Working memory in schizophrenia: a meta-analysis

N. F. Forbes, L. A. Carrick, A. M. McIntosh and S. M. Lawrie*

University of Edinburgh, Department of Psychiatry, Royal Edinburgh Hospital, Edinburgh, UK

Background. Memory impairment is being recognized increasingly as an important feature of the neuropsychology of schizophrenia. Dysfunction of working memory, a system for the short-term storage and manipulation of information, may relate to a number of core symptoms of schizophrenia. Many studies have examined working memory function in schizophrenia but a clear understanding of the nature and extent of any deficit has been elusive.

Method. A systematic review and meta-analysis of studies comparing working memory function in subjects with schizophrenia and healthy controls was performed. Following a comprehensive literature search, meta-analyses were conducted on 36 measures of phonological, visuospatial and central executive working memory functioning, encompassing 441 separate results from 187 different studies.

Results. Statistically significant effect sizes were found for all working memory measures, indicating deficits in schizophrenia groups. Some of these were robust findings in the absence of evidence of significant heterogeneity or publication bias. Meta-regression analyses showed that the working memory deficit was not simply explained by discrepancies in current IQ between schizophrenia and control groups.

Conclusions. Large deficits in working memory were demonstrated in schizophrenia groups across all three working memory domains. There were, however, no clear differences across subdomains or between particular working memory tasks. There was substantial heterogeneity across results that could only be partly explained.

Received 20 November 2007; Revised 4 August 2008; Accepted 5 September 2008; First published online 23 October 2008

Key words: Meta-analysis, schizophrenia, systematic review, working memory.

Introduction

Schizophrenia is a heterogeneous disorder, probably encompassing several disease processes that disturb a range of human characteristics including thought, perception, emotion, drive and behaviour. Although neurocognitive dysfunction in schizophrenia is sometimes overshadowed by more obvious symptoms such as hallucinations, delusions and thought disorder, it is a major contributor to the disability experienced by people with schizophrenia (Green & Nuechterlein, 1999) and may in fact represent a core feature of the disease from which the others arise.

There has been much debate as to whether the putative neuropsychological impairment in schizophrenia is a generalized one (Blanchard & Neale, 1994), or whether certain functions are differentially impaired. For example, increased distractibility (Oltmanns, 1978), memory impairment (comparable to classical amnesic syndrome; McKenna et al. 1990), attentional disturbance (Nuechterlein & Dawson, 1984; cf. Kenny & Meltzer, 1991) and executive dysfunction (Weinberger et al. 1986) have all been suggested as core features. Resolution of this question is complicated because different tests vary in their degree of difficulty and their sensitivity to the effects of impaired functioning (Chapman & Chapman, 1973; Calev, 1984). In addition, the variation in intellectual difficulties, sometimes on top of an apparent decline from pre-morbid or general levels, may range from a gross impairment in some patients with schizophrenia to complete preservation of function in a minority, rendering a precise formulation difficult.

A meta-analysis by Heinrichs & Zakzanis (1998) examined a comprehensive range of studies looking at selected neurocognitive variables in schizophrenia. They demonstrated that neurocognitive deficit was a reliable finding in schizophrenia and that deficits were present on most of the tasks examined, although no task could discriminate completely between schizophrenia patients and controls. Impaired verbal memory was a reliable finding in patients with schizophrenia compared to healthy controls. They found far fewer studies reporting performance on non-verbal memory tasks; the magnitude of performance deficits among schizophrenia patients on these tasks was comparable overall to their deficits on verbal...
memory tasks but much more variable across studies and therefore less reliable. Different measures of global intellectual functioning produced different results, with the Wechsler Adult Intelligence Scale – Revised (WAIS-R) IQ test being the most reliable at demonstrating a deficit in schizophrenia groups.

A meta-analysis by Aleman et al. (1999), looking specifically at memory impairment in schizophrenia, also found evidence of a significant association between schizophrenia and memory dysfunction. According to Cohen’s (1988) effect size nomenclature, there was a large impairment on tests of recall and a moderate impairment on tests of recognition, suggesting the possibility of a particular deficit in retrieval. There was significant impairment on measures of short-term memory but no difference between measures of verbal and non-verbal memory.

Lee & Park’s (2005) meta-analysis looked specifically at working memory in schizophrenia. They demonstrated significant deficits in working memory across all 124 studies included, but suggested that these deficits were more consistent for visuospatial than for verbal working memory tasks. They also showed that the extent of the working memory deficit in schizophrenia groups was unrelated to the length of delay involved in the task, implicating a deficit in encoding rather than storage or retrieval.

Working memory is a concept of short-term memory that is a flexible and dynamic memory resource, where sensory information can be held briefly and manipulated. Baddeley (1986) described working memory as ‘a system for the temporary holding and manipulation of information during the performance of a range of neurocognitive tasks such as comprehension, learning and reasoning’, implying an active process rather than simply a static, temporary store. Goldman-Rakic (1994) postulated a neuropsychological model of schizophrenia with working memory dysfunction as the core feature. Specifically, she related working memory dysfunction to formal thought disorder, which she argued was characterized by a failure to adequately retain recent ideas or utterances in working memory.

Baddeley’s (1986) model of working memory comprised three subsystems: a supervisory, executive system, and two ‘slave systems’, the phonological loop and the visuospatial sketchpad. The phonological loop consists of a short-term phonological store, supported by an articulatory rehearsal process. Material enters the phonological store either as auditory inputs, such as speech, or through non-auditory modalities, for example printed words or pictures representing specific words. The visuospatial sketchpad acts as a short-term store for spatial information and visual characteristics such as shape, colour, brightness and pattern. The central executive has a supervisory attentional role, coordinating the activities of the other working memory subsystems. Specifically, its role is conceived as regulating memory and action by integrating information and selecting strategies for processing it, rather than as attending to perception in a more general sense. Baddeley’s (1986) model was refined recently to include a fourth component, the episodic buffer (Repovs & Baddeley, 2006). Together with the central executive, the episodic buffer integrates information from working memory and other systems, including semantic and long-term memory.

Many studies have examined working memory function in people with schizophrenia, and robust, reliable deficits across a range of measures have been demonstrated (Lee & Park, 2005). We set out to replicate these findings by conducting a thorough and wide-ranging systematic review and meta-analysis of working memory in schizophrenia, comparing performance on tests of working memory in people with schizophrenia versus healthy control subjects. We aimed to determine whether the working memory deficit in schizophrenia, if it exists, is a global impairment, or whether there are specific domains of working memory, measured by specific neuropsychological tests or subtests, that are differentially impaired. By adopting broader inclusion criteria and incorporating data on more working memory tasks than Lee & Park (2005), we hoped to extend their findings by including measures of executive working memory functioning, and calculating overall, pooled effect sizes for many specific tests of working memory.

We hypothesized that any working memory deficit in schizophrenia would represent a specific neurocognitive deficit, rather than simply a correlate of general intellectual ability (IQ). We also set out to explore the relationship between working memory in schizophrenia and illness duration, antipsychotic medication, and positive and negative symptoms, as well as comparing first-episode and chronic patient groups. These variables have not previously been examined so comprehensively and specifically with respect to working memory function in schizophrenia.

**Method**

**Literature search**

A comprehensive search for studies was conducted in Medline, EMBASE, Cinahl and PsycINFO, with all articles published after 1966 and before the end of 2005 considered. The terms ‘schizophrenia’ and ‘memory’ and related terms were searched, both as free text and as exploded subject headings. The specific search terms were adapted to the specific headings and
thesauri of the databases searched. The full search strategy as conducted in Medline is detailed in the Appendix.

This search produced 6094 results, 5470 of which were English language publications. Titles, abstracts and, where necessary, the full text of the articles identified were examined for possible inclusion in the analysis. In addition, references in the bibliographies of included studies and relevant review articles were examined.

**Inclusion criteria**

Studies were included if they met the following criteria:

1. They measured the performance of people with schizophrenia (however diagnosed) and that of healthy control subjects on a test of working memory under comparable conditions. Studies that used other psychiatric or medical patients, prisoners, or relatives of people with schizophrenia as controls were excluded because of the possibility of subtle neurocognitive impairment existing in these groups. Only studies where the mean age in both schizophrenia and control groups was under 65 years were included to reduce possible contamination by age-related neurocognitive decline.

2. They included a valid test of working memory. For tests of phonological or visuospatial working memory, the delay period between presentation of material and recall or recognition testing had to be 2 minutes or less. For visual information, the stimulus had to be presented for at least 1 second to ensure adequate registration. When in doubt, the appropriateness of including tests as measures of working memory was determined by consulting Lezak (1995) or Baddeley et al. (1995).

3. They reported sufficient data to enable calculation of an effect size for each test; that is, clear reporting of the number of subjects undertaking each test, mean scores and standard deviation or standard error for each group.

4. They were published in English language journals.

**Data extraction**

For each working memory test reported in each study, the following data, where provided, were recorded for both schizophrenia and control groups:

1. Number of subjects completing the test.
2. Current IQ, as measured by the WAIS (Wechsler, 1955), WAIS-R (Wechsler, 1981) or Quick Test (Ammons & Ammons, 1962); where full-scale IQ was extrapolated from WAIS or WAIS-R subtests, only estimates based on four or more subtest scores were included.

3. Mean test score and standard deviation or standard error.

In addition, the following data were recorded for the schizophrenia group:

1. Duration of illness in years.
2. Whether or not experiencing first episode of illness.
3. Symptom ratings by the Positive and Negative Syndrome Scale (PANSS; Kay et al. 1987) positive scale score (PANSS-P) and negative scale score (PANSS-N).
4. Number of subjects on antipsychotic medication and mean dose in chlorpromazine equivalents for both first- and second-generation antipsychotics (Woods, 2003).

Where descriptive data were unclear, for example where all of the group to whom the demographic data applied did not complete the memory test, then data were extrapolated to describe the subgroup that participated in memory testing. Some papers sub-divided the schizophrenia group (e.g. by symptom profile, chronicity of illness or medication status) and gave results for two or more subgroups. Where first-episode and chronic schizophrenia patients were differentiated, data for both schizophrenia groups were recorded. Otherwise, means and standard deviations were pooled and entered for the schizophrenia group as a whole, to avoid multiple entries for the same control sample.

Several multiple publications were located, where data from the same subjects were reported by more than one study. To avoid multiple entries for the same sample, only the data from the study with the largest number of subjects were entered.

**Data analysis**

Meta-analyses were performed on tests of working memory for which three or more different studies reported results. Statistical calculations were performed using Stata version 8.2 (Stata Corp., College Station, TX, USA).

From the data reported for each working memory test in each study, a random effects model (allowing for heterogeneity between studies) was used to calculate effect size for the difference in test performance between schizophrenia and control groups. Effect sizes were estimated using the unbiased estimator, $d$ (Hedges & Olkin, 1985), and are presented with 95% confidence intervals (CIs). The $d$ statistic is calculated as the difference between the two group means.
divided by the pooled standard deviation. The $d$ statistic was corrected for small sample bias using a multiplier that corrects for the tendency of small studies to overestimate the effect size.

The summary effect size ($d+$), its significance and CI were calculated for each working memory measure using a random effects meta-analysis. The DerSimonian & Laird (1986) method was used to calculate all summary estimates, and includes an estimate of the variance in effect size between study populations (i.e. heterogeneity). The significance and magnitude of any heterogeneity were calculated using the Q and $I^2$ statistics respectively. $I^2$ provides a measure of the proportion of inconsistency between studies attributable to heterogeneity, and is described by Higgins et al. (2003). Galbraith plots were constructed to look for outliers in analyses with significant heterogeneity.

To further investigate causes for heterogeneity, meta-regression analyses were performed for the following variables: duration of illness in years, difference in current IQ between schizophrenia and control groups, PANSS-P score, PANSS-N score, and current antipsychotic medication dose in chlorpromazine equivalents. The Stata program metareg.ado was used throughout and the Restricted Maximum Likelihood (REML) method was used to estimate the model parameters. Subgroup analysis was also performed to explore possible heterogeneity between first-episode and non-first-episode groups.

Publication bias was investigated in two ways. First, funnel plots of effect sizes from individual studies against each study’s sample size were constructed. Second, a statistical analogue of the funnel plot was used, the linear regression approach described by Egger et al. (1997). This approach examines the association between effect size and standard error for each study, and therefore takes into account both the sample size and the statistical power of each study in relation to effect size.

**Results**

The literature search and review of bibliographies identified 200 references that met criteria for inclusion. Results for specific working memory tests were categorized together, and a random effects meta-analysis was conducted for those tests of working memory where results had been reported by three or more different studies. In total, meta-analyses were conducted for 36 different tests of working memory derived from 187 different studies. These yielded 441 sets of data incorporating a total of 17 152 working memory test scores from subjects with schizophrenia and 14 605 from healthy controls.

**Phonological working memory**

Table 1 displays the results of meta-analyses of schizophrenia and control subjects in performance on tests of phonological working memory. Twenty-one tests or subtests with at least three contributing studies (range up to 38 studies) were identified. For all of these tests, scores were worse in schizophrenia than control groups at a statistically significant level, with absolute effect sizes ranging from 0.55 to 1.41. It is notable, however, that the 95% CIs for all of these tests overlap, that is this meta-analysis provides no evidence for differential impairment on one particular phonological working memory task.

Fourteen of these tests were associated with significant heterogeneity between studies with $I^2 >50\%$ and $p$ values $<0.05$. Seven tests showed significant differences between schizophrenia and control groups that were not potentially attributable to heterogeneity or publication bias: the long Digit Span Distraction Test (DSDT) (non-distraction condition), long DSDT (distraction condition), short DSDT (non-distraction condition), digit span backwards, verbal learning tests list A trial 1, verbal learning tests list B, and verbal span tasks not involving manipulation of information. Representative Forest plots from the phonological working memory meta-analyses are shown for digit span backwards (Fig. 1) and letter number span (Fig. 2).

Fig. 2 demonstrates one potential explanation of this high level of heterogeneity, in that Antonova et al. (2005) appears to be an outlier. We also considered the possibility that some of this heterogeneity may relate to the tasks we combined in individual meta-analyses, as less heterogeneity is generally evident in the analyses of the better circumscribed tests or subtests, such as digit span. However, subgroup analyses comparing those verbal learning tests where the items belong to clear semantic categories (as in the California Verbal Learning Test; Delis et al. 1987) against those where the items are not categorized (as in the Rey Auditory Verbal Learning Test, described by Lezak, 1995) identified no significant differences between the two.

**Visuospatial working memory**

We found 11 tests or subtests of visuospatial working memory with at least three studies contributing usable data (range up to 33 studies), all of which found statistically significantly impairments in schizophrenia versus control groups, with absolute effect sizes ranging between 0.51 and 1.29 (Table 2). Again, the CIs of these effect sizes overlapped and there was a suggestion that those tests with higher effect sizes
were also associated with greater heterogeneity. Seven tests were not associated with significant heterogeneity or publication bias. These were: complex figure reproduction tests, spatial delayed response tasks, tests of spatial recognition, tests of pattern recognition, spatial span backwards, spatial span forwards and backwards combined, and visual paired associate learning tests. However, the single visuospatial working memory test for which there were most data, spatial span forwards, showed evidence of significant heterogeneity and publication bias. Figs 3 and 4 show representative Forest plots from the spatial

<table>
<thead>
<tr>
<th>Measure</th>
<th>Study (n)</th>
<th>Effect size (95% CI)</th>
<th>p value (effect size)</th>
<th>I² (%) (heterogeneity)</th>
<th>p value (heterogeneity)</th>
<th>t (publication bias)</th>
<th>p value (publication bias)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consonant Trigram Test</td>
<td>7</td>
<td>−0.77 (−1.09 to −0.44)</td>
<td>&lt;0.001</td>
<td>54.0</td>
<td>0.042</td>
<td>−1.54</td>
<td>0.183</td>
</tr>
<tr>
<td>Long DSDT: distraction condition</td>
<td>3</td>
<td>−1.41 (−2.12 to −0.70)</td>
<td>&lt;0.001</td>
<td>65.3</td>
<td>0.056</td>
<td>−1.18</td>
<td>0.448</td>
</tr>
<tr>
<td>Long DSDT: non-distraction condition</td>
<td>3</td>
<td>−1.07 (−1.54 to −0.59)</td>
<td>&lt;0.001</td>
<td>29.5</td>
<td>0.242</td>
<td>8.53</td>
<td>0.074</td>
</tr>
<tr>
<td>Short DSDT: distraction condition</td>
<td>6</td>
<td>−1.3 (−1.73 to −0.87)</td>
<td>&lt;0.001</td>
<td>64.9</td>
<td>0.014</td>
<td>−1.10</td>
<td>0.335</td>
</tr>
<tr>
<td>Short DSDT: non-distraction condition</td>
<td>6</td>
<td>−1.0 (−1.25 to −0.76)</td>
<td>&lt;0.001</td>
<td>0.0</td>
<td>0.544</td>
<td>−1.29</td>
<td>0.265</td>
</tr>
<tr>
<td>Digit Span Backwards</td>
<td>31</td>
<td>−0.87 (−0.96 to −0.77)</td>
<td>&lt;0.001</td>
<td>6.7</td>
<td>0.36</td>
<td>0.69</td>
<td>0.493</td>
</tr>
<tr>
<td>Digit Span Forwards</td>
<td>38</td>
<td>−0.55 (−0.7 to −0.4)</td>
<td>&lt;0.001</td>
<td>65.7</td>
<td>&lt;0.001</td>
<td>0.78</td>
<td>0.439</td>
</tr>
<tr>
<td>Digit Span Forwards and Backwards</td>
<td>17</td>
<td>−0.92 (−0.69 to −1.16)</td>
<td>&lt;0.001</td>
<td>68.0</td>
<td>&lt;0.001</td>
<td>3.36</td>
<td>0.004</td>
</tr>
<tr>
<td>Letter Number Span</td>
<td>12</td>
<td>−1.36 (−1.63 to −1.09)</td>
<td>&lt;0.001</td>
<td>69.4</td>
<td>&lt;0.001</td>
<td>−1.70</td>
<td>0.121</td>
</tr>
<tr>
<td>Tests of Passage Recall</td>
<td>37</td>
<td>−1.35 (−1.50 to −1.20)</td>
<td>&lt;0.001</td>
<td>67.0</td>
<td>&lt;0.001</td>
<td>−2.27</td>
<td>0.029</td>
</tr>
<tr>
<td>Verbal Paired Associate Learning Tests</td>
<td>18</td>
<td>−0.88 (−1.07 to −0.69)</td>
<td>&lt;0.001</td>
<td>64.0</td>
<td>&lt;0.001</td>
<td>−0.88</td>
<td>0.390</td>
</tr>
<tr>
<td>Verbal Learning Tests: List A Trial 1</td>
<td>11</td>
<td>−1.01 (−1.26 to −0.76)</td>
<td>&lt;0.001</td>
<td>62.1</td>
<td>0.36</td>
<td>−1.01</td>
<td>0.339</td>
</tr>
<tr>
<td>Verbal Learning Tests: List A Trial 5</td>
<td>13</td>
<td>−1.25 (−1.54 to −0.97)</td>
<td>&lt;0.001</td>
<td>78.8</td>
<td>&lt;0.001</td>
<td>−0.88</td>
<td>0.399</td>
</tr>
<tr>
<td>Verbal Learning Tests: List A Trial 6, cued</td>
<td>6</td>
<td>−0.98 (−1.32 to −0.65)</td>
<td>&lt;0.001</td>
<td>59.2</td>
<td>0.031</td>
<td>−6.14</td>
<td>0.004</td>
</tr>
<tr>
<td>Verbal Learning Tests: List A Trial 6, free</td>
<td>13</td>
<td>−1.08 (−1.35 to −0.80)</td>
<td>&lt;0.001</td>
<td>76.6</td>
<td>&lt;0.001</td>
<td>−0.56</td>
<td>0.629</td>
</tr>
<tr>
<td>Verbal Learning Tests: Total List A, Trials 1-5</td>
<td>28</td>
<td>−1.32 (−1.48 to −1.17)</td>
<td>&lt;0.001</td>
<td>60.4</td>
<td>&lt;0.001</td>
<td>−0.88</td>
<td>0.399</td>
</tr>
<tr>
<td>Verbal Learning Tests: List B</td>
<td>6</td>
<td>−0.92 (−1.19 to −0.65)</td>
<td>&lt;0.001</td>
<td>40.4</td>
<td>0.136</td>
<td>0.59</td>
<td>0.596</td>
</tr>
<tr>
<td>Verbal Recognition Tests</td>
<td>19</td>
<td>−0.63 (−0.84 to −0.42)</td>
<td>&lt;0.001</td>
<td>58.9</td>
<td>0.001</td>
<td>3.00</td>
<td>0.008</td>
</tr>
<tr>
<td>Verbal Recall Tests</td>
<td>24</td>
<td>−1.21 (−1.39 to −1.03)</td>
<td>&lt;0.001</td>
<td>53.3</td>
<td>0.001</td>
<td>−0.21</td>
<td>0.838</td>
</tr>
<tr>
<td>Verbal Span Tasks involving manipulation</td>
<td>9</td>
<td>−1.02 (−1.32 to −0.71)</td>
<td>&lt;0.001</td>
<td>67.6</td>
<td>0.002</td>
<td>0.24</td>
<td>0.819</td>
</tr>
<tr>
<td>Verbal Span Tasks not involving manipulation</td>
<td>7</td>
<td>−0.72 (−0.91 to −0.51)</td>
<td>&lt;0.001</td>
<td>0.0</td>
<td>0.433</td>
<td>0.77</td>
<td>0.475</td>
</tr>
</tbody>
</table>

DSDT, Digit Span Distraction Test; CI, confidence interval.
Overall task. Battery (CANTAB) spatial working memory strategy Cambridge Neuropsychological Test Automated in Table 3. Fig. 5 shows the Forest plot for the heterogeneity or publication bias. These are shown none were associated with statistically significant effect sizes ranging between 0.73 and 0.92, and (range up to six studies). All found statistically significant with at least three studies contributing usable data Four tests of executive working memory were found meta-analyses.

Executive working memory

Four tests of executive working memory were found with at least three studies contributing usable data (range up to six studies). All found statistically significant effect sizes ranging between 0.73 and 0.92, and none were associated with statistically significant heterogeneity or publication bias. These are shown in Table 3. Fig. 5 shows the Forest plot for the Cambridge Neuropsychological Test Automated Battery (CANTAB) spatial working memory strategy task.

Meta-regression analysis

We were interested a priori in determining whether some of the working memory deficits in schizophrenia could be related to key clinical features of the illness. We also sought, post hoc, to see whether some of the heterogeneity we found by calculation of the I statistic could be explained using meta-regression analyses.

We conducted meta-regression analyses for duration of illness in years against all tests where there were sufficient data. We found significant associations for the following five tests: spatial working memory strategy task ($b = 0.027, Z = 2.31, p = 0.021$), verbal learning tests cued recall post-distraction ($b = -0.037, Z = -3.27, p = 0.001$), verbal learning tests free recall post-distraction ($b = -0.0339, Z = -2.44, p = 0.015$), visual paired associate learning tests ($b = -0.15, Z = -2.39, p = 0.017$) and immediate visual recall tests ($b = -0.062, Z = -2.81, p = 0.005$) (note that the sign of these beta coefficients reflects the scoring of the tests). These associations suggest that studies including patients with longer illness durations generally find greater deficits on these tasks.

We also looked at the difference in current IQ between schizophrenia and control groups, and conducted meta-regression analyses for these values against all tests where there were sufficient data. There was only one statistically significant association, between passage recall and IQ difference ($b = -0.074, Z = -3.42, p = 0.001$). Therefore, in general, the pattern...
of working memory deficits was not explained by any discrepancies in current IQ between schizophrenia and control groups.

We looked at positive and negative symptoms scores, measured by the PANSS (Kay et al. 1987). For verbal list learning list 1, significant heterogeneity was associated with greater symptom severity on both PANSS-P ($b=0.09$, $Z=2.11$, $p=0.035$) and PANSS-N ($b=0.13$, $Z=2.34$, $p=0.019$). For digit span forwards and backwards combined, significant heterogeneity was also associated with greater symptom severity on both PANSS-P ($b=0.154$, $Z=3.65$, $p<0.001$) and PANSS-N ($b=0.139$, $Z=3.75$, $p<0.001$).

Finally, we looked at current antipsychotic dose in chlorpromazine equivalents, which was also associated with heterogeneity in verbal list learning list 1 ($b=-0.0026$, $Z=-4.29$, $p<0.001$), verbal list learning list 5 ($b=-0.0027$, $Z=-3.98$, $p<0.001$), total recall on verbal learning tests recall trials ($b=-0.0013$, $Z=-3.98$, $p=0.008$) and verbal leaning tests, short delay, free recall ($b=-0.002$, $Z=-2.56$, $p=0.01$).

**Subgroup analysis**

We then examined whether there were greater deficits in multiple-episode versus first-episode schizophrenia groups, for working memory tests with data from at least two first-episode and two multiple-episode groups. For 11 of the 13 working memory tests analysed, performance in the sample was worse in the multiple-episode than first-episode patients. We found five tests or subtests where performance was worse, to a statistically significant degree, in those with multiple episodes compared to first episodes: verbal learning tests list 5, verbal learning tests list 6 (free recall), tests of pattern recognition, spatial span forwards and tests of immediate visual recall. It is, however, striking that none of these test were also found to be associated with duration of illness.

**Discussion**

The number of studies included in the present meta-analysis makes it larger than any published

---

### Table 2. Results of meta-analyses of tests of visuospatial working memory

<table>
<thead>
<tr>
<th>Measure</th>
<th>Study (n)</th>
<th>Effect size (95% CI)</th>
<th>p value (effect size)</th>
<th>$I^2$ (%) (heterogeneity)</th>
<th>$p$ value (heterogeneity)</th>
<th>$t$ (publication bias)</th>
<th>$p$ value (publication bias)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benton Visual Retention Test</td>
<td>9</td>
<td>$-1.29$ (1.73 to -0.85)</td>
<td>&lt;0.001</td>
<td>83</td>
<td>&lt;0.001</td>
<td>0.40</td>
<td>0.701</td>
</tr>
<tr>
<td>Complex figure reproductiona (errors)</td>
<td>3</td>
<td>0.95 (0.49, 1.41)</td>
<td>&lt;0.001</td>
<td>18.1</td>
<td>0.295</td>
<td>0.43</td>
<td>0.743</td>
</tr>
<tr>
<td>Tests of Immediate Facial Recognition</td>
<td>15</td>
<td>$-0.82$ (-1.02 to -0.61)</td>
<td>&lt;0.001</td>
<td>44.1</td>
<td>0.034</td>
<td>-1.26</td>
<td>0.230</td>
</tr>
<tr>
<td>Tests of Pattern Recognition</td>
<td>9</td>
<td>$-0.84$ (-1.04 to -0.64)</td>
<td>&lt;0.001</td>
<td>42.5</td>
<td>0.084</td>
<td>-1.18</td>
<td>0.277</td>
</tr>
<tr>
<td>Spatial Delayed Response Tasksa (error distance)</td>
<td>3</td>
<td>0.85 (0.5, 1.19)</td>
<td>&lt;0.001</td>
<td>0.0</td>
<td>0.955</td>
<td>-0.07</td>
<td>0.956</td>
</tr>
<tr>
<td>Tests of Spatial Recognition</td>
<td>5</td>
<td>$-0.76$ (-0.96 to -0.56)</td>
<td>&lt;0.001</td>
<td>0.0</td>
<td>0.41</td>
<td>-1.55</td>
<td>0.219</td>
</tr>
<tr>
<td>Spatial Span Backwards</td>
<td>6</td>
<td>$-0.99$ (-1.27 to -0.72)</td>
<td>&lt;0.001</td>
<td>27.6</td>
<td>0.227</td>
<td>-1.59</td>
<td>0.186</td>
</tr>
<tr>
<td>Spatial Span Forwards</td>
<td>19</td>
<td>$-0.94$ (-1.08 to -0.79)</td>
<td>&lt;0.001</td>
<td>39.1</td>
<td>0.042</td>
<td>-3.60</td>
<td>0.002</td>
</tr>
<tr>
<td>Visual Paired Associate Learning Tests</td>
<td>6</td>
<td>$-0.91$ (-1.24 to -0.57)</td>
<td>&lt;0.001</td>
<td>51.5</td>
<td>0.067</td>
<td>0.50</td>
<td>0.646</td>
</tr>
<tr>
<td>Immediate Visual Recall Tests</td>
<td>33</td>
<td>$-0.87$ (-1.01 to -0.72)</td>
<td>&lt;0.001</td>
<td>63.2</td>
<td>$&lt;0.001$</td>
<td>$-1.20$</td>
<td>0.239</td>
</tr>
<tr>
<td>Visuospatial Span – forwards and backwards</td>
<td>3</td>
<td>$-0.51$ (-0.88 to -0.16)</td>
<td>0.004</td>
<td>53.4</td>
<td>0.117</td>
<td>$-0.22$</td>
<td>0.865</td>
</tr>
</tbody>
</table>

CI, Confidence interval.
a Tests where lower scores indicate better performance.
meta-analysis of working memory in schizophrenia to date. In the majority of the present analyses, funnel plots and statistical tests of publication bias demonstrated little evidence of reporting bias, suggesting that studies of working memory in schizophrenia are equally likely to be published regardless of the direction, magnitude or statistical significance of their results. Most of the analyses in the present study demonstrated relatively large and statistically robust overall effect sizes. The likelihood that these would be altered in a meaningful way by the inclusion of a small number of unpublished or unidentified studies is therefore low.

Statistically significant overall effect sizes, indicating impaired performance among subjects with schizophrenia, were found for all 36 analyses. These were present and comparable across all three domains of working memory examined, and across the different tests or subtests within each domain. There were also some consistent deficits in analyses where there was no evidence of significant heterogeneity or publication bias. These 18 tasks, which include simple tests such as digit span backwards and spatial span backwards, can therefore be regarded as relatively robust and reliable methods of identifying deficits in patients with schizophrenia.

However, one of the most striking results of this meta-analysis is the amount of heterogeneity between studies and the difficulty in finding consistent associations with this. Some of this inhomogeneity may of course reflect the way that we grouped tests, but we did this with reference to standard tests. Furthermore, it is clear that some single tests, such as forwards digit span, are associated with substantial heterogeneity, whereas some collections of, for example, verbal span tasks are not. Some of this inhomogeneity may of course reflect differences in the way tests are administered and scored but this detail is rarely provided in original papers. Overall, we were left with the impression that the bulk of heterogeneity in these results is potentially attributable to differences in the populations studied, and in particular the possibility that larger effect sizes in some studies could reflect the inclusion of particularly impaired patients.

Subgroup (first-episode versus chronic) and meta-regression (for duration of illness) analyses nonetheless produced some interesting findings in keeping with the possibility of a progressive decline in working memory performance in people with schizophrenia. This would be in accordance with Mohs’ (1999) suggestion that age-related neurocognitive changes may follow an accelerated trajectory in people with schizophrenia compared to healthy individuals. However, we conducted numerous analyses and there is no agreement between the specific tests identified on the subgroup and meta-regression analyses. We cannot therefore say that we have found evidence for progression of cognitive impairment in schizophrenia. Rather, we would use these findings to suggest that future studies should address these possibilities.

The lack of association with deficit in current IQ suggests that the robust finding of working memory deficit in schizophrenia groups across all 36 tasks examined is not simply attributable to IQ deficits. This supports the hypothesis that people with schizophrenia have a specific working memory deficit, rather than a generalized cognitive impairment. Indeed, many of the studies contributing data to the meta-analyses had matched schizophrenia and control subjects for pre-morbid IQ and/or educational attainment. This, however, may have led to the selection of unrepresentative samples from both groups, as suggested by Buckley et al. (1992). Thus, our findings may be more representative of relatively high functioning patients, or indeed of patients with more than average decline from pre-morbid levels of function.

The severity of positive or negative psychotic symptoms did not seem to affect the magnitude of performance deficit across the range of working memory tests, suggesting that the deficits relate to an aspect of the schizophrenic trait, rather than, for
example, distractibility and disorganization in the acute psychotic state. There was, however, a trend towards association between performance on verbal learning tests and both positive and negative symptom scores and antipsychotic dose, suggesting that researchers interested in verbal learning should take careful account of both the symptom and treatment status of their subjects.

The validity of comparing results between different types of working memory test has, however, been called into question. Chapman & Chapman (1973, 1978) observed that subjects with schizophrenia scored less well than healthy controls on most tasks, and that for a deficit to be of particular importance it should be of greater magnitude than would be expected on the basis of the generalized deficit. They pointed out that such differential deficits in ability may not be reflected accurately by differential deficits in test performance because of varying psychometric properties between different tests. For example, tests may differ in the extent to which they discriminate between high and low performers, or in their reliability. Chapman & Chapman (1978) suggested that, to overcome these problems with comparability of results between different tests, tests should be matched for level of difficulty and variance in scores among normal control subjects.

In the present analysis, only the digit span distraction tests had all been developed according to this principle, so that comparisons of the distraction and non-distraction component of each are valid. The other tests of working memory included in the present study were, however, unmatched. This must be considered when comparing effect sizes, both between analyses and within those analyses incorporating different working memory tasks. These measures may be influenced by the psychometric properties of the tests themselves, and their relationship to each other will not necessarily equate to underlying differences in ability between schizophrenia and control groups. Any conclusions drawn from relative magnitudes of overall effect size between analyses must therefore be guarded.

**Conclusions**

Statistically significant deficits in task performance were found in schizophrenia groups in all 36 working memory tasks examined, across the domains of phonological, visuospatial and central executive working memory. For many tasks, these findings could not be explained by publication bias or significant heterogeneity. Meta-regression analyses showed that the working memory deficits were not explained...
by discrepancies in current IQ between schizophrenia and control groups. No consistent associations were found between duration of illness, antipsychotic medication or symptom profile and the working memory deficit in schizophrenia groups. Our conclusions are, however, limited by the significant heterogeneity across studies included in the meta-analyses.

Appendix

**Literature search conducted in Medline**

1. explode SCHIZOPHRENIA or explode SCHIZOPHRENIA, CATATONIC or explode SCHIZOPHRENIA, DISORGANIZED or explode SCHIZOPHRENIA, PARANOID
2. schizophren$ [title, abstract, MeSH]
3. #1 or #2
4. explode MEMORY or explode MEMORY, SHORT-TERM or explode MEMORY DISORDERS or explode ATTENTION
5. ‘memory’ or ‘attention’ or ‘learning’ [title, abstract, MeSH]
6. ‘central executive’ or ‘supervisory attentional system’ [title, abstract, MeSH]
7. ‘phonological loop’ or ‘articulatory loop’ [title, abstract, MeSH]
8. ‘visuo-spatial scratch-pad’ or ‘visuospatial scratch-pad’ or ‘visuo-spatial scratchpad’ or ‘visuo-spatial sketch-pad’ or ‘visuospatial sketch-pad’ or ‘visuospatial sketchpad’ [title, abstract, MeSH]
9. ‘digit sequence’ [title, abstract, MeSH]
10. ‘span’ or ‘supra-span’ [title, abstract, MeSH]
11. ‘immediate recall’ [title, abstract, MeSH]
12. ‘paired associate’ [title, abstract, MeSH]
13. ‘consonant trigrams’ or ‘consonant trigram’ [title, abstract, MeSH]
14. (‘Peterson’ and ‘task’) [title, abstract, MeSH]
15. ‘proactive inhibition’ [title, abstract, MeSH]
16. ‘Corsi’ [title, abstract, MeSH]
17. (‘match’ and ‘sample’) [title, abstract, MeSH]
18. ‘visual distraction’ or ‘visual reproduction’ or ‘visual retention’ or ‘visual learning’ or ‘visual recall’ [title, abstract, MeSH]
19. ‘spatial recall’ or ‘spatial learning’ [title, abstract, MeSH]
20. ‘pattern recognition’ or ‘spatial recognition’ or ‘facial recognition’ [title, abstract, MeSH]
21. (‘Rey’ and ‘Osterreith’) or (‘Rey’ and ‘Osterrieth’) or (‘Rey’ and ‘Oesterreith’) or (‘Rey’ and ‘Oesterrieth’) [title, abstract, MeSH]
22. ‘complex figure’ [title, abstract, MeSH]
23. (‘dual’ and ‘task’) [title, abstract, MeSH]
24. ‘sentence verification’ [title, abstract, MeSH] (‘random’ and ‘generation’) [title, abstract, MeSH]
25. (‘source’ and ‘monitoring’) [title, abstract, MeSH]
26. ‘Wechsler Adult Intelligence Scale’ or ‘WAIS’ [title, abstract, MeSH]
27. ‘Wechsler Memory Scale’ or ‘WMS’ [title, abstract, MeSH]
28. #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28
29. #3 and #29
30. ‘case’ or ‘affected’ or ‘patient’ or ‘proband’ or ‘subject’ or ‘control’ or ‘normal’ or ‘healthy’ or ‘unaffected’ or ‘cohort’ or ‘follow-up’ or ‘case-control’ or ‘comparative study’ or ‘risk’ or ‘odds ratio’ [title, abstract, MeSH]
31. explode CASE-CONTROL STUDIES or explode COHORT STUDIES or explode RISK or explode RISK FACTORS
32. #31 or #32
33. #30 and #33
34. limit #34 to human

Acknowledgements

We thank Wendy Mill for assistance in obtaining articles. A.M.M. is supported by the Health Foundation and S.M.L. is supported by the Sackler Foundation.

Declaration of Interest

None.

References


*Altshuler LL, Ventura J, van Gorp WG, Green MF, Theberge DC, Mintz J (2004). Neurocognitive function in...*
clinically stable men with bipolar I disorder or schizophrenia and normal control subjects. Biological Psychiatry 56, 560–569.


Working memory in schizophrenia


*Landro NI, Pape-Ellefsen E, Hagland KO, Oddland T (2001). Memory deficits in young schizophrenics with normal...


