The neuropsychology of adult obsessive–compulsive disorder: A meta-analysis

Amitai Abramovitch,⁎, Jonathan S. Abramowitz, Andrew Mittelman

Department of Psychiatry, Harvard Medical School, Boston, MA, USA
Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA
Department of Psychology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

HIGHLIGHTS

• Neuropsychological investigations in OCD yield inconsistent and contrasting results.
• We conducted the first meta-analysis of 115 studies including 3452 patients.
• Across domains moderate effect sizes (ES) indicated reduced performance in OCD.
• The magnitude of effect sizes may be clinically insignificant.
• Despite significant heterogeneity, no moderating effects were found.

ABSTRACT

A vast and heterogeneous body of literature on the neuropsychology of obsessive–compulsive disorder (OCD) has accumulated in recent decades, yielding inconsistent results. In an attempt to quantitatively summarize the literature, we conducted a meta-analysis of 115 studies (including 3452 patients), comparing adult OCD patients with healthy controls on tests of 10 neuropsychological domains. Across studies, medium mean effect sizes were found for all executive function subdomains, processing speed, and sustained attention. Small effect sizes were found for visuospatial abilities and working memory. A large effect size was found for non-verbal memory whereas a small effect size was found for verbal memory, where only the former was found to be associated with impairments in executive functions. Moderators of effect sizes were also investigated. Results are discussed in terms of their clinical significance as well as their implications for current neurobiological models of OCD and methodological caveats.

© 2013 Elsevier Ltd. All rights reserved.
1. Introduction

Obsessive–compulsive disorder (OCD) is an often-debilitating condition with a worldwide prevalence of 1.5–3% (Okasha, 2003; Ruscio, Stein, Chiu, & Kessler, 2010). The main symptoms of OCD include intrusive thoughts, urges and images that cause marked anxiety or distress (i.e., obsessions) and repetitive mental acts or behavioral rituals (i.e., compulsions) which the person feels compelled to perform in order to reduce or prevent the distress provoked by the obsessive thoughts (American Psychiatric Association, 2013). Although the classification, assessment, and treatment of OCD have traditionally focused on the presence of obsessions, compulsions, and anxiety, investigators (e.g., Chamberlain, Blackwell, Fineberg, Robbins, & Sahakian, 2005) have hypothesized the presence of neuropsychological deficits which might be associated with the clinical phenotype or predict treatment outcome (D’Alicante et al., 2012). Memory deficits, for example, might explain compulsive checking behaviors (e.g., Woods, Vevea, Chambless, & Bayen, 2002). Accordingly, identifying neuropsychological deficits might prove to be an important avenue in better understanding and treating this condition. Despite the accumulation of well over one hundred studies over several decades, however, there have been no attempts to quantitatively review this body of literature. Thus, the present study represents the first meta-analysis of neuropsychological research on OCD.

Prevailing neuropsychological models of OCD (i.e., the frontostriatal model; Saxena & Rauch, 2000) are derived from a distinct pattern of increased resting state frontostriatal activation and observations that symptom severity is positively correlated with activation in corticostriato–thalamo–cortico (CSTC) circuits (e.g., Harrison et al., 2009). Imaging studies have also reported abnormal activation in OCD across major sites of the frontostriatal system during performance on neuropsychological tasks such as the Tower of London (TOL), trail making test, Go/No-Go, Wisconsin card sorting test (WCST), Stroop, and Stop Signal Task (SST) (Ayicigei, Dinn, Harris, & Erkmen, 2003; Kwon et al., 2003; Lucey et al., 1997; Nabeyama et al., 2008; Schlosser et al., 2010; van den Heuvel et al., 2005), Interest in these biological models and findings has led to the hypothesis that neuropsychological functions subserved by the frontostriatal system would be affected in OCD. However, research has yielded highly inconsistent findings across both adult and pediatric samples (Abramovitch, Mittelman, Henin, & Geller, 2012; Kuelz, Hohagen, & Voderholzer, 2004).

Given the focus on the orbitofrontal cortex (OFC), dorsolateral prefrontal cortex (DLPFC), and anterior cingulate cortex (ACC) and their connection to the basal ganglia (thought to primarily subserve higher order executive functioning), the majority of neuropsychological studies in OCD have targeted performance on tests of executive functioning, generally observing statistically significant differences between patients and controls, with OCD patients performing more poorly (Kuelz et al., 2004). Of specific interest are tests of response inhibition, a construct in which deficient performance has been proposed as an endophenotype of OCD (Chamberlain et al., 2005). Yet whereas some studies found statistically poorer performance on response inhibition tasks, including the Go/No-Go, SST, and Stroop tests (Abramovitch, Dar, Schweiger, & Hermesh, 2011; Martinot et al., 1990; Menzies et al., 2007; Penades et al., 2007; van den Heuvel et al., 2005), others noted no differences in performance on these tests between OCD patients compared to controls (Boone, Ananth, Philpott, Kaur, et al., 1991; Krishna et al., 2011; Rao, Reddy, Kumar, Kandavel, & Chandrashekar, 2008). Similarly, whereas some studies found poorer performance in OCD patients on tests of planning ability (e.g., the TOL test; Nielen & Den Boer, 2003; van den Heuvel, Veltman, Groenewegen, Cath, et al., 2005) and set shifting (e.g., the WCST; Anderson, 2002; Cavedini, Zorzi, Piccinni, Cavallini, & Bellodi, 2010; Okasha et al., 2000), several other investigations found that OCD patients evidence no difference from healthy individuals on these tasks (Abbruzzese, Ferri, & Scarone, 1995; Henry, 2006; Purcell, Maruff, Kyrios, & Pantelis, 1998).

Memory functioning in OCD has also been the subject of controversy. Some studies report poorer performance relative to healthy individuals on verbal memory tests such as the California Verbal Learning Test (CVLT; Cha et al., 2008; Deckersbach et al., 2004; Hartl et al., 2004), yet a number of other studies found that OCD patients performed worse than healthy controls on such tasks (de Geus, Denys, Sitskoorn, & Westenberg, 2007; Kitsis et al., 2007; Moritz, Kloss, von Eckstaedt, & Jelinek, 2009). On the other hand, there is more consistent evidence that individuals with OCD score significantly lower than controls on non-verbal memory tasks, specifically as measured by the Rey–Osterrieth Complex Figure Test (ROCF). Numerous studies report reduced performance on the Copy, Immediate, and Delayed Recall components of the ROCF (Penades, Catalan, Andres, Salamero, & Gasto, 2005; Rajender et al., 2011; Savage et al., 1999; Shin et al., 2004). However, Savage et al. (1999) suggested that executive function impairments mediate this effect in that OCD patients demonstrate poor organizational abilities that consequently impair coding and retrieval of information on the ROCF. These findings have been reliably replicated (Buhlmann et al., 2006; Penades et al., 2005). Notably, very few studies compared neuropsychological test performance of OCD patients with other clinical groups. Moreover, these studies used relatively small sample sizes and usually compared OCD to other anxiety disorders. Thus, there is very limited information available in which a cogent inference could be made regarding OCD disorder-specific domains of impairment.

The inconsistent pattern of results described above has been attributed to several factors, including the highly heterogeneous nature of OCD (e.g., Hashimoto et al., 2011) and the high rate of comorbidity with depression (e.g., Basso, Bornstein, Carona, & Morton, 2001). Other authors have suggested that variability in neuropsychological findings stems from disorder-specific reduced processing speed that negatively impacts performance on most tests, especially those measuring executive functions (Bedard, Joyal, Godbout, & Chantal, 2009). It is also possible that inconsistent results in the domain of memory functions in OCD stem from a lack of confidence in memory, rather than a memory deficit per se (Dar, Rish, Hermesh, Taub, & Fux, 2000). Still an additional caveat is that in many studies, authors did not make alpha corrections for multiple comparisons (e.g., Kuelz et al., 2004), which is particularly important in neuropsychological studies where multiple outcome measures are usually examined.

Not surprisingly, attempts to synthesize and summarize the results of neuropsychological research on OCD have reached discrepant conclusions (Greisberg & McKay, 2003; Henry, 2006; Kuelz et al., 2004; Tallis, 1997). The ability to draw strong inferences from narrative summaries of this literature is further obfuscated by the large number of domains and corresponding neuropsychological tests, as well as potentially important clinical (e.g., OCD symptom severity), demographic (e.g., medication status), and methodological differences (e.g., number of tests in the battery) that might moderate findings across studies. Surprisingly, however, no attempts have yet been made to apply meta-analytic procedures to quantitatively review this literature. Accordingly, the primary aim of the present meta-analysis was to determine the magnitude of differences between individuals with OCD and healthy individuals.
(controls) for the various domains of neurocognitive function. We also examined numerous clinical, demographic, and methodological variables as potential moderators of effect size across studies. Due to the abundance of neurocognitive constructs and outcome measures, we chose to focus on 6 primary neuropsychological domains and a total of 10 subdomains. We excluded reward-based tests, decision-making tasks, and studies that incorporated outcome measures not considered inherent to the classic tests (e.g., organizational abilities in the ROCF test).

2. Method

2.1. Retrieval and selection of studies

Published studies (in English) on neuropsychological functioning in adult OCD patients were identified by searching the MEDLINE, ISI Web of Knowledge and PsycINFO electronic databases, publication reference lists, and issues of relevant scientific journals through February, 2012. We also solicited unpublished data from researchers publishing in the field of neuropsychology of OCD. This pursuit yielded 207 research articles. In further designating research as appropriate for inclusion in the meta-analysis, only controlled studies were considered. That is, we included a study only if it contained at least one comparison between a group of DSM-diagnosed adult (18 years or older) OCD patients (i.e., via a structured interview) and a healthy (i.e., screened for the absence of a psychiatric or neurologic diagnosis) adult control group using one or more neuropsychological tests. Studies that lacked a healthy control group and only reported comparisons between multiple patient groups (e.g., patients with OCD vs. patients with schizophrenia) were omitted, as were single-group or within-subject group (e.g., pre/post treatment) studies. Treatment studies were included only if a pre-treatment comparison was available between OCD patients and healthy controls. One hundred seventy-seven of the 207 studies met these initial criteria.

Of the 177 studies, thirteen had to be excluded because they provided insufficient information for neither calculation nor estimation of effect sizes. Twelve additional studies were excluded because they utilized neuropsychological tasks that were beyond the scope of this review (e.g., decision-making tasks, and emotional Stroop). Thirty studies were excluded due to the exclusive use of neuropsychological tasks that were either significantly modified from the original version or rarely used. Finally, seven studies were excluded because they contained previously published data. This left a final count of 115 studies, the year of publication ranging from 1989 to 2012.

2.2. Variables recorded and coded from studies

The following general information was recorded from the 115 studies: (a) publication status, (b) year of publication, and (c) country where the study was conducted. Characteristics of participants in each study were recorded as follows: (a) OCD and control group sample sizes, (b) mean age, (c) mean age of OCD onset, (d) years of education, (e) percent of males in the OCD group, (f) mean scores on measures of OCD (e.g., the Yale-Brown Obsessive Compulsive Scale) and depressive symptoms (e.g., the Beck Depression Inventory), and (g) percent of OCD patients receiving SRI or neuroleptic (i.e., antipsychotic) medication. Table 1 summarizes the characteristics of the 115 studies. As can be seen, group sizes ranged widely, as did the percent of males included in the groups and the severity of OCD and depressive symptoms. The average OCD sample had moderately severe OCD symptoms and generally mild depressive symptoms. Of the countries, the largest quantity of studies (23) was conducted in the United States, followed by Germany (18), South Korea (13), Spain (8), and the United Kingdom (8). Finally, 40 studies reported that their OCD group was unmedicated. Across the 64 other studies that reported medication status, the average percent of OCD patients on either an SRI or neuroleptic was 66.48% (SD = 26.40, range = 9.5%-90.5%).

We recorded the specific neuropsychological tests used in each study along with the domain of functioning (e.g., memory) that the test measured. We also coded the subdomain of functioning measured by each test (e.g., verbal memory, non-verbal memory). Table 2 lists the domains, subdomains, and tests coded from the studies in the meta-analysis. In some cases, some but not all of the outcome variables were unreported. In these cases (e.g., the TMB–TMA index score), we recorded only the conventional outcome variables. Finally, from each study we recorded (a) the number of tests, (b) number of testing sessions, and (c) average length of testing sessions.

2.3. Effect size computation

Because studies used such a diversity of instruments, we expressed the results for all tests in all studies in terms of Cohen's d, a standardized effect size measure. Cohen's d is the difference between two means divided by the standard deviation of the difference scores. Because studies used such a diversity of instruments, we expressed the results for all tests in all studies in terms of Cohen's d, a standardized effect size measure. Cohen's d is the difference between two means divided by the standard deviation of the difference scores. Because studies used such a diversity of instruments, we expressed the results for all tests in all studies in terms of Cohen's d, a standardized effect size measure.
measure of effect size, which was calculated by subtracting the control group’s mean from the OCD group’s mean and dividing by the pooled standard deviation. Thus, each test on which an OCD group was compared with a control group yielded its own effect size within each study. Positive effect sizes (e.g., \( d = 0.25 \)) indicated higher neuropsychological functioning in the OCD group relative to the control group, whereas negative effect sizes (e.g., \( d = -0.50 \)) indicated that the opposite effect was present. Cohen (1977) suggested that effect size magnitudes of 0.2, 0.5, and 0.8 correspond to small, medium, and large effects, respectively.

With few exceptions, we calculated effect sizes directly from means and standard deviations reported in studies. When this information was not available, results of other statistical tests (e.g., values of \( t \) or \( F \)) were used to calculate or estimate effect size (See Ray & Shadish, 1996). Effect sizes were calculated after the coding was completed, so as to reduce the potential for knowledge of a study’s results to bias the coding of its characteristics. To determine whether effect sizes were consistent across comparisons, we calculated a homogeneity statistic \( Q \), which had an approximate chi-square distribution with \( k - 1 \) df, where \( k \) is equal to the number of effect sizes.

2.4. Preliminary analyses

Within each domain and subdomain of neuropsychological functioning, the differences between OCD and control groups in a given study were often assessed by multiple tests (\( M = 5.23 \); range, 1–24). We therefore calculated an average effect size for each domain and subdomain within each study using the available test results as listed in Table 2. This procedure yielded 272 domain effect sizes and 321 subdomain effect sizes. Additionally, we retained separate effect sizes for each test, each subdomain, and each domain so that all could be addressed in separate analyses.

In two studies (Mataix-Cols et al., 2006; Roth, Milovan, Baribeau, & O’Connor, 2005), authors included multiple OCD and control groups for the purpose of examining theoretically important research questions (i.e., neuropsychological performance associated with gender and age of onset). Thus, in these studies it was possible to derive multiple OCD-control comparisons on each dependent measure. However, variability in effect sizes derived from multiple comparisons within the same study is likely to be less than variability in effect sizes drawn from different studies (Robinson, Berman, & Neimeyer, 1990). To eliminate this potential for non-independence of observations, we aggregated our data such that each study contributed only a single effect size in each analysis. Accordingly, in each of the two studies mentioned above, we used the mean effect size from across the multiple comparisons.

We examined the likelihood that our effect sizes were inflated due to a publication bias. Such a bias may occur if studies reporting significant findings (large effect sizes) are published, whereas those obtaining null results (small effect sizes) are not. We computed the fail safe \( N \) (Orwin, 1983) to determine the number of unpublished studies obtaining small effect sizes required to reduce our overall mean effect size (\( d = 0.499 \)) to a trivial level. This analysis indicated that over 172 unpublished studies with effect sizes of 0.20 would be needed to reduce our effect size to a level not significantly different from ‘no effect’ and thus overturn our findings. If the \( N \) fail-safe number is larger than the observed number of already published studies, this is thought to be sufficient to suggest the absence of a significant publication bias. Thus, we concluded that our results were unlikely to have been affected by publication bias. Finally, to correct for familywise inflation of type I error, a conservative significance threshold of 0.01 was determined for moderator analyses.

3. Results

Our meta-analytic results are presented in the following sequence: First, we present overall mean effect sizes for the domains and subdomains of functioning. Next, we report the results of correlational analyses to identify potential moderators of the domain and subdomain effect sizes. Potential moderators included clinical variables (OCD and depressive symptom severity, and medication status), sample variables (age of onset, percent males, and mean age), and variables related to the administration of neuropsychological batteries (number of tests in the battery, number of testing sessions, and length of testing sessions). We used the random effect model, assuming heterogeneity across \( k \) samples because as mentioned above, neuropsychological studies in this population yield notoriously inconsistent and somewhat contradictory results. Thus, we assume that these differences would yield effect size heterogeneity across studies. For these reasons, we were unable to assume a common effect size across studies as in the case of the fixed effect model. The random effects model was applied to produce the pooled effect size \( d \) by weighting each study’s effect size with its standard error, corrected for possible heterogeneity by \( \tau^2 \).

3.1. Mean effect sizes for domains and subdomains

Stem-and-leaf plots of the effect sizes derived from each study for each domain are shown in Fig. 1. As can be seen, with the exception of attention (which only included effect sizes from five studies), the majority of study effect sizes clustered around the means. The only extreme effect sizes (i.e., \( \geq 4.0 \) standard deviation difference from the mean) were observed within the executive functioning and processing speed domains. In fact, these studies (Boldrini et al., 2005; Cavedini et al., 2010) contributed extreme effect sizes across a number of their reported outcome measures and were subsequently excluded from our analyses, resulting in a final sample of 113 studies.

Table 3 presents a summary of the mean effect sizes for each of the six neuropsychological domains. As can be seen, all mean effect sizes were negative, indicating that across studies, OCD patients performed worse than the healthy control groups. Although generally of medium magnitude, all six mean effect sizes differed significantly from no effect. Of the six domains, the largest effect size was found on tests of memory where on average, OCD patients performed almost two-thirds of a standard deviation worse than control groups. Working memory evidenced the smallest effect size, with OCD patients performing just over one third of a standard deviation worse than controls. Excluding the attention domain, the test of homogeneity (\( Q \)) was significant across domains, suggesting that effect sizes calculated for these domains were heterogeneous over and above the heterogeneity explained by the standard error.

Table 4 presents a summary of the mean effect sizes for the subdomains of neuropsychological functioning. As can be seen, all mean effect sizes were negative (i.e., control groups outperformed OCD patients) and significantly different from no effect. There was, however, a greater range in subdomain effect sizes (\( -.33 \) to \( -.76 \)) than was observed among the domain effect sizes (\( -.34 \) to \( -.63 \)). Even within the same domain, the variability in subdomain effect sizes was noteworthy. For example, within the memory domain, the mean effect size for nonverbal memory (\( d = 0.76 \), which also happened to be the largest subdomain effect size) was substantially larger than that for verbal memory (\( d = 0.33 \); which happened to be the smallest subdomain effect size).

This discrepancy within the same domain is uncommon in the neuropsychological literature. As mentioned above, it has been suggested that executive functioning (organizational strategies) is related to poor performance on the ROCF, and not necessarily non-verbal memory impairments in OCD. To address this question, we examined how strongly effect sizes for executive functioning were correlated with effect sizes for non-verbal memory and for verbal memory. Among the 26 studies available for this computation, a significant positive correlation (\( r = 0.60, p = .001 \)) was found between effect sizes derived from tests of executive function and effect sizes derived from tests of non-verbal memory. On the other hand, effect sizes derived from tests of executive function were not significantly correlated with those derived from tests of non-verbal memory.
from tests of verbal memory on tests of executive function was not significantly correlated with those derived from tests of verbal memory (21 studies; \( r = .05, p = .84 \)). Subsequent Fisher's Z transformation analysis revealed a significant difference between the two correlation coefficients \( Z = 2.04, p = .03 \), indicating that in studies where more OCD participants were using neuroleptics, these participants performed more poorly relative to healthy controls. However, when a correction for multiple comparisons was applied \( (p = .01) \), these correlations were no longer significant.

### 3.2. Moderators of effect size

With respect to symptom-severity, neither global OCD severity (Y-BOCS) nor depression severity (Hamilton Depression Scale and BDI) were associated with effect sizes for any domains or subdomains of neuropsychological functioning. With respect to medication status, although the percent of OCD patients using SRI medication in each group was unrelated to effect size, the percent of patients using neuroleptic medication was significantly negatively associated with the effect size for executive functioning across 27 studies \( (r = -.43, p = .03) \), as well as for the executive function subdomain set-shifting/cognitive flexibility across 21 studies \( (r = -.48, p = .03) \). This indicated that in studies where more OCD participants were using neuroleptics, these participants performed more poorly relative to healthy controls. However, when a correction for multiple comparisons was applied \( (p = .01) \), these correlations were no longer significant.

Two patient characteristics were initially associated with effect size. First, age of onset of OCD was negatively associated with the effect size for executive functioning (50 studies; \( r = -.35, p = .014 \)) and set shifting/cognitive flexibility (44 studies; \( r = -.35, p = .021 \)), indicating that the onset of OCD symptoms at an older age was associated with worse performance in these areas of functioning. Second, the percent of males in the OCD groups was positively associated with the effect size for the working memory domain (26 studies; \( r = .46, p = .02 \)), indicating that the greater the proportion of men in the OCD groups, the better the group performed relative to controls in this domain. However, following a correction for multiple comparisons the above moderating effects became non-significant. Further analysis

### Table 3

<table>
<thead>
<tr>
<th>Domain</th>
<th>Sig of Q</th>
<th>Q</th>
<th>DF</th>
<th>Numerator</th>
<th>C</th>
<th>Tau²</th>
<th>Effect size</th>
<th>Variance</th>
<th>Stand. error</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>Z-value</th>
<th>P value (2 tail)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention</td>
<td>9.49</td>
<td>8.98</td>
<td>5</td>
<td>3.98</td>
<td>7930</td>
<td>.05</td>
<td>-.409</td>
<td>.020</td>
<td>.140</td>
<td>-.775</td>
<td>-.224</td>
<td>-.354</td>
<td>0.0004</td>
</tr>
<tr>
<td>Executive functions</td>
<td>107.52</td>
<td>170.36</td>
<td>86</td>
<td>84.36</td>
<td>1190.26</td>
<td>.07</td>
<td>-.408</td>
<td>.007</td>
<td>.142</td>
<td>-.580</td>
<td>-.415</td>
<td>-.1187</td>
<td>0.0000</td>
</tr>
<tr>
<td>Memory</td>
<td>65.17</td>
<td>129.23</td>
<td>49</td>
<td>80.23</td>
<td>71430</td>
<td>.11</td>
<td>-.630</td>
<td>.004</td>
<td>.062</td>
<td>-.751</td>
<td>-.509</td>
<td>-.1026</td>
<td>0.0000</td>
</tr>
<tr>
<td>Processing speed</td>
<td>69.83</td>
<td>77.90</td>
<td>53</td>
<td>24.90</td>
<td>660.47</td>
<td>.04</td>
<td>-.517</td>
<td>.002</td>
<td>.048</td>
<td>-.611</td>
<td>-.423</td>
<td>-.1047</td>
<td>0.0000</td>
</tr>
<tr>
<td>Visuospatial abilities</td>
<td>53.76</td>
<td>59.35</td>
<td>41</td>
<td>18.35</td>
<td>639.09</td>
<td>.03</td>
<td>-.350</td>
<td>.002</td>
<td>.048</td>
<td>-.445</td>
<td>-.256</td>
<td>-.7272</td>
<td>0.0000</td>
</tr>
<tr>
<td>Working memory</td>
<td>40.11</td>
<td>46.88</td>
<td>28</td>
<td>18.88</td>
<td>422.08</td>
<td>.04</td>
<td>-.341</td>
<td>.004</td>
<td>.065</td>
<td>-.468</td>
<td>-.214</td>
<td>-.5257</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Sig of Q = \( \chi^2 \) threshold for significance of Q; DF = degrees of freedom; CI = confidence interval.

**Fig. 1.** Stem and leaf plots of individual study effect sizes (Cohen's \( d \)) for the six neuropsychological domains. Digits to the left (stems) of the vertical line are read as the ones and tenths place of each study's effect size. Numbers to the right of the vertical line (leaves) are the hundredths place for each study's effect size. Multiple leafs indicate that there were multiple effect sizes with the same stem.
revealed that the mean age of the OCD group was unrelated to any effect sizes.

Finally, we examined whether variables related to the administration of the neuropsychological test battery were associated with effect size. However, correlational analyses revealed that neither the number of tests in the battery, the length of the testing session, nor the number of testing sessions was significantly associated with any effect sizes.

4. Discussion

Over the past several decades, a large and heterogeneous body of literature on the neuropsychology of OCD has accumulated, yielding inconsistent results. Narrative reviews of this literature, while helpful, are subject to selection and interpretation bias. Accordingly, the aim of the present meta-analysis was to quantitatively summarize the findings from this literature and draw conclusions regarding the neuropsychological performance of OCD patients relative to healthy controls. Our review encompassed 113 studies including over three thousand patients with OCD. Statistically significant effect sizes were found across all domains (i.e., attention, executive function, memory, visuospatial abilities, processing speed and working memory) and subdomains (i.e., sustained attention, planning, response inhibition, set shifting/cognitive flexibility, verbal memory, non-verbal memory, visuospatial abilities, processing speed, memory, visuospatial abilities, working memory, spatial working memory).

Table 4: Weighted mean effect sizes and tests of homogeneity by neuropsychological sub-domain.

<table>
<thead>
<tr>
<th>Domain/subdomain</th>
<th>Sig of Q</th>
<th>Q</th>
<th>Df</th>
<th>Numerator</th>
<th>C</th>
<th>Tau²</th>
<th>Effect size</th>
<th>Variance</th>
<th>Stand. error</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>Z-value</th>
<th>P value (2 tail)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sustained attention</td>
<td>9.49</td>
<td>8.98</td>
<td>5</td>
<td>3.98</td>
<td>79.30</td>
<td>0.05</td>
<td>−0.499</td>
<td>0.020</td>
<td>0.140</td>
<td>−0.775</td>
<td>−0.224</td>
<td>−3.554</td>
<td>0.0004</td>
</tr>
<tr>
<td>Executive functions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planning</td>
<td>19.68</td>
<td>12.33</td>
<td>12</td>
<td>0.33</td>
<td>180.78</td>
<td>0.00</td>
<td>−0.440</td>
<td>0.005</td>
<td>0.072</td>
<td>−0.581</td>
<td>−0.299</td>
<td>−6.113</td>
<td>0.0000</td>
</tr>
<tr>
<td>Response inhibition</td>
<td>47.40</td>
<td>90.52</td>
<td>34</td>
<td>56.52</td>
<td>472.24</td>
<td>0.12</td>
<td>−0.492</td>
<td>0.006</td>
<td>0.078</td>
<td>−0.646</td>
<td>−0.338</td>
<td>−6.268</td>
<td>0.0000</td>
</tr>
<tr>
<td>Set shifting/cognitive flexibility</td>
<td>84.82</td>
<td>137.23</td>
<td>66</td>
<td>71.23</td>
<td>949.22</td>
<td>0.08</td>
<td>−0.517</td>
<td>0.002</td>
<td>0.048</td>
<td>−0.611</td>
<td>−0.424</td>
<td>−10.830</td>
<td>0.0000</td>
</tr>
<tr>
<td>Memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal memory</td>
<td>35.17</td>
<td>51.79</td>
<td>24</td>
<td>27.79</td>
<td>385.19</td>
<td>0.07</td>
<td>−0.332</td>
<td>0.006</td>
<td>0.076</td>
<td>−0.480</td>
<td>−0.184</td>
<td>−4.393</td>
<td>0.0000</td>
</tr>
<tr>
<td>Nonverbal memory</td>
<td>51.00</td>
<td>80.99</td>
<td>37</td>
<td>43.99</td>
<td>529.75</td>
<td>0.08</td>
<td>−0.761</td>
<td>0.004</td>
<td>0.065</td>
<td>−0.889</td>
<td>−0.634</td>
<td>−11.683</td>
<td>0.0000</td>
</tr>
<tr>
<td>Processing speed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Processing speed</td>
<td>69.83</td>
<td>77.90</td>
<td>53</td>
<td>24.90</td>
<td>660.47</td>
<td>0.04</td>
<td>−0.517</td>
<td>0.002</td>
<td>0.048</td>
<td>−0.611</td>
<td>−0.423</td>
<td>−10.747</td>
<td>0.0000</td>
</tr>
<tr>
<td>Visuospatial abilities</td>
<td>55.76</td>
<td>59.35</td>
<td>41</td>
<td>18.35</td>
<td>639.09</td>
<td>0.03</td>
<td>−0.350</td>
<td>0.002</td>
<td>0.048</td>
<td>−0.545</td>
<td>−0.256</td>
<td>−7.272</td>
<td>0.0000</td>
</tr>
<tr>
<td>Working memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working memory</td>
<td>36.42</td>
<td>47.18</td>
<td>25</td>
<td>22.18</td>
<td>390.25</td>
<td>0.06</td>
<td>−0.343</td>
<td>0.005</td>
<td>0.072</td>
<td>−0.484</td>
<td>−0.202</td>
<td>−4.774</td>
<td>0.0000</td>
</tr>
<tr>
<td>Spatial working memory</td>
<td>15.51</td>
<td>17.18</td>
<td>9</td>
<td>8.18</td>
<td>144.85</td>
<td>0.06</td>
<td>−0.369</td>
<td>0.013</td>
<td>0.114</td>
<td>−0.592</td>
<td>−0.147</td>
<td>−3.254</td>
<td>0.0011</td>
</tr>
</tbody>
</table>

Sig of Q = χ² threshold for significance of Q; Df = degrees of freedom; CI = confidence interval.

The mean effect size for the ROCF recall trials is mediated by executive function-related organizational impairments (Savage et al., 1999). Notably, performance on verbal learning tests is associated with left temporal lobe functioning, whereas complex figure test performance is associated with right temporal lobe functioning (Majdan, Sziklas, & Jones-Gotman, 1996). Both types of tests are also associated with frontal lobe functioning (Buckner, Kelley, & Petersen, 1998), representing the strategic organizational component. However, tangible verbal associations facilitate the organization and coding of word lists on tests such as the CVLT, while complex figural tests lack such cues. Thus, overall, our results are in support of moderate differences between OCD patients and healthy individuals (patients performing more poorly) across executive functioning, sustained attention and processing speed, and smaller differences for verbal memory, working memory and visuospatial abilities. Finally, the nature of the large effect size found for non-verbal memory may be related to executive functioning and less with memory impairments per se.

Response inhibition is yet another construct that has received attention in the OCD literature, with some research suggesting response inhibition impairment is an endophenotypic marker for OCD (Chamberlain et al., 2005). Our results show a medium effect size for response inhibition (d = −.49), however a discrepancy was found between the components comprising this domain (i.e., Go/No-Go/CPT/SST commission errors, and Stroop interference): across 23 studies, a medium weighted mean effect size was found for the Stroop interference (d = −.54), but across 15 studies, a small effect size was found for commission errors across (d = −.33). In addition, while significant heterogeneity was found on both domains, confidence intervals for the latter revealed an upper limit approaching zero (CIU: −.61, CI: −.04). These results are in accord with the low percentage of studies reporting a significantly increased number of commission errors in OCD patients as compared to controls. In fact, these results suggest that a reconsideration of response inhibition deficits as an endophenotype of OCD is in order. In addition, in light of the observable differences found between the effect sizes of Stroop interference and commission errors, it would be useful for future studies to conduct a comparative examination between different tests of response inhibition among OCD samples.

Depending on the particular domain or subdomain, the mean effect sizes across studies indicate that OCD patients’ test performance ranged from approximately one-third to three quarters of a standard deviation worse than healthy individuals. Although the majority of effect sizes were of moderate strength, the common rule of thumb in the field of neuropsychology is that differences of less than two standard deviations on neuropsychological tests—which were developed to identify deficits in individuals with brain injuries—are not clinically meaningful (Lezak, Howieson, Bigler, & Tranel, 2012, p. 167). Given that the average mean effect size in the present meta-analysis was −0.499, the existing
literature indicates that across domains, neuropsychological deficits in OCD patients may be clinically insignificant. This has important implications for the use of neuropsychological functioning as an endophenotype of this disorder. Specifically, our findings suggest that neuropsychological impairments may not be reliable endophenotypic markers of OCD.

Our analyses of moderators of effect size across studies yielded no reliable moderator variables. Of particular interest was whether OCD symptom severity would moderate differences between OCD and control groups; however, neither OCD nor depression symptom severity emerged as moderators. Our meta-analytic data do not allow for inferences about causal relationships. Yet if such a relationship exists, it might be complex, bidirectional, and difficult to detect across a methodologically diverse literature. Some authors have proposed that neuropsychological impairments are inherent trait properties of OCD (Bannon, González, Croft, & Boyce, 2006; Rao et al., 2008), whereas others assert that they are state dependent, or an epiphenomenon of OCD symptoms (Abramovitch, Dar, Hermesh, & Schweiger, 2012; Moritz, Hottenrott, Jelinek, Brooks, & Scheurich, 2012). In support of the latter view, a number of studies have demonstrated improvement in neuropsychological functioning following successful treatment of OCD that did not target neuropsychological functions (Kuelz et al., 2006; Moritz, Kloss, Katenkamp, Birken, & Hand, 1999; Nakao et al., 2003).

Future studies might examine this association more carefully and in a longitudinal manner, for example by retesting patients over time, considering possible nonlinear associations, including more ecologically valid tests, using different symptom severity scales, and focusing on more specific clinical factors (e.g., OCD symptom dimensions). Given that approximately 90% of OCD patients are diagnosed with at least one additional psychiatric disorder, and often with multiple comorbidities (Ruscio et al., 2010), the question of whether neuropsychological impairments are associated with the pure phenotype of OCD, as opposed to comorbidity, is also deserving of further research attention.

A limitation of the existing OCD neuropsychology research literature that came to light while reviewing studies for this meta-analysis was that, with very few exceptions (Cha et al., 2008; Hashimoto et al., 2011; Nedeljkovic et al., 2009), it has largely regarded OCD as a homogeneous condition and included only global measures of symptom severity (e.g., the Yale–Brown Obsessive Compulsive Scale). That is, whereas the heterogeneous and dimensional structure of OCD is well established (Abramowitz et al., 2010; McKay et al., 2004), the vast majority of studies have made no attempts to examine possible associations between test performance and specific OCD symptom presentations (e.g., contamination vs. violent obsessions), nor assess the types of obsessions and compulsions present in their samples. As a result, we were unable to meta-analytically examine the extent to which differences in the makeup of OCD samples moderated neuropsychological performance. The use of heterogeneous OCD samples might also have suppressed clinically meaningful findings from coming to light. That is, if patients with some manifestations of OCD (e.g., contamination) have much poorer neuropsychological performance than those with other types of symptoms (e.g., sexual obsessions), the presence of both types of patients in a single group would lead to smaller effect sizes than if a homogeneous group of patients with sexual obsessions was studied. In light of studies suggesting that different manifestations of OCD are associated with different neuropsychological and neurobiological processes and pathways (Mataix-Cols, Alonso, Pifarre, Menchon, & Vallejo, 2002; Mataix-Cols et al., 2004) we suggest that future studies either include OCD samples with more homogeneous presentations (e.g., only those with contamination obsessions), or report the makeup of their OCD samples in terms of the types of obsessions and compulsions present. We also recommend that future neuropsychological investigations carefully examine (and control for) the effects of neuroleptic medications and their combination with SSRIs in OCD. More research is also needed to determine the role of executive functioning/organizational strategies in non-verbal memory tasks. Given that the vast majority of studies utilized the ROCF, one suggestion would be to examine this construct using different non-verbal memory tasks. These include tasks that may not require extensive organizational abilities. Of note, a specific aspect of the ROCF (where participants are not informed of the requirement to memorize the figure) may need to be taken into account, especially in OCD patients. Finally, future studies are advised to pay careful attention to comorbidities and their relative weight in terms of distress and functional impairments as well as to verify that OCD is indeed the primary disorder.

Although the current meta-analysis quantifies the results from comparisons between OCD patients and healthy individuals, the question remains as to whether there are meaningful neuropsychological differences between OCD patients and other psychiatric groups. Indeed, we came across some studies that included comparisons between OCD patients and other psychiatric disorders, yet we did not include these comparisons in the present review because there were too few with any given condition from which to draw meaningful conclusions. For this reason, and especially given the lack of a significant moderating effect of OCD symptom severity, we are limited from drawing conclusions about the specificity of our results. That is, because of the lack of psychiatric control groups, it is not possible to determine whether the differences between OCD patients and healthy controls we have quantified here are particular to OCD, or associated more broadly with anxiety-related symptoms or psychopathology more generally. Thus, we encourage investigators to include psychiatric control groups (e.g., patients with other anxiety or OCD-related problems) in future neuropsychological studies.

Finally, it has been suggested that hoarding symptoms in OCD and hoarding disorder, may be associated with different neuropsychological deficits than OCD without hoarding (Tolin, Villavicencio, Umbach, & Kurtz, 2011). In an attempt to explore hoarding as a potential moderator, for each included study, we recorded whether each study included or excluded hoarders, or did not screen/report. Unfortunately, due to the small number of studies that included hoarders, and the large number of studies that did not address this issue, as well as due to the inconsistent use of measures for screening (e.g., the obsessive-compulsive inventory versus the proposed DSM-5 criteria), we were unable to conduct this analysis. However, this issue deserves further investigation, especially in light of hoarding disorder’s new status as a separate disorder in the DSM-5.

5. Conclusion

In conclusion, although the present meta-analytic examination offers evidence for poorer performance in OCD relative to healthy individuals across the majority of neuropsychological domains, these differences may not be clinically meaningful. Moreover, the specificity of the statistically significant differences to particular presentations of OCD, or even to OCD per se, remains unclear. Notably, we cannot rule out the possibility that clinically meaningful effects would emerge when examining specific presentations of OCD. Furthermore, it is not clear whether these observed differences play any role in the development of OCD, or whether they are epiphenomena of having OCD. From a theoretical perspective, our results are to some extent conceptually in line with current neuropsychological models of OCD, yet a large number of psychiatric disorders are also associated with frontostriatral pathophysiology (e.g., schizophrenia, depression, bipolar disorder, ADHD) and nearly every psychiatric disorder is associated with executive function impairments. Taken together with the small to moderate levels of effect sizes seen across domains, and the scarce evidence for clinically significant familial neuropsychological impairments, neuropsychological factors may not be sound candidates for OCD endophenotypes. Thus, we suggest that future neuropsychological research in OCD give emphasis to the interaction between neuropsychological and specific clinical aspects of the disorders using theoretically driven hypothesis, with a translational goal in mind.
Acknowledgments

The authors would like to thank Drs. Katherine Burdick, Antje Hunger, Benna Burin Kivirik, Kathrin Koch, Himani Kashyap, Leonard Lennartz, Maja Nedeljkovic, Nick Ramsey, Janardhan Reddy, Cinto Segalas, David Tolim and Nic van der Wee for providing us with valuable information, clarifications and unpublished data/analyses. Additionally, we would like to thank Dr. George Demakis for his invaluable comments on a previous draft of this manuscript.

References


