some, but not all, prior studies that examined the relationship of complex WM measures and other aspects of cognitive performance in PSZ suggest an amplification of these relationships (Gold et al., 1997; Perry et al., 2001; Silver et al., 2003). Indeed, recent electrophysiological evidence from our group suggests that WM capacity reduction in PSZ is mediated by processes different from those that produce variations in K among healthy individuals (Leonard et al., 2012), raising the possibility that the relationship between WM and other aspects of cognitive performance might be altered in PSZ. However, the present findings, based on large samples of PSZ and HCs, suggest that the normative patterns of covariation between abilities are preserved in PSZ. PSZ may have arrived at their low K scores as the result of a pathological process, but the effect on broader cognitive function remains the same. By analogy, automobile tires normally lose air pressure over time because of simple diffusion, but air pressure can also be decreased by a nail puncture; in either case, the reduced air pressure has the same effect on the automobile’s performance. This suggests that new treatments designed to “reinflar” WM capacity in PSZ could potentially have a substantial impact on broader cognitive abilities.

Interestingly, K was correlated approximately equally with most of the individual cognitive domains sampled by the WASI and MCCB, with no compelling evidence of a stronger correlation with the MCCB Working Memory domain. There were two exceptions: a larger correlation with the Processing Speed and a smaller correlation with the Social Cognition domains. Note, however, that even the smallest correlation in the pooled-groups analysis exceeded Cohen’s $r = .30$ criteria for a medium effect size. Thus, although some of the correlations differed statistically across domains, the theoretical or practical implications of these differences appear to be minimal. WM storage capacity may play a particularly large role in measures of processing speed because greater storage capacity may allow people to process multiple items in parallel, avoiding the need for slower serial processing. The lower correlation with the MCCB Social Cognition domain may indicate that “cold” cognitive factors such as WM capacity play a smaller role in social cognition than in traditional cognitive tasks.

The most striking aspect of the data is that a relatively pure measure of WM storage correlated with all of the abilities sampled by the WASI and MCCB. There are at least two means of accounting for this observation. First, one could argue that the MCCB and WASI subscales do a poor job of isolating specific cognitive abilities, evidenced by the fact that the different domain and composite scores are robustly intercorrelated (see online Supplemen tal Tables 1 and 2 and August et al., 2012). From this perspective, the WM storage correlations might be interpreted as evidence for a relationship with general cognitive ability, not broad correlation across discrete abilities. Alternatively, WM capacity may play a similar role across these various domains, perhaps explaining the correlations among the domain scores as well as with the WM capacity score. Given that this broad pattern of correlations was observed in both HCs and PSZ, it is likely to be true in other clinical populations, although this remains to be documented.

The ANCOVA results shown in Figure 2 are consistent with the correlational data. Covarying K substantially reduced the magnitude of between-groups differences across all WASI and MCCB measures and accounted for nearly 40% of the between-groups difference. Thus, the between-groups difference in WM capacity does not fully account for the generalized cognitive deficit observed in schizophrenia, but it accounts for an impressively large proportion. It is possible that similar ANCOVA findings would be observed using measures of processing speed (or others) as potential covariates. However, one important advantage of the use of WM in these types of analyses is that the neural system involved with variation in WM capacity (especially visual WM storage capacity) is much better defined than that involved with processing speed. Progress in understanding the neural basis and cognitive mechanisms implicated in WM impairment in schizophrenia may offer critical targets for psychological and pharmacological treatment development as any improvement in this area will likely have a positive impact on broad aspects of cognitive performance.

One possible concern about the interpretation of these data is the extent of other cognitive processes. Specifically, might reductions in processing speed or degradation of perceptual processing be implicated in reduced WM capacity? In a prior study, we directly compared performance with 100-ms and 500-ms sample durations and did not find better patient performance with the longer arrays, arguing against a role for perceptual encoding speed deficits (Gold, Wilk, McMahon, Buchanan, & Luck, 2003). Furthermore, we have shown that the rate of decay of iconic memory is not different between PSZ and HCs (Hahn et al., 2011). In another study, we documented reductions in WM capacity in the absence of any evidence that the perceptual precision of WM representations was altered in PSZ (Gold et al., 2006). Thus, it does not appear likely that impairments in processing speed or perceptual processing can account for the WM capacity deficit in PSZ.

This study has two potential limitations. Most important, correlations are not evidence of causation, nor can they establish the directionality of the effects. For example, one could argue that people who are more cognitively able may be able to manage their use of WM capacity more efficiently than those who are less cognitively able. This explanation, although a logical possibility, fails to address why they are more cognitively able. Thus, the idea that variation in WM capacity is predictive of broader cognitive performance appears to be more parsimonious but cannot be proven by the correlational approach taken here. Furthermore, we cannot rule out the possibility of an unknown third variable that is driving variation in both WM capacity and broad cognitive performance. The fact that our results are consistent with prior studies in healthy adults may serve to reduce the concern that an obscure third variable, specific to PSZ, is driving the findings.
Our HCs, despite being recruited by random digit dialing, scored well above the population mean on the WASI, and it is possible that different relationships would be observed in a lower ability sample. It is also possible that the magnitude of our effect size for WM capacity may be amplified by this IQ difference. Note, however, that our groups were matched on parental education years, suggesting they were drawn from similar backgrounds. Moreover, our controls scored very close to the standardization sample mean of the MCCB. Thus, it is unlikely that our effect size was substantially overestimated because of a “super normal” control group. Indeed, the effect size may have been somewhat underestimated because we had only a single, relatively brief measure of WM capacity. Moreover, the fact that our measure of WM capacity accounted for approximately 40% of the patient deficit on the MCCB cannot be explained by the unexpectedly high IQ scores of the HCs (given that their MCCB scores were typical), nor can the strong correlations between WM capacity and both the MCCB and WASI IQ. Thus, the higher than expected IQ scores of the HCs are not a major cause for concern.

References


Working Memory and Broad Cognitive Ability


Received July 19, 2012
Revision received January 21, 2013
Accepted January 28, 2013