Neuropsychological impairments in panic disorder: a systematic review

Neuropsychological impairments in panic disorder: a systematic review

Kate O'Sullivan [a b] *, Emily F. Newman [b]

[a]Department of Clinical Psychology, Tayside NHS, Dundee, UK

[b]Clinical and Health Psychology Section, School of Health in Social Science, University of Edinburgh, UK

* Corresponding author. Glasgow Addiction Service NHS GG&C, Parkhead Alcohol Unit, Parkhead Hospital, 81 Salamanca Street, Parkhead, Glasgow, G31 5ES. Tel.: +44 141 211 8331; fax: +44 141 2118370

Email address: Kate.O'Sullivan@ggc.scot.nhs.uk (K. O'Sullivan).

Abstract

Background: There is a growing body of literature investigating the neuropsychological profile of Panic Disorder (PD), some of which suggests potential cognitive dysfunction. This paper systematically reviews the existing literature on neuropsychological performance in PD.

Method: PsycINFO, EMBASE, MEDLINE and PsycARTICLES databases were searched to identify articles reporting on neuropsychological function in PD published in English during the time period 1980 to March 2012. 14 studies were identified.

Results: There was limited support for impairment in short term memory among individuals with PD, although this was not found across all studies. Overall, the reviewed studies did not support the presence of impairment in other areas of cognitive functioning, including executive function, long term memory, visuospatial or perceptual abilities and working memory.

Limitations: Studies with samples of fewer than 15 participants per group were excluded from this review. A limited amount of research has been published on this topic and small sample sizes (under 25 per group) have been used by many studies. Therefore, the current review is based on a small number of studies with limited power.

Conclusions: There is limited evidence of specific neuropsychological impairments in participants with PD. Impairments in short term memory warrant further investigation to establish their relevance to clinical practice. Larger sample sizes and appropriate statistical adjustment for multiple comparisons in future studies is highly recommended.

Keywords: Panic Disorder, Cognitive impairment, Neuropsychology, Review

1. Introduction

Panic disorder (PD) is a disabling mental health problem characterised by unexpected, recurrent panic attacks, fear about the implications of attacks and modifications of behaviour as a result of the attacks (American Psychiatric Association, 2000). PD can occur with or without agoraphobia and is associated with high levels of psychiatric comorbidity and severe role impairment (Baillie and Rapee, 2005; Kessler et al., 2006). A number of recent studies have focused on the neurobiology of common psychiatric disorders, including anxiety disorders, and underlying cognitive impairments associated with them (Millan et al., 2012). Impairments in neuropsychological functioning are of interest as they may have implications for treatment outcomes, as has been seen in schizophrenia and anorexia nervosa (Cavedini et al., 2006; Tabarés-Seisdedos et al., 2008). These impairments may also act as measurable symptoms of underlying neurobiological dysfunction. Several studies have found structural brain abnormalities in patients with anxiety disorders, including patients with PD (Mataix-Cols and van den Heuvel, 2006; Phan et al., 2009; Szeszko et al., 2005; van den Heuvel et al., 2005). Patterns of impairments in executive function have been reported in a number of recent reviews of neuropsychological performance in OCD (Martinez-Gonzalez and Piqueras-Rodriguez, 2008; Menzies et al., 2008; Olley et al., 2007). Executive function impairments have also been implicated in Post Traumatic Stress Disorder (PTSD; Aupperle et al., 2012); however PD has been less well researched.

In PD, imaging studies have indicated abnormalities in specific brain regions compared to controls, including different metabolic activity in the hippocampal and parahippocampal areas (Bisaga et al., 1998) and abnormalities in the temporal lobe structures (Vythilingam et al., 2000). Brain abnormalities such as these may lead to learning and memory deficits, if present in panic disordered individuals. However Reiman and colleagues have noted that in their work, similar regional blood flow patterns have been seen in panic

disordered patients as in healthy controls with anticipatory anxiety. It remains unclear whether abnormalities seen relate to structural differences or effects of state or trait anxiety (Reiman et al., 1989a; Reiman et al., 1989b).

It has been suggested that state and trait anxiety may influence performance on neuropsychological testing, confounding neuropsychological testing in anxiety disorders (Orsillo and McCaffrey, 1992). Recent research has investigated the potential effects of state or trait anxiety on neuropsychological test performance in a number of different populations including a mixed psychiatric sample (O'Jile et al., 2005; Smitherman et al., 2007). Neither state nor trait anxiety were found to have a significant effect on executive function, (Horowitz and McCaffrey, 2008; Smitherman et al., 2007; Waldenstein et al., 1997) nor on memory and verbal learning (Kizilibash et al., 2002; O'Jile et al., 2005; Waldenstein et al., 1997), once age, gender and IQ were controlled for. Performance on an attention test was found to be unaffected by state or trait anxiety in healthy males (Waldenstein et al., 1997). Gass and Curiel (2011) reported that test anxiety did not impair performance on the coding or block design subtests of the WAIS-III in their sample of veterans, which are recognised as requiring sustained attention and concentration, often thought to be affected by anxiety. State anxiety, was found to significantly affect performance by veterans on working memory tasks, however this effect was reduced when education was controlled for (Gass and Curiel, 2011). An effect of anxiety on working memory was not seen by Waldenstien et al.(1997) using the digit span task on healthy male participants.

The co-occurrence of anxiety and depression may influence neuropsychological performance. Although anxiety alone was not found to affect memory performance, when anxiety and depression were reported together by veteran participants, a significant negative effect on memory performance was seen (Kizilibash et al., 2002). Overall, current research

suggests that state or trait anxiety alone in the absence of depression may not have a large impact on neuropsychological test performance in PD.

Individual studies have found associations between PD and impairment in a number of cognitive areas, including executive function and episodic memory (Airaksinen et al., 2005). However conflicting results have been produced, with some studies supporting memory impairment in people with PD (Asmundson et al., 1994; Lucas et al., 1991) and others reporting no memory problems of any kind (Gladsjo et al., 1998). No review was found of neuropsychological performance in PD. A greater understanding of cognitive deficits in PD afforded by the collation of existing research evidence may have implications for its treatment and measurement of treatment outcomes. This paper aimed to provide a systematic review of neuropsychological performance in PD compared to healthy control (HC) participants.

2. Methods

2.1. Search Strategy

Relevant studies published between January 1980 and March 2012 were identified by systematic searches of PsycInfo, Embase, PsycArticles and Medline databases. Articles reporting neuropsychological performance of all anxiety disorder groups were initially identified in order to avoid missing any relevant papers, as a preliminary search of the literature indicated that PD groups were often used as comparison groups in studies which focused on other anxiety disorders, such as OCD (e.g. Purcell et al., 1998). Keywords for the search were “neurocognition”, “attention”, “executive function”, “leaning”, “memory”, “inhibition” AND “neuropsychological tests” AND “anxiety”, “OCD”, “PTSD”. Terms were adapted and ‘exploded’ in keeping with subject headings for each database (i.e. all of the

lower ‘branches’ of the keyword were automatically included in the search. The lower branch terms are defined by each database). Associated terms “task performance”, “central coherence”, “association learning”, “cognitive defect”, “short term memory”, “autobiographical memory”, “decision making”, “motor control”, “social anxiety”, “social phobia”, “panic”, “phobia”, “obsession” were also included in the search. A total of 3431 papers were originally identified from all 4 databases. Reviews and study protocols were eliminated.

The reference lists of ten papers, identified as appropriate after inclusion and exclusion criteria were applied, were checked for relevant studies, resulting in 3 additional papers. A ‘cited by’ search was conducted using Web of Science (1899-present) resulting in 2 additional papers. Subsequently two studies were discovered to be reporting on the same data (Lautenbacher et al., 2002; Spernal et al., 2003). The paper containing the most data was retained in the review, resulting in a total of 14 papers (Lautenbacher et al., 2002).

2.2. Inclusion and exclusion criteria

Studies were included if they reported on: (1) adults (18-65 years) , (2) diagnosed with current PD according to DSM or ICD criteria, (3) a comparison group of healthy controls free from any current psychiatric disorder (HC) , (4) had ≥ 15 participants in each group, and (5) were published in English. A relevant paper was found during the search but could not be included as it was only available in Spanish (Castillo et al., 2010). Studies on the effect of psychotropic medication or a treatment intervention were excluded. Investigations of cognitive performance in the presence of anxiety provoking words or stimuli were excluded. Studies of neuropsychological performance during brain imaging or brain activity recording were also excluded. The paper selection procedure is described in Figure 1.

<Figure 1 about here>

2.3. Data Extraction

Data were extracted from each paper by the first author (KO'S) according to a structured pro-forma, specifically developed by the first author for this review, covering key study characteristics. These key characteristics and the data extracted using the pro-forma are detailed in Table 1.

<Table 1 about here>

2.4. Assessment of methodological quality

To rate the methodological quality of included studies, criteria were developed by the first author, drawing from the Cochrane Handbook for Systematic Reviews guidance on assessing risk of bias (Higgins and Altman, 2008) and the Centre for Reviews and Dissemination (CRD) guidance on conducting quality assessment (CRD, 2008). A checklist of 8 quality criteria was developed a priori. The quality criteria were as follows: 1. Eligibility criteria for the PD group were appropriate and clearly specified. This was considered to be adequately addressed if the inclusion criteria were not outlined clearly but could be ascertained from the details provided in the paper. 2. The comparison group was matched to the PD group on age, gender and either IQ or education. This was considered to be adequately addressed if the group was matched on some but not all of these. 3. Diagnosis of PD, or absence of diagnosis for HCs, was ascertained using DSM/ICD criteria using a structured interview by a clinician. This was considered to be adequately addressed if diagnosis was determined using a questionnaire based on DSM/ICD criteria. 4. The

neuropsychological measures used were robust for this population with reliability and validity specified in the paper or easily obtainable using Lezak et al.'s (2004) reference text, or another reference provided in the paper. This was considered to be adequately addressed if tests were well described but reliability and validity not found or tests were not the most valid for this population. 5. The sample size of both the PD and HC groups was 25 or more as this was calculated to be the minimum number needed to detect a large effect size. This was considered to be adequately addressed if the sample size in each group was between 15 and 25. 6. The levels of uptake from those invited to participate were reported in the paper and the effects of uptake levels were considered. This was considered to be adequately addressed if the levels of uptake were described. 7. The means, standard deviations and confidence intervals were reported. This was considered to be adequately addressed if enough statistical outputs were reported to facilitate comparison. 8. Appropriate statistics were used including compensations for multiple comparisons with a new alpha level clearly stated. This was considered to be adequately addressed if no corrections were made for multiple comparisons but otherwise appropriate statistics were used. The quality ratings for the included studies and criteria are listed in Table 2.

For each criterion, included studies were assigned one of four outcome ratings: 'well covered' (2 points); 'adequately addressed' (1 point); 'poorly addressed' or 'not addressed' (both 0 points). Two additional raters independently reviewed four studies each. Exact agreement was reached on 87.5% (Cohen's kappa = .79) and 84.4% (Cohen's kappa = .74) of the ratings respectively. A difference of one point occurred on 12.5% of the items and by 2 points on 1.5% items. Differences in rating of criteria were discussed and amended.

<Table 2 about here>

2.5. Calculation of effect sizes

For each measure, the standardised mean difference effect size using pooled standard deviations was calculated (Cohen's *d*), using the formula

$$d = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}}} = \frac{\overline{X}_1 - \overline{X}_2}{s_{pooled}}$$

and adjusted for small sample size using the formula below (Lipsey and Wilson, 2001).

$$d' = d\left[1 - \frac{3}{4N - 9}\right]$$

Weighted mean effect sizes were calculated for each neuropsychological domain and an overall study mean effect size was calculated following the procedure of Lipsey and Wilson (2001). Whenever PD participants performed poorer than HCs, between group differences were reported as positive effect sizes. Cohen (1988) defines a small effect size as $d \geq .2$, a medium effect size as $d \geq .5$, and a large effect size as $d \geq .8$. In the context of neuropsychological research, Bezeau and Graves (2011) reported that large effect sizes are typically found, with an average effect size of .88 found in the 66 studies they reviewed. They suggested that effect sizes of .8 or larger are preferable in order to ensure less overlap between group distributions and therefore to be clinically useful in distinguishing between groups. Where PD groups were split into two groups in Kaplan et al.'s (2006) paper, the means and standard deviations were combined to represent all of the PD participants. The formulas used for combining this data were taken from the Cochrane Handbook Table 7.7.a (Higgins and Deeks, 2011).

3. Results

3.1. Search results and characteristics of studies

Fourteen studies were identified comparing a PD group with a HC group. The 14 studies involved 439 patients with PD in total and 510 HCs (see Table 1 for details). The median sample size was 23.5 for PD patients (range 15 - 93) and 27.5 for HCs (range 15 - 175). Three studies reported on the presence or absence of agoraphobia in their PD sample but 11 did not. Both Boldrini et al. (2005) and Gorini et al. (2010) reported all PD participants to have PD with agoraphobia. Twenty six of the 28 PD participants in Galderisi et al.'s (2008) study had PD with agoraphobia.

3.2. Neuropsychological Variables

As many neuropsychological measures can be said to assess a number of cognitive functions, Lezak et al.'s (2004) categorisation of neuropsychological assessments has been broadly followed when tabulating and discussing the measures used in the reviewed studies.

3.2.1. Verbal Memory

The California Verbal Learning Test (CVLT) (Delis et al., 1987), two Selective Reminding (SR) tasks (Buschke and Fuld, 1974) and 4 recall of word lists tasks were included in the analysis. The Warrington Recognition Memory task (Words) (Warrington, 1984) which tests the participant's ability to recognise a previously presented word when paired with a distractor was also included, as well as the paired associates and logical memory subscales of the Wechsler Memory Scale (WMS; Wechsler, 1987).

3.2.2. Visual Memory

The Benton Visual Retention test (BVRT; Benton, 1945), Rey-Osterrieth Complex Figure test recall measure (RCFT; Rey, 1941), Visual Selective Reminding Test and Continuous Visual Memory Test (CVMT; Trahan and Larrabee, 1989) were included in this domain. Visual memory was also assessed using a non-standardised task in which an array of numbers was visually presented followed by immediate recall (Gordeev, 2008) and the visual reproduction subscale of the WMS (Wechsler, 1987). Visual recognition memory was investigated using 3 subtests of the computerised Cambridge Neuropsychological Test Automated Battery (CANTAB; Cambridge Cognition, Cambridge, UK) - Spatial Recognition, Pattern recognition and delayed match to sample and the Warrington Recognition Memory test (Faces) (Warrington, 1984).

3.2.3. Working Memory

Working memory abilities were assessed using span and superspan tasks. The digit span task (Wechsler, 1981), the Hebb Digit Recurring task, the Corsi Block Tapping Task (CBTT; Berch, 1998) and the spatial span subtask of the computerised CANTAB (Cambridge Cognition, Cambridge, UK) were included in the analysis. These tasks are also influenced by attention and verbal and visual memory abilities.

3.2.4. Attention

Attention was assessed using the Digit Symbol task from the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981), a Continuous Performance Test (Rosvold, Mirsky, Sarason, Bransome, Beck, 1956), time to complete the Trail Making Test part A (TMT A; Reitan, 1958), the Digit Vigilance task (Lewis and Rennick, 1979), the Rapid Visual Information Processing subtask of the CANTAB (Cambridge Cognition, Cambridge, UK) and the mental control task of the WMS (Wechsler, 1987).

3.2.5. Perception

Tests included in this section were classed primarily as tests of perception by Lezak et al. (2004) as they measure visual field perception, visual searching and facial recognition but also include selected and divided attention. The Benton Facial Recognition test (BFRT; Benton 1983) was used to test facial recognition in the absense of a memory condition. The Visual Field Neglect task from the 'Testbatterie zur Aufmerksamkeitsprufung' (TAP; Zimmermann and Finn, 1993) was used to test divided attention and the Signal Detection test from the Weiner Testsystem (Schuhfried, 1999), however information on psychometric properties could not be found for either of these measures. The Munsterberg test was also used to test selective attention. A Schulte tables task of sustained attention and a digit cancellation test were also used. The Munsterberg test requires participants to find words in a random set of letters within a limited time. Psychometric properties for the Munsterberg task and Schulte tables were not found in a search of the literature. Some literature was found indicating that Schulte tables do not have well established psychometric properties (Ennok, 2010).

3.2.6. Visuospatial Ability

Visuospatial ability was assessed using the Block Design task from the WAIS-R (Wechsler, 1981), the copy trial of the RCFT (Rey, 1941), and an unstandardised Virtual Water Maze analog testing spatial orientation and learning (Jacobs et al., 1997). The ability to perform spatial rotations mentally was investigated using the mental rotation test (MRT; Vandenberg and Kuse, 1978).

3.2.7. Executive function

The category of executive functioning was expanded from that of Lezak et al. (2004) in line with Burgess (2003) to incorporate tests of inhibition, coordinated dual tasks (e.g. the Trail Making Task) and verbal fluency. In the context of the reviewed studies, such measures were used with the purpose of assessing executive functioning. Results were considered under the headings of planning and organising, set shifting, verbal fluency and decision making. The Wisconsin Card Sort Task (WCST; Heaton et al., 1993), time to complete the Trail Making Test part B (TMT B; Reitan, 1958), the Intradimensional-Extradimensional shift and the Spatial Working Memory subtasks of the computerised CANTAB (Cambridge Cognition, Cambridge, UK) were used to assess set shifting. The Tower of London subtask of the computerised CANTAB (Cambridge Cognition, Cambridge, UK) and the organisation trial of the RCFT (Rey, 1941) were used to assess planning and organising ability. Verbal fluency was assessed using the FAS letter fluency task (Benton and Hamsher, 1983), a category fluency task (Goodglass and Kaplan, 1983) and the Benton Controlled Oral Word Association Test (COWAT; Benton, 1989). Decision Making ability was examined by Cavedini et al. (2002) using the Iowa Gambling Task (IGT; Bechara et al., 1994), by Kaplan et al. (2006) using the Cambridge Gambling task (Rogers et al., 1999) and by Ludewig et al. (2003) using a two-choice prediction task (Paulus, 1997).

3.3.Memory

3.3.1. Verbal Memory

Short and long term verbal memory were investigated in eight studies using ten measures (see Table 3). Two studies investigated short term verbal memory using non-standardised short term recall tasks. These indicated worse performance in PD participants compared to HC (Airaksinen et al., 2005; Gordeev, 2008). Using the CVLT (Delis et al.,

1987) Asmundson et al. (1994) also reported worse performance in PD participants compared to HCs. Three studies reported on learning in short term memory trials 1-5 of the CVLT, none finding significantly poorer performance of PD participants compared to HCs (Deckersbach et al., 2011; Gladsjo et al., 1998; Asmundson et al., 1994). No differences between PD and HC groups were reported by Lucas et al. (1991) or Deckersbach et al. (2011) in measures of immediate recall of stories or paired words, taken from the WMS (Wechsler, 1987). Mixed results were reported in relation to cued short term memory, with group differences seen on a non-standardised task (Airaksinen et al., 2005) but not on the cued recall subscale of the CVLT (Asmundson et al., 1994). Overall short term verbal memory was measured by six studies and impairment of PD patients compared to HC was reported in three studies. Three of the five tasks used by the reviewed studies found a significant difference in performance between PD and HC groups, however two of these tasks were non-standardised and differences seen using the CVLT were not reported by two other studies using this task.

Delayed verbal memory was investigated by seven studies. No differences were found between PD and HC groups on delayed verbal memory using the CVLT (Asmundson et al., 1994; Deckersbach et al., 2011; Gladsjo et al., 1998), the Auditory Verbal Learning Test (AVLT) (Galderisi et al., 2008), or the paired associates and logical memory subscales of the WMS (Lucas et al., 1991). Using a selective reminding procedure (Fletcher, 1985), conflicting results were produced by two studies with Lucas et al. (1991) observing poorer performance in long term verbal memory in a PD sample compared to HC while Boldrini et al. (2005) did not. Using the Recognition Memory Test (RMT; Warrington, 1984) no difference in recognition memory after a delay was found between PD and HC groups (Gladsjo et al., 1998). In summary, delayed verbal memory was investigated in six studies with only one studies reporting significantly poorer performance in the PD group. Of the six

tasks used only the selective reminding test found significant differences between groups and this was not replicated by another study using the same task. Effect sizes for verbal memory ranged from small and negative, -0.27, to large and positive, 10.27. Although few significant group differences were found, all but four of thirty effect sizes derived from the verbal memory data indicated a trend towards worse visual memory performance in PD participants compared to HC participants (Table 3).

< Table 3 about here>

3.3.2. Visual Memory

Both short and long term visual memory in people with PD was investigated by eight studies using ten measures (see Table 4). Short term visual memory was investigated in two studies, one using the RCFT (Deckersbach et al., 2011) and one using immediate recall of numbers (Gordeev, 2008). Both these tasks were associated with an impaired performance in people with PD compared to HC.

Seven studies reported on measures of long term visual memory, including measures of retention and recognition. Three studies used the BVRT (Benton, 1945) with two of these reporting significantly worse long term visual memory in people with PD relative to HC (Asmundson et al., 1994; Deckersbach et al., 2011; Lucas et al., 1991). A selective reminding procedure (Fletcher, 1985) was used in two studies producing mixed results, as Lucas et al. (1991) reported poor performance in people with PD but this was not replicated by Gladsjo et al. (1998). No differences were reported between PD and HC groups on the visual recognition tasks (Gladsjo et al., 1998; Kaplan et al., 2006; Purcell et al., 1998). People with PD also performed similarly to HC on the RCFT percent recall (Boldrini et al., 2005; Deckersbach et al., 2011) and the CVMT (Gladsjo et al., 1998). The PD group performed worse than HCs on delayed recall of the Visual Reproduction subscale of the

WMS (Lucas et al., 1991). Both studies investigating visual short term memory found worse performance in PD compared to HC groups on the two tasks used. Two of seven studies found differences between PD and HC groups on visual long term memory. Nine tasks were used to investigate visual long term memory, six of these found no differences. Three tasks, the BVRT, the visual reproduction subscale of the WMS and the selective reminding procedure, found poorer performance in the PD group compared to HCs, however for two of these tasks, these findings were not consistent across all studies using them. The effect sizes reported varied widely from medium and negative to large and positive (range -0.64 to 11.53).

In summary, there was some support for significantly different short term memory performance in people with PD in both verbal and visual memory but the reliability and validity of a number of the tasks providing this support is unclear. There was little support for impairment in either verbal or visual long term memory.

<Table 4 about here>

3.4. Working memory, attention and perception

See Table 5 for results relating to working memory, attention, perception and visuospatial ability.

<Table 5 about here>

3.4.1. Working Memory

Working memory in people with PD was explored by seven studies using working memory span tasks. No differences between the performance of people with PD and HC were

reported by four studies using the Digit Span (Wechsler, 1981) task. Two studies using the CANTAB computerised Spatial Span and Spatial Working Memory tasks also found no differences. Boldrini et al. (2005) and Galderisi et al. (2008) employed the CBTT (Berch, 1998) to investigate working memory span while Deckersbach et al. (2011) used the similar Spatial Span Task from the WMS (Wechsler, 1987). Performance was reported as no different to HC on measures of span (Boldrini et al., 2005; Deckersbach et al., 2011), however on the supraspan subscale, Boldrini et al. (2005) reported a poor learning process in spatial working memory in people with PD compared to HC. Galderisi and colleagues (2008) also administered the CBTT and the Hebb Digit Recurring test (HDR) but only reported on the accuracy index of each. People with PD were found to perform better than the HCs in their sample on the CBTT but not the similar HDR (Galderisi et al., 2008). In summary, there was very little support for working memory impairment in PD. Only one study of seven found impaired performance in PD compared to HC and only on one subtest of the CBTT. In this way only one of five tasks used showed a difference between the groups and this difference was not seen on all subtests of the CBTT. The effect sizes reported ranged from medium and negative, -0.51, to large and positive, 3.78.

3.4.2. Attention

Attention was investigated in seven studies using six different tests. No impairments of people with PD were found compared to HC in any study and no clear trend was seen in the direction of the small to medium effect sizes (range -0.48 to 0.65) (Airaksinen et al., 2005; Asmundson et al., 1994; Deckersbach et al., 2011; Galderisi et al., 2008; Gladsjo et al., 1998; Kaplan et al., 2006; Lucas et al., 1991).

3.4.3. Perception

Four studies investigated perception in PD using six tasks. Individuals with PD performed poorly compared to HC on the Munsterberg Test, described as a test of selective attention and on a Schulte tables task of sustained attention (Gordeev, 2008). In contrast with Gordeev (2008), Lautenbacher et al.(2002) did not find overall group differences on selective attention using a computerised signal detection task. However Lautenbacher et al. (2002) reported significant differences on a divided attention task in which individuals with PD demonstrated impaired performance relative to HC. Individuals with PD were found to be as good at identifying faces as HCs using the Benton Facial Recognition test (short form) (Benton, 1983; Boldrini et al., 2005). No difference in ability was seen between the two groups on a digit cancellation test (Asmundson et al., 1994). Overall two of four studies reported some findings of impaired perceptual abilities in people with PD. Three of the six tests used over the four studies found poor performance in the PD group but two of the tests reporting significant results used tasks for which the reliability was uncertain and produced large effect sizes of 21.70 and 29.47 which were inconsistent with the magnitude of effect sizes produced by all other studies. Effect sizes for the perception domain ranged from small and negative, -0.06 to large and positive, 29.47.

3.4.4. Visuospatial Ability

Visuospatial ability was investigated in PD by five studies using four measures. Two studies found no group differences in visuospatial ability using the RCFT copy score (Boldrini et al., 2005; Deckersbach et al., 2011). Poor visuospatial ability of people with PD was seen on the Block Design task by Asmundson et al. (1994) but was not replicated (Gladsjo et al., 1998). The PD group performed as well as HCs on the Mental Rotation test. A virtual water maze task was used by Gorini et al. (2010) to investigate spatial orientation

and learning in people with PD. Learning the position of the platform over trials was impaired in the PD group compared to HCs, however this was a novel task which has not been standardised and high variability of performance was noted in both groups. In summary, two studies out of five provided support for visuospatial impairment in people with PD. This support was provided by two of the four tests of visuospatial ability used, the Block Design task (WAIS-R) and the virtual water maze. However the Block Design task did not produce consistent findings with impairment being found in only one of the two studies where it was used. Effect sizes ranged from small and negative, -0.32 to large and positive 2.32.

3.5. Executive Function

Executive function was assessed by nine studies using nine measures (see Table 6).

< Table 6 about here>

3.5.1. Planning and Organising

People with PD demonstrated impaired performance relative to HC in one of two studies on the organisation score of the RCFT (Deckersbach et al., 2011). Purcell et al. (1998) administered the Tower of London task (CANTAB; Cambridge Cognition, Cambridge, UK) and found comparable performance between the two groups. Only two studies examined this area with one reporting a significant impairment in PD relative to HC. Out of the two tasks used, only one, the RCFT, found impaired performance relative to HC. Effect sizes ranged from small, 0.16 to large, 1.42, although both suggested poorer performance in PD participants.

3.5.2. Set Shifting

The set shifting performance of people with PD in comparison to HC was examined by six studies using three tasks. Individuals with PD performed as well as HC on the WCST (Boldrini et al., 2005; Heaton et al., 1993). There were also no group differences found in both studies using the CANTAB Intradimensional and Extradimensional Shifting Task and the Spatial Working Memory task (Kaplan et al., 2006; Purcell et al., 1998). Three studies reported on the TMT B (Reitan, 1958), with only one (Airaksinen et al., 2005) reporting significantly slower times in people with PD compared to HC. Overall one of six studies (Airaksinen et al., 2005) and one of three tasks (TMT B) found impaired set shifting of PD in comparison to HC. This result was not replicated by the other two studies using the TMT B. Although results were not significant, ten of fourteen effect sizes suggested worse performance by PD participants. Effect sizes ranged from small and negative, -0.34 to large and positive, 2.77.

3.5.3. Verbal Fluency

Five studies investigated letter fluency abilities, finding individuals with PD produced as many words as HCs. One of these also investigated category fluency (Benton, 1989; Gladsjo et al., 1998) and again found no group differences in number of words produced. Effect sizes ranged from small and negative, -0.08 to large and positive, 2.38.

3.5.4. Decision Making

Performance of the PD group was not significantly different to the control group for any task in any of the studies, although PD participants showed increased sensitivity to error, being more likely to search for a better responding strategy even at low error rates in

Ludewig et al.'s (2003) study using a two-choice prediction task (Paulus, 1997). Kaplan et al. (2006) reported that within their sample comorbid major depressive disorder (MDD) seemed to slow decision making. Effect sizes could not be calculated from the available data. In summary, the evidence reviewed does not support significantly different executive function performance in PD.

3.6. Overall mean effect sizes

For each domain, the range, median, weighted mean effect size, and 95% confidence interval for the weighted mean effect size were calculated. These are provided, along with the Q statistic for heterogeneity and the critical Q value, in Table 7.

< Table 7 about here>

3.7. Comparison of PD and other disorders in the reviewed studies

Within the reviewed papers participants with PD were compared to participants with Social Phobia, Specific Phobia, Generalised Anxiety Disorder (GAD), OCD and Unipolar Depression (UD). They were also compared to people with both PD and UD in Kaplan et al.'s (2006) paper. Comparisons with other anxiety disorders are shown in Tables 8 and 9, and comparisons with depression shown in Table 10.

3.7.1. Comparison of PD and OCD

Patients with PD were compared to those with OCD in four of the reviewed studies (Airaksinen et al., 2005; Boldrini et al., 2005; Cavedini et al., 2002; Purcell et al., 1998). Only one study of three testing verbal memory found any significant group differences, reporting better long term memory performance by PD participants compared to OCD

participants on two of the four reported outputs of the Selective Reminding test, although only one of these remained significant after Bonferroni corrections (Boldrini et al., 2005). Three studies investigated visual memory with most tests indicating better performance by PD participants compared to OCD, although only one study reported significant results using the CANTAB tasks of Pattern Recognition memory and spatial Recognition Memory. Overall, effect sizes for memory performance suggest better performance of PD participants than OCD participants (effect sizes ranged from large and negative, -4.09 to small and positive, 0.3) (see Table 8).

PD was compared to OCD by two studies using three measures of working memory (Table 9). Boldrini et al. (2005) found PD participants to perform significantly worse than OCD participants on the Digit Span task but significantly better on the Corsi Block Tapping test (CBTT) spatial span task. Purcell et al. (1998) did not find significantly different performance between the two groups on the spatial span measure of a computerised version of the CBTT. Effect sizes for working memory ranged from -2.45 to 1.62. Attention performance was compared between PD and OCD participants in one study using the TMT A and no significant difference was found (Airaksinen et al., 2005). One study compared performance in the area of perception, using the Facial Recognition test, finding no difference in recognition of faces (Boldrini et al., 2005). Using the Rey-Osterrieth Complex Figure test copy task, one study compared visuospatial ability of the PD and OCD groups, reporting significantly better performance in the PD group than the OCD group (Boldrini et al., 2005). Set shifting ability was compared in three studies using four tasks. No significant difference in set shifting ability between PD and OCD participants was reported by Airaksinen et al. (2005) on the TMT B and mixed results were reported by Boldrini et al. (2005), as performance differed across different measures within the Wisconsin Card Sort Test. Overall half the WCST measures indicated poorer performance by PD participants, two of those

differences being significant. However PD participants also performed significantly better on one WCST measure. Purcell et al. (1998) reported better set shifting among PD participants compared to OCD participants on the Intradimensional-Extradimensional shift task and the Spatial working memory task of the CANTAB, although these results were non-significant. Both studies comparing PD and OCD participants on verbal fluency found no significant differences using the FAS, although both indicated better performance by PD participants. Planning and organising of PD and OCD participants was investigated by one study using the Tower of London task, finding significantly better performance among PD participants than OCD participants (Purcell et al. (1998). In a comparison of PD participants and OCD participants on a task of decision making, Cavedini et al. (2002) reported significantly better performance by PD participants; however there was insufficient data to calculate an effect size. Effect sizes for executive function ranged from large and negative, -3.14 to small and positive, 0.95.

< Tables 8 and 9 about here>

3.7.2. Comparison of PD and Social Phobia

Two studies compared neuropsychological performance of PD participants and participants with Social Phobia (Asmundson et al., 1994; Airaksinen et al., 2005). No significant differences were found between the two groups on verbal memory performance. Overall, effect sizes suggested worse performance by PD compared to participants with Social Phobia (range -0.44 to 0.72). Similarly, the one task used to compare visual memory performance found a non-significant difference between the two groups although a medium effect size (0.46) was found suggesting worse visual memory performance by PD participants (Asmundson et al., 1994). Both studies investigated attention using the TMT A, with one

study suggesting better performance by PD participants and the other suggesting the opposite (range -0.2 to 0.09). Neither result was significant. One study investigated perception using the Digit Cancellation Task and reported significantly better performance by the PD group than the Social Phobia group (medium effect size, -0.69). No significant difference in performance between the two groups was found in visuospatial ability (Asmundson et al., 1994). Both studies investigated set shifting performance using the TMT B. Each study reported non-significant small effect sizes of similar magnitude suggesting opposite patterns of performance (range of -0.35 to 0.31). PD participants performed better than participants with Social Phobia on the FAS (small effect size, -0.3) although this was a non significant result (Airaksinen et al., 2005).

3.7.3. Comparison of PD with GAD

Only one study compared PD participants to a small group of participants with GAD (Airaksinen et al., 2005). They found no difference between groups on a verbal memory task of cued recall and a small non-significant difference (small effect size, 0.11) on a task of verbal free recall. No significant difference was found between the two groups on attention performance although GAD participants did not perform as well as PD participants (medium effect size, -0.47). Executive function performance was compared using the TMT B and the FAS. No significant differences were found on these tasks (effect sizes were small, ranging from 0.08 to 0.22).

3.7.4. Comparison of PD and Specific Phobia

Airaksinen et al. (2005) compared PD participants to patients with Specific Phobia. Using a verbal memory task of free and cued recall, PD participants did not perform significantly differently to participants diagnosed with Specific Phobia (small effect sizes,

0.21 to 0.37). Attention performance was not found to be different between the two groups (small effect size, 0.28). Executive function performance was not significantly different between groups on the TMT B (small effect size, 0.17) or the FAS (small effect size, 0.18).

3.7.5. Comparison of PD with Depression

Kaplan et al. (2006) compared two small groups (n=11) of PD patients, one group diagnosed with PD alone and one group diagnosed with PD and comorbid Unipolar Depression (UD). The paper provided performance data for both these groups, though the authors compared their performance with matched healthy controls. From the information provided, a comparison of visual memory, working memory, attention and set shifting task scores revealed no significant differences between the groups in any domain and effect sizes were small, ranging from -0.03 to 0.23.

Three studies compared participants with PD to participants with diagnosed UD or Major Depressive Disorder (MDD) (Lautenbacher et al., 2001; Ludewig et al., 2003; Purcell et al., 1998). Purcell et al. (1998) compared the groups on three tasks of visual memory from the CANTAB computerised battery, reporting no significant differences and small effect sizes (range -0.03 to 0.27). Working memory performance was found to be significantly poorer in PD than UD participants using a computerised spatial span task (medium effect size 0.66) (Purcell et al., 1998). On two measures of perception no significant differences were found on measures of divided or selective attention (range of small effect sizes -0.07 to 0.3) (Lautenbacher et al., 2001). Executive function was investigated using task of set shifting, planning and organising and decision making. Set shifting performance was compared between PD and UD groups using the CANTAB Intradimensional-Extradimensional shift task (IES). Effect sizes suggested better set shifting performance by PD participants, one of these differences reached statistical significance (range from medium, -0.73, to small, -0.11)

(Purcell et al., 1998). Very similar performance between groups was reported on a Tower of London task investigating planning and organising ability (effect size 0.10) (Purcell et al., 1998). On a two choice prediction task, PD participants were reported to perform similarly to UD participants in terms of response bias; however PD participants generated less predictable response sequences than UD at low error rates (Ludewig et al., 2003). Effect sizes could not be calculated from the available data.

<Insert Table 10 about here>

3.8. Examination of the effects of state anxiety

Four studies reported on some analysis comparing high reported anxiety at the time of testing and neuropsychological performance. Three studies reported no effect of state anxiety on neuropsychological performance in the PD group (Ludewig et al., 2003; Boldrini et al., 2005; Gladsjo et al., 1998). Lucas et al. (1991) reported that state anxiety accounted for some of the difference in visual memory performance seen in their PD participants.

3.9. Summary of Neuropsychological Findings

The findings of the reviewed studies suggest some support for a significant difference in short term memory performance in people with PD in both verbal and visual memory, with five out of seven studies indicating a difference in performance. There was some support for impairment in perceptual ability, but although three out of four studies found differences, the reliability and validity of the measures used was unclear. Only three of fourteen studies found differences in long term memory compared to HC, which did not support long term memory impairment in PD. The findings reviewed did not indicate consistent differences in executive function performance, as only two of nine studies and two of twelve found group

differences. Similarly, working memory was impaired only in one of seven studies and one of four tests. Attention was not impaired in any of the seven studies incorporating six tests of attention. Visuospatial abilities were impaired in two out of five studies and two of four tests, which does not suggest consistent impairment. The weighted mean effect sizes for each domain were calculated and presented in Table 7.

Four studies compared participants diagnosed with PD compared to others diagnosed with OCD, one of three found significantly better performance in PD than OCD on tasks of verbal and visual memory. Mixed results were reported in relation to working memory and set shifting differences between groups. PD participants performed significantly better than OCD participants on measures of planning and organising, visuospatial ability and decision making. No significant differences between groups were found on measures of attention, perception or verbal fluency.

A PD group was reported to perform significantly better than a Social Phobia group on a measure of perception but no significant differences were reported on tasks of memory, attention, visuospatial ability, set shifting or the FAS. When PD was compared to other disorder groups, no significant differences were seen in neuropsychological performance between PD and groups diagnosed with Generalised Anxiety Disorder, Specific Phobia, or comorbid PD and Unipolar Depression. When PD was compared to patients with only Unipolar depression, PD participants were found to be significantly poorer on a task of working memory and significantly better on a set shifting task.

Unusually large effect sizes were found for the Gordeev (2008) study which had an influence on the overall weighted means in the visual memory, verbal memory and perception domains. Median effect size values may therefore be a more reliable estimate of performance trends in these areas.

3.10. Assessment of Methodological Quality

Table 2 contains study ratings on the eight quality criteria selected. This rating system provides an indication of the methodological strengths of the studies reviewed relative to each other, although it does not allow for detailed comparison.

Based on the chosen criteria, Purcell et al. (1998) was methodologically the strongest study, although the majority of studies were of average to high quality. Studies which reported significant results for more than half of the measures they reported, tended to be of lower quality, as defined by the quality criteria. Four such studies (Airaksinen et al., 2005; Gorini et al., 2010; Lautenbacher et al., 2002; Lucas et al., 1991) did not describe adjustment for the multiple comparisons they used. In addition, three of these studies (Airaksinen et al., 2005; Gorini et al., 2010; Lautenbacher et al., 2002) reported significant results on a measure for which no reliability or validity information was available. Gordeev (2008) used corrections for multiple comparisons and reported significant results for all the measures used, examining areas of perception and short term memory; however no reliability or validity data were available for any of these measures. These studies provided some of the support in favour of short term memory and perception difficulties in people with PD.

Studies which reported few significant differences associated with diagnosis of PD, tended to be of high quality as defined by the quality criteria. These studies (Asmundson et al., 1994; Boldrini et al., 2005; Deckersbach et al., 2011; Galderisi et al., 2008) reported significant differences between PD and HC participants for fewer than half of the measures they investigated. Although two (Asmundson et al., 1994; Galderisi et al., 2008) used one measure in their study that did not have reliability and validity data available, significant results were only reported on validated measures. These studies contributed findings supporting difficulties in short term memory, working memory span and learning, visuospatial abilities and executive function.

Four studies of high quality, as defined by the quality criteria, reported no significant differences between PD samples and HC (Cavedini et al., 2002; Gladsjo et al., 1998; Kaplan et al., 2006; Purcell et al., 1998). Three described no differences in relation to long term memory and set shifting (Gladsjo et al., 1998; Kaplan et al., 2006; Purcell et al., 1998); two in relation to decision making (Cavedini et al., 2002; Kaplan et al., 2006) and two relating to other aspects of executive functioning (Gladsjo et al., 1998; Purcell et al., 1998). One lower quality study also reported no significant findings on a decision making task for which no psychometric information was found (Ludewig et al., 2003). Studies with high methodological quality tended towards findings of little or no differences between PD and HC groups. However three of the four high quality studies with negative findings had small sample sizes, reducing their power to detect differences.

Overall, within the studies reviewed, group matching, method of diagnosis and description of exclusion criteria were addressed adequately. Presentation of results was generally adequate but all studies failed to provide confidence intervals or effect sizes with their results. Eight of the 14 studies reviewed reported sample sizes less than 25, indicating that they would have lacked the power to detect a large effect size with an alpha level of .05 in a 2 tailed comparison of two means (see Table 1). No studies reported using an a priori power calculation. Uptake levels were poorly reported or not addressed in all but three studies (Airaksinen et al., 2005; Gladsjo et al., 1998; Purcell et al., 1998). Just over half of the studies reported corrections for multiple comparisons (Asmundson et al., 1994; Boldrini et al., 2005; Cavedini et al., 2002; Deckersbach et al., 2011; Galderisi et al., 2008; Gordeev, 2008; Kaplan et al., 2006; Purcell et al., 1998). As most studies included a number of measures and various post hoc tests, correction for multiple comparisons is an important issue for the avoidance of type I errors.

No studies reported on the reliability and validity of the measures used, therefore these properties were further investigated. Most measures were described with psychometric properties in Lezak and colleagues' (2004) detailed description of neuropsychological assessment and were described as valid and reliable for the groups in question. However, six studies used a measure or measures not described in Lezak et al. (2004) and did not provide reference to an appropriate source of reliability and validity data (Airaksinen et al., 2005; Asmundson et al., 1994; Galderisi et al., 2008; Gordeev, 2008; Gorini et al., 2010; Ludewig et al., 2003). One study used a measure from a German test battery and no reliability or validity data was found in English (Lautenbacher et al., 2002). An appropriate search for these data could not be performed in German due to translation difficulties.

4. Discussion

This systematic review examined the neuropsychological profile of individuals with panic disorder using the available literature. With only 14 studies included in the review, it demonstrated the scarcity of research in the area. The results obtained in these studies mostly indicated an absence of difficulties in PD participants relative to HC, with no statistically significant group differences being consistently reported across studies. There was some support for potential impairments in short term verbal and visual memory in people with PD compared to HCs. Results provided little support for impairment in any other area of neuropsychological function. Differences in performance between PD and HC participants in a number of neuropsychological areas was suggested by small to medium weighted mean effect sizes calculated for each domain, however the reviewed studies did not have the power to detect a statistically significant differences of this size.

In relation to other anxiety disorders, using the limited comparisons in the reviewed studies, PD participants performed significantly better than OCD participants on measures of

planning and organising, visuospatial ability and decision making. Mixed results were reported in relation to memory, working memory and set shifting. No significant differences in neuropsychological performance were reported between PD groups and groups diagnosed with Generalised Anxiety Disorder or Specific Phobia and a difference in performance in the domain of perception was reported between PD and a group diagnosed with Social Phobia. Overall these comparisons suggested that patients with PD may perform differently on neuropsychological measures to patients with OCD but not necessarily from patients with other common anxiety disorders.

Evidence from the reviewed studies also suggested no significant differences in neuropsychological performance between a group with PD alone, and a group with PD and comorbid depression. In light of literature which suggests that comorbid anxiety and depression lead to poorer neuropsychological performance than anxiety alone, this is unexpected (Kizilibash et al., 2002). However the number of participants involved was small and statistical power was very limited. When a PD group was compared to patients with Unipolar Depression alone, PD participants were found to perform significantly poorer than depressed patients on a working memory task and significantly better on a set shifting task. These results were difficult to interpret based on such limited information but may warrant further study.

A number of factors may have influenced the obtained results, including methodological quality and characteristics of the sample used by each study, such as the presence of comorbid disorders and the medication status of participants. These factors may also affect the generalisability of the results.

Quality criteria were applied to the studies reviewed, in order to further evaluate the reported findings. Key issues arising from the assessment of methodological quality were risk of type I error by failing to correct for multiple comparisons and use of measures without

evident reliability or validity data for this population. Considering only studies which do not suffer from these methodological weaknesses the overall profile changed very little: some support remained for visual and verbal short term memory difficulties in people with PD (Asmundson et al., 1994; Deckersbach et al., 2011) but the lack of consistency of results did not support a conclusion of impairment in this area. No remaining studies supported verbal long term memory or perception differences between people with PD and HCs. Only one of the five remaining studies reporting on visual long term memory provided support for group differences in that area (Deckersbach et al., 2011). This removal of the less methodologically strong studies did not change the overall findings of no group differences on tasks of working memory, attention, visuospatial ability and executive functioning.

As a number of Axis I and Axis II disorders have been associated with cognitive impairment (Trivedi, 2006), the inclusion of PD participants with comorbid disorders in some of the reviewed studies introduces potential confounders. Eight studies reported participants as having no comorbid disorders, two allowed all comorbidities, two excluded only depression and two did not clearly state their exclusions. The exclusion of comorbidities would help to isolate difficulties that are due to PD alone without the influence of other psychological disorders. Within the reviewed studies, patients without comorbidity tended to perform similarly to HCs (Cavedini et al., 2002; Gladsjo et al., 1998; Kaplan et al., 2006; Purcell et al., 1998). This pattern of results suggested that the presence of comorbid disorders increased the likelihood of finding poorer neuropsychological performance in patients with PD or that the poor performance seen in some studies may be related to the comorbid disorder rather than PD. However this was not seen in Kaplan et al.'s (2006) study where comorbid UD did not lead to significantly poorer performance in PD. Further research will be required to clarify this.

Half of the studies reviewed reported including participants on medication, although two of these excluded people who were taking benzodiazepines. There were no trends in findings relating to medication status of participants. This was somewhat surprising as benzodiazepines (Deckersbach et al., 2011) and tricyclic antidepressants (Stein and Strickland, 1998) have been associated with additional cognitive impairment while Selective Serotonin Reuptake Inhibitors have not been consistently associated with impairment (Mataix-Cols et al., 2002).

Ten of the fourteen studies matched groups on age, gender and education. Of the four poorly matched groups, three of these were among those who produced a high number of significant findings. Poor group matching at the outset may have influenced results, as differences in age, gender and education have been shown to have an impact on neuropsychological test performance (Corral et al., 2006; Lowe et al., 2003; Reitan and Wolfson, 1995).

State anxiety at the time of testing was measured in eleven of fourteen studies. Eight of these made comparisons between PD and HC groups. In these studies, statistical tests suggested that PD groups were more anxious than HC at the time of testing but there was no pattern in the data relating to participant groups identified as being more anxious subsequently performing worse on tasks. Four studies reported on investigating the effect of state anxiety on test performance, three of these finding no effect (Boldrini et al., 2005; Gladsjo et al., 1998; Ludewig et al., 2003; Lucas et al., 1991). Research using other populations also suggests that state anxiety is unlikely to have an impact on test performance (Gass and Curiel, 2011; O'Jile et al., 2005; Smitherman et al., 2007).

4.1. Limitations

Only papers written in English were included in this review, limiting its scope. At least one potentially relevant study, not published in English, was excluded (Castillo et al., 2010). Studies containing PD samples of fewer than 15 participants were excluded from the review. This also reduced the number of studies reviewed, however the statistical power of such studies would have been low and findings, particularly negative findings, would have been difficult to interpret (Bezeau and Graves, 2001). This review is based on a relatively small number of studies; however this is primarily due to the scarcity of literature rather than the exclusion of potentially relevant studies. The consistency of the findings across these studies allows for greater confidence in conclusions drawn from this small number of studies.

4.2. Recommendations/implications for future research

These studies suggested no consistent differences in cognitive performance between individuals with PD and healthy control participants, which was in keeping with similar findings in people with Social Phobia and Generalised Anxiety Disorder (Airaksinen et al., 2005; O'Toole and Pedersen, 2011). The weighted mean effect sizes calculated suggested a potential small to medium effect size, but this is smaller than what is usually considered a clinically significant effect size in neuropsychological research (Bezeau and Graves, 2001). Therefore the present review did not suggest any general implications of neuropsychological functioning for the treatment of PD (in a broad sense); though of course individual variations may mean that neuropsychological functioning has an impact on the treatment of some individuals with PD. An impairment in short term memory, if it were present in some PD patients, may have an impact on the psychoeducation phase of CBT treatment, as recommended by the National Institute of Health and Clinical Excellence (NICE, 2011). The provision of written materials and other memory aids could potentially be helpful.

An avoidance of type I error in relation to multiple comparisons has been discussed already, but type II errors were also of relevance to the studies in this review. Future research should consider using sample sizes appropriate to detecting small to medium effect sizes, as weighted mean effect sizes calculated ranged from 0.08 to 0.48. Reporting on the effect sizes obtained would further illustrate the potential magnitude of any differences detected (Bezeau and Graves, 2001). Specific hypotheses focussing on the highlighted areas of potential impairment, particularly short term memory, with an effort to use the same or directly comparable measures to other studies, would contribute to the clarification of findings. In addition the specificity of any potential impairment requires further examination. While PD has been compared to OCD on a number of occasions (Bannon et al., 2006), comparisons with disorders such as Social Phobia and GAD, which have demonstrated similar patterns of neuropsychological performance may help to further illustrate if there are any specific impairments related to PD.

4.3. Conclusion

This systematic review of the neuropsychological profile of Panic Disorder (PD) demonstrates that within the current literature there is little support for any neuropsychological impairment in PD. Some support was found for impairment in short term memory which requires further investigation using larger sample sizes (25 or more) and the use of appropriate clinical comparison groups to determine the specificity of any impairment found.

References

Airaksinen, E., Larsson, M., Forsell, Y., 2005. Neuropsychological functions in anxiety disorders in population-based samples: evidence of episodic memory dysfunction. J. Psychiatr. Res. 39, 207-214.

American Psychiatric Association, 2000. Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR, 4th Edition, Text Revision. American Psychiatric Press, Inc., Washington, DC.

Aupperle, R.L., Melrose, A.J., Stein, M.B., Paulus, M.P., 2012. Executive function and PTSD:disengaging from trauma. Neuropharmacology 62, 686-694.

Asmundson, G.J., Stein, M.B., Larsen, D.K., Walker, J.R., 1994. Neurocognitive function in panic disorder and social phobia patients. Anxiety 1, 201-207.

Baillie, A.J., Rapee, R.M., 2005. Panic attacks as risk markers for mental disorders*. Soc. Psychiatry Psychiatr. Epidemiol. 40, 240-244.

Bannon, S., Gonsalvez, C.J., Croft, R.J., Boyce, P.M., 2006. Executive functions in obsessive-compulsive disorder: state or trait deficits? Aust. N. Z. J. Psychiatry 40, 1031-1038.

Bechara, A., Damasio, A.R., Damasio, H., Anderson, S.W., 1994. Insensitivity to future consequences following damage to human prefrontal cortex. Cognition 50, 7-15.

Benton, A.L., 1945. A visual retention test for clinical use. Archives of Neurology And Psychiatry 54, 212-216.

Benton, A.L., 1983. Contributions to neuropsychological assessment. Oxford University Press, New York.

Benton, A.L., Hamsher, K., 1983. Multilingual Aphasia Examination. AJA Associates, Iowa City.

Benton, A.L., 1989. Multilingual aphasia examinations. AJA Associates, Iowa City, IA.

Berch, D., 1998. The Corsi Block-tapping Task: methodological and theoretical considerations. Brain Cogn. 38, 317-338.

Bezeau, S., Graves, R., 2001. Statistical power and effect sizes of clinical neuropsychology research. J. Clin. Exp. Neuropsychol. 23, 399-406.

Bisaga, A., Katz, J.L., Antonini, A., Wright, C.E., Margouleff, C., Gorman, J.M., Eidelberg, D., 1998. Cerebral glucose metabolism in women with panic disorder. Am. J. Psychiatry 155, 1178-1183.

Boldrini, M., del Pace, L., Placidi, G., Keilp, J., Ellis, S., Signori, S., Cappa, S., 2005. Selective cognitive deficits in obsessive-compulsive disorder compared to panic disorder with agoraphobia. Acta Psychiatr. Scand. 111, 150-158.

Burgess, P.W., 2003. Assessment of executive function, In: Halligan, P.W., Kischka, U., Marshall, J.C. (Eds.), Handbook of Clinical Neuropsychology. Oxford University Press, Oxford; New York.

Buschke, H., Fuld, P.A. (1974). Evaluating storage, retention and retrieval in disordered memory and learning. Neurology 24: 1019-1025.

Castillo, E.P., Coy, P.E.C., Shejet, F.O., Duran, E.T., Cabrera, D.M., 2010. Cognitive function evaluation: attention and memory in panic disorder patients. Salud Mental 33, 481-488.

Cavedini, P., Riboldi, G., D'Annucci, A., Belotti, P., Cisima, M., Bellodi, L., 2002. Decision-making heterogeneity in obsessive-compulsive disorder: Ventromedial prefrontal cortex function predicts different treatment outcomes. Neuropsychologia 40, 205-211.

Cavedini, P., Zorzi, C., Bassi, T., Gorini, A., Baraldi, C., Ubbiali, A., Bellodi, L.,

2006. Decision-making functioning as a predictor of treatment outcome in anorexia nervosa. Psychiatry Res. 145, 179-187.

Cohen, J., 1992. A power primer. Psychological Bulletin, Vol. 112. No. 1, 155-159.

Corral, M., Rodriguez, M., Amenedo, E., Sanchez, J.L., Diaz, F., 2006. Cognitive reserve, age, and neuropsychological performance in healthy participants. Dev. Neuropsychol. 29, 479-491.

CRD, 2008. CRD's guidance for undertaking reviews in health care. Centre for Reviews and Dissemination, University of York.

Deckersbach, T., Moshier, S.J., Tuschen-Caffier, B., Otto, M.W., 2011. Memory dysfunction in panic disorder: an investigation of the role of chronic benzodiazepine use. Depress. Anxiety 28, 999-1007.

Delis, D.C., Kramer, J.H., Kaplan, E., Ober, B.A., 1987. California Verbal Learning Test. The Psychological Corporation, New York.

Ennok, N., 2010. The Untested Assumptions of Schulte Tables, The International Neuropsychological Society, The Polish Neuropsychological Society and The Polish Neuroscience Society Joint Mid-Year Meeting, Krakow. Poland.

Fletcher, J.M., 1985. Memory for verbal and nonverbal stimuli in learning disability subgroups: analysis by selective reminding. J. Exp. Child Psychol. 40, 244-259.

Galderisi, S., Mancuso, F., Mucci, A., Garramone, S., Zamboli, R., Maj, M., 2008. Alexithymia and cognitive dysfunctions in patients with panic disorder. Psychother. Psychosom. 77, 182-188.

Gass, C.S., Curiel, R.E., 2011. Test anxiety in relation to measures of cognitive and intellectual functioning. Arch. Clin. Neuropsychol. 26, 396-404.

Gladsjo, J.A., Rapaport, M.H., McKinney, R., Lucas, J.A., Rabin, A., Oliver, T., Davis, J., Auerbach, M., Judd, L.L., 1998. A neuropsychological study of panic disorder: negative findings. J. Affect. Disord. 49, 123-131.

Goodglass, H. and Kaplan, E. (1983) Boston Diagnostic Aphasia Examination (BDME). Philadelphia: Lea & Fabinger.

Gordeev, S.A., 2008. Cognitive functions and the state of nonspecific brain systems in panic disorders. Neurosci.Behav. Physiol. 38, 707-714.

Gorini, A., Schruers, K., Riva, G., Griez, E., 2010. Nonhomogeneous results in place learning among panic disorder patients with agoraphobia. Psychiatry Res. 179, 297-305.

Heaton, R.K., Chelune, G.J., Talley, J.L., Kay, G.G., Curtiss, G., 1993. Wisconsin Card Sorting Test Manual: revised and expanded.. Psychological Assessment Resources Inc, Florida.

Higgins, J., Altman, D., 2008. Chapter 8: Assessing risk of bias in included studies., In: Higgins, J., Green, S. (Eds.), Cochrane handbook for systematic reviews of interventions Version 5.0.1 (updated September 2008). The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

Higgins, J., Deeks, J. J., 2011. Chapter 7: Selecting studies and collecting data. In: Higgins, J., Green, S. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Horwitz, J.E., McCaffrey, R.J., 2008. Effects of a third party observer and anxiety on tests of executive function. Arch. Clin. Neuropsych., 23, 409-417.

Jacobs, W.J., Laurance, H.E., Thomas, K.G.F., 1997. Place learning in virtual space I: acquisition, overshadowing, and transfer. Learn. Motiv. 28, 521-

541.

Kaplan, J.S., Erickson, K., Luckenbaugh, D.A., Weiland-Fiedler, P., Geraci, M., Sahakian, B.J., Charney, D., Drevets, W.C., Neumeister, A., 2006. Differential performance on tasks of affective processing and decision-making in patients with panic disorder and panic disorder with comorbid major depressive disorder. J. Affect. Disord. 95, 165-171.

Kessler, R.C., Chiu, W.T., Jin, R., Ruscio, A.M., Shear, K., Walters, E.E., 2006. The epidemiology of panic attacks, panic disorder, and agoraphobia in the National Comorbidity Survey Replication. Arch. Gen. Psychiatry 63, 415-424.

Kizilbash, A.H., Vanderploeg, R.D., Curtiss, G., 2002. The effects of depression and anxiety on memory performance. Arch. Clin. Neuropsych. 17, 57-67.

Lautenbacher, S., Spernal, J., Krieg, J.C., 2002. Divided and selective attention in panic disorder. A comparative study of patients with panic disorder, major depression and healthy controls. Eur. Arch. Psychiatry Clin. Neurosci. 252, 210-213.

Lewis, R., Rennick, P.M., 1979. Manual for the Repeatable Cognitive-Perceptual-Motor Battery. Axon, Grosse Point, MI.

Lezak, M.D., Howieson, D.B., Loring, D.W., 2004. Neuropsychological assessment 4th Edition. Oxford University Press, New York.

Lipsey, M.W., Wilson, D.B., 2001. Practical Meta-analysis. Sage Publications, Thousand Oaks, California.

Lowe, P.A., Mayfield, J.W., Reynolds, C.R., 2003. Gender differences in memory test performance among children and adolescents. Arch. Clin.Neuropsychol. 18, 865-878.

Lucas, J.A., Telch, M.J., Bigler, E.D., 1991. Memory functioning in panic disorder: A neuropsychological perspective. J. Anxiety Disord. 5, 1-20.

Ludewig, S., Paulus, M.P., Ludewig, K., Vollenweider, F.X., 2003. Decision-making strategies by panic disorder subjects are more sensitive to errors. J. Affect. Disord. 76, 183-189.

Martinez-Gonzalez, A.E., Piqueras-Rodriguez, J.A., 2008. Neuropsychological update on obsessive-compulsive disorder. Rev. Neurol. 46, 618-625.

Mataix-Cols, D., Alonso, P., Pifarré, J., Menchón, J.M., Vallejo, J., 2002. Neuropsychological performance in medicated vs. unmedicated patients with obsessive–compulsive disorder. Psychiatry Res. 109, 255-264.

Mataix-Cols, D., van den Heuvel, O.A., 2006. Common and distinct neural correlates of obsessive-compulsive and related disorders. Psychiatr. Clin. North Am. 29, 391-410.

Menzies, L., Chamberlain, S.R., Laird, A.R., Thelen, S.M., Sahakian, B.J., Bullmore, E.T., 2008. Integrating evidence from neuroimaging and neuropsychological studies of obsessive-compulsive disorder: the orbitofronto-striatal model revisited. Neurosci. Biobehav. Rev. 32, 525-549.

Millan, M.J., Agid, Y., Brüne, M., Bullmore, E.T., Carter, C.S., Clayton, N.S., Connor, R., Davis, S., Deakin, B., DeRubeis, R.J., Dubois, B., Geyer, M.A., Goodwin, G.M., Gorwood, P., Jay, T.M., Joëls, M., Mansuy, I.M., Meyer-Lindenberg, A., Murphy, D., Rolls, E., Saletu, B., Spedding, M., Sweeney, J., Whittington, M., Young, L.J., 2012. Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. Nat. Rev. Drug Discov. 11, 141-168.

NICE, 2011. Generalised anxiety disorder and panic disorder (with or without agoraphobia) in adults. NICE clinical guideline 113.

O'Jile, J.R., Schrimsher, G.W., O'Bryant, S.E., 2005. The relation of self-report of mood and anxiety to CVLT-C, CVLT, and CVLT-2 in a psychiatric sample. Arch. Clin. Neuropsych. 20, 547-553.

O'Toole, M.S., Pedersen, A.D., 2011. A systematic review of neuropsychological performance in social anxiety disorder. Nord. J. Psychiatry 65, 147-161.

Olley, A., Malhi, G., Sachdev, P., 2007. Memory and executive functioning in obsessive–compulsive disorder: a selective review. J. Affect. Disord. 104, 15-23.

Orsillo, S. M, McCaffrey, R. J., 1992. The impact of state-trait anxiety on neuropsychological test performance, In: Puente, A. E. and McCaffrey, R. J. (Eds.), Handbook of Neuropsychological Assessment: A Biopsychosocial Perspective. Plenum Press, New York

Paulus, M.P., 1997. Long-range interactions in sequences of human behavior. Phys. Rev. E 55.

Phan, K.L., Orlichenko, A., Boyd, E., Angstadt, M., Coccaro, E.F., Liberzon, I., Arfanakis, K., 2009. Preliminary evidence of white matter abnormality in the uncinate fasciculus in generalized social anxiety disorder. Biol. Psychiatry 66, 691-694.

Purcell, R., Maruff, P., Kyrios, M., Pantelis, C., 1998. Neuropsychological deficits in obsessive-compulsive disorder: a comparison with unipolar depression, panic disorder, and normal controls. Arch. Gen. Psychiatry 55, 415-423.

Reiman, E.M., Fusselman, M.J., Fox, P.T., Raichle, M.E., 1989a. Neuroanatomical correlates of anticipatory anxiety. Science 243, 1071-1074.

Reiman, E.M., Raichle, M.E., Robins, E., Mintun, M.A., Fusselman, M.J., Fox, P.T., Price, J.L., Hackman, K.A., 1989b. Neuroanatomical correlates of a lactate-induced anxiety attack. Arch. Gen. Psychiatry 46, 493-500.

Reitan, R.M., 1958.Validity of the trail making test as an indicator of organic brain disease. Perceptual Motor Skills 8, 271–276.

Reitan, R.M., Davidson, L.A., 1974. Clinical neuropsychology: current status and applications. Wiley, New York.

Reitan, R.M., Wolfson, D., 1995. Influence of age and education on neuropsychological test results. Clin. Neuropsychol. 9, 151-158.

Rey, A., 1941. L'examen psychologique dans les cas d'encephalopathie traumatique.(Les problems.). Archives de Psychologie 28, 215-285.

Rogers, R., Everitt, B., Baldacchino, A., Blackshaw, A., Swainson, R., Wynne, K., Baker, N., Hunter, J., Carthy, T., Booker, E., London, M., Deakin, J., Sahakian, B., Robbins, T., 1999. Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: evidence for monoaminergic mechanisms. Neuropsychopharmacology 20, 322-339.

Rosvold, H.E., Mirsky, A.E., Sarason, I., Bransome, E.D.J., Beck, L.H., 1956. A continuous performance tests of brain damage. J Consult Psychol. 20, 343–350.

Schuhfried, G., 1999. Vienna test system (WINWTS), Version 4.50, Modling, Austria.

Smitherman, T.A., Huerkamp, J.K., Miller, B.I., Houle, T.T., O'Jile, J.R., 2007. The relation of depression and anxiety to measures of executive functioning in a mixed psychiatric sample. Arch. Clin. Neuropsych. 22, 647-654.

Spielberger, C. D. (1983). Manual for the state–trait anxiety inventory (STAI). Palo Alto, CA: Consulting Psychologists Press.

Spernal, J., Krieg, J.C., Lautenbacher, S., 2003. Pain thresholds as a putative functional test for cerebral laterality in major depressive disorder and panic disorder. Neuropsychobiology 48, 146-151.

Stein, R.A., Strickland, T.L., 1998. A review of the neuropsychological effects of commonly used prescription Medications. Arch. Clin. Neuropsychol. 13, 259-284.

Szeszko, P.R., Ardekani, B.A., Ashtari, M., Malhotra, A.K., Robinson, D.G., Bilder, R.M., Lim, K.O., 2005. White matter abnormalities in obsessive-compulsive disorder: a diffusion tensor imaging study. Arch. Gen. Psychiatry 62, 782-790.

Tabarés-Seisdedos, R., Balanzá-Martínez, V., Sánchez-Moreno, J., Martinez-Aran, A., Salazar-Fraile, J., Selva-Vera, G., Rubio, C., Mata, I., Gómez-Beneyto, M., Vieta, E., 2008. Research report: Neurocognitive and clinical predictors of functional outcome in patients with schizophrenia and bipolar I disorder at one-year follow-up. J. Affect. Disord. 109, 286-299.

Trahan, D.E., Larrabee, G.J., 1989. Continuous Visual Memory Test. Psychological Assessment Resources, Odessa, FL.

Trivedi, J.K., 2006. Cognitive deficits in psychiatric disorders: Current status. Indian J. Psychiatry 48, 10-20.

van den Heuvel, O.A., Veltman, D.J., Groenewegen, H.J., Witter, M.P., Merkelbach, I., Cath, D.C., van Balkom, A., van Oppen, P., van Dyck, R., 2005. Disorder-specific neuroanatomical correlates of attentional bias in obsessive-compulsive disorder, panic disorder, and hypochondriasis. Arch. Gen. Psychiatry 62, 922-933.

Vandenberg, S.G., Kuse, A.R., 1978. Mental rotations, a group test of three-dimensional spatial visualization. . Percept. Mot. Skills 47, 599-604.

Vythilingam, M., Anderson, E.R., Goddard, A., Woods, S.W., Staib, L.H., Charney, D.S., Bremner, J.D., 2000. Temporal lobe volume in panic disorder - a quantitative magnetic resonance imaging study. Psychiatry Res.-Neuroimaging 99, 75-82.

Waldenstein, S.R., Ryan, C.M., Jennings, R., Muldoon, M.F., Manuck, S.B., 1997. Self-reported levels of anxiety do not predict neuropsychological performance in healthy men. Arch. Clin. Neuropsych. 12, 6, 567-574.

Warrington, E.K., 1984. Recognition Memory Test. NFER-Nelson, Windsor, UK.

Wechsler, D., 1981. Wechsler Adult Intelligence Scale-Revised (WAIS-R). The Psychological Corporation.

Wechsler, D., 1987. Wechsler Memory Scale-Revised Manual. The Psychological Corporation, San Antonio, TX.

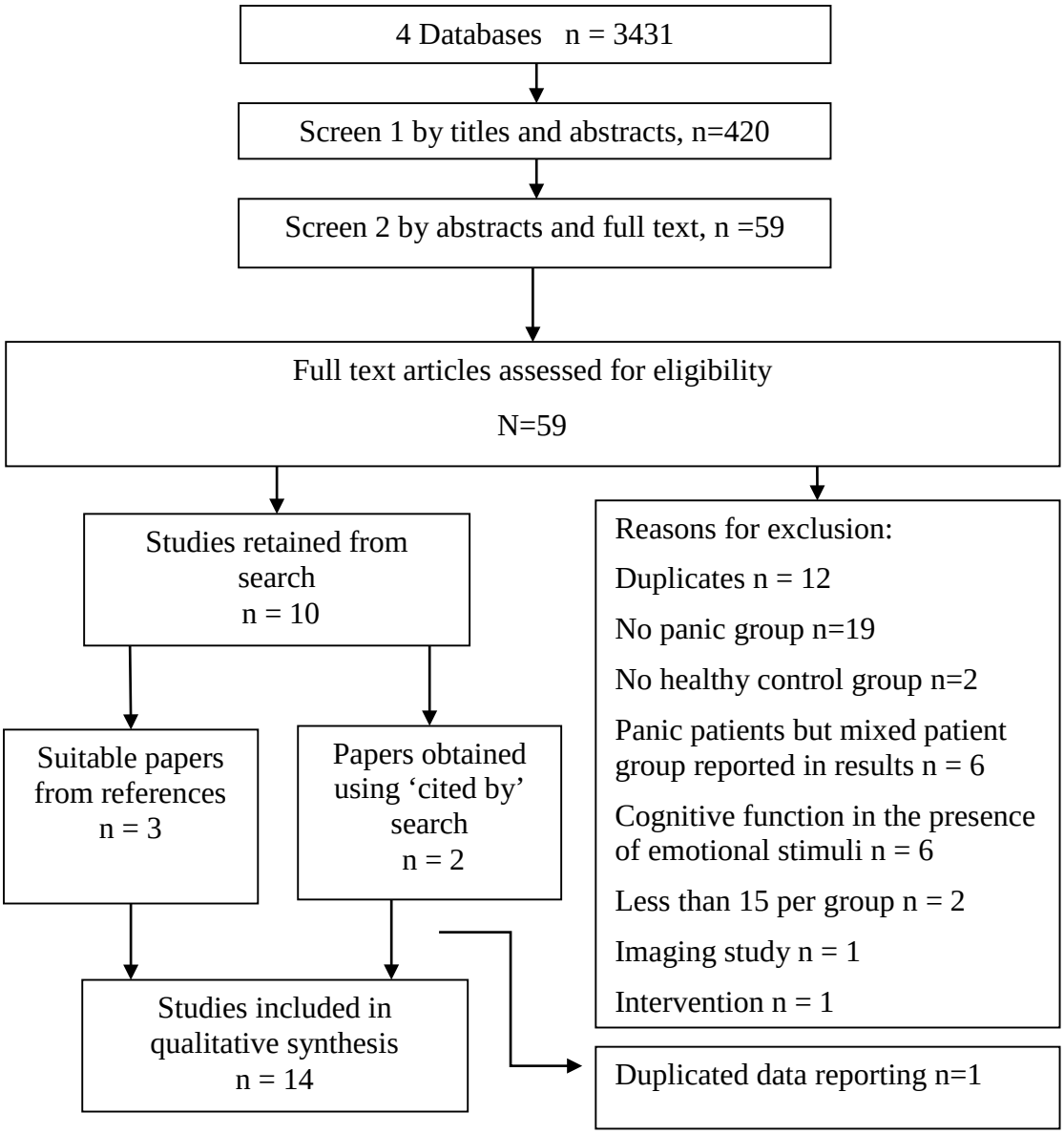

Figure 1 Flowchart showing search process and the number of included and excluded studies.

Table 1 Key Study Characteristics

Authors	Sample size[a]	Groups matched on	Key exclusion criteria	Levels of anxiety and depression compared between groups	Medication
Airaksinen et al. (2005)	33 PD (30 PD only and 3 Ag only) 32 SP, 7 GAD, 16 OCD 24 Specific P, 175 HC	age, education	No exclusions	No comparisons	No medications excluded
Asmundson et al. (1994)	18 PD 18 SP 16 HC	gender, age, education	Current major depression excluded in all groups. SP excluded in PD group, PD excluded in SP group	BDI: PD = SP > HC BAI: PD = SP > HC	1 PD taking pro re nata benzodiazepines
Boldrini et al. (2005)	15 PD with Ag 25 OCD 15 HC	gender, age, education handedness, intelligence	All Axis I or II excluded	No comparisons	Free from benzodiazepines (no time period) but SSRIs not excluded
Cavedini et al. (2002)	16 PD 34 OCD 34 HC[b]	gender, age, education	All Axis I or II excluded	No comparisons	All medication free for at least 2 weeks

Deckersbach et al. (2011)	20 PD 20 HC	gender, age, education	Depression, psychosis and bipolar disorder excluded	STAI: PD > HC BDI: PD > HC	All free from benzodiazepines for at least 4 weeks. One PD taking Sertraline
Galderisi et al. (2008)	28 PD (26 with Ag) 32 HC	gender, age, education handedness	MDD and other anxiety disorders excluded	No comparisons	Medication free for 4 weeks or drug naive
Gladsjo et al. (1998)	69 PD 19 HC	gender, age, education ethnicity, handedness	All Axis I or II excluded	No comparisons	Medication free for at least 2 weeks
Gordeev (2008)	93 PA 36 HC	education	No information	BDI: PD > HC STAI: PD > HC	Medication free for 2 weeks
Gorini et al. (2010)	31 PD with Ag, 31 HC	gender, age, education	Other primary diagnoses excluded	No comparisons	Medication free for 1-2 weeks
Kaplan et al. (2006)	22 PD (11 PD only, 11 PD + MDD)[c], 22 HC	gender, age, education	Other anxiety or depressive disorder excluded	MADRS: PD > HC Ham A: PD > HC	All medication free (no time period given)
Lautenbacher et al (2001)	21 PD, (16 with Ag, 5 without Ag) 21 MDD	gender	Lifetime comorbidity of Axis 1 excluded	No comparison	Medication free for 6 days

	20 HC[b]				
Lucas et al. (1991)	25 PD 25 HC	gender, age, education handedness	Current mood disorder, or other anxiety disorder excluded	BDI: PD > HC STAI: PD > HC	Patients remained on medication
Ludewig et al. (2003)	18 PD 18 MDD 35 HC[b]	gender, age	No information	No comparisons	14 of 18 PDs on medication including SSRIs and TCA
Purcell et al. (1998)	30 PD 30 OCD 30 UD 30 HC	gender, age, education handedness, IQ	Comorbid disorder excluded but anxiety or depression symptoms accepted	Ham D: PD = OCD = HC Ham A: PD > OCD = HC	19 PD on medication

a PD group is mixed with and without agoraphobia with details unavailable unless otherwise specified, *b* Additional groups were included in the study which did not form part of this review, *c* MDD is episode secondary to PD, MDD = Major Depressive Disorder, UD = Unipolar Depression, SP = Social Phobia, GAD = Generalised Anxiety Disorder, OCD = Obsessive Compulsive Disorder, Specific P = Specific Phobia, STAI= State Trait Anxiety Inventory, TCA = tricyclic antidepressant, SSRI = selective serotonin reuptake inhibitor

Ag = agoraphobia, HamA = Hamilton Anxiety rating scales, HamD = Hamilton Depression rating scale, BDI= Beck Depression Inventory, BAI= Beck Anxiety Inventory, MADRS = Montgomery Asberg Depression Rating Scale

Table 2 Quality criteria applied to reviewed studies

Name of study	i exclusion criteria	ii group matching	iii diagnosis	iv neuropsych measures	v sample size	vi uptake levels	vii outputs reported	viii Analysis	Total
Airaksinen et al. (2005)	WC	AA	WC	AA	WC	AA	AA	AA	11
Asmundson et al. (1994)	WC	WC	WC	AA	AA	NA	AA	AA	10
Boldrini et al. (2005)	WC	WC	WC	WC	AA	NA	AA	AA	11
Cavedini et al. (2002)	AA	WC	WC	WC	AA	NA	AA	WC	11
Deckenbacher et al. (2011)	AA	WC	WC	WC	AA	NA	AA	AA	10
Galderisi et al.(2008)	WC	WC	WC	AA	WC	NA	AA	AA	11
Gladsjo et al. (1998)	WC	WC	WC	WC	AA	AA	AA	AA	12
Gordeev (2008)	PA	PA	AA	AA	WC	NA	AA	WC	7
Gorini et al. (2010)	PA	WC	AA	AA	WC	NA	AA	AA	8
Kaplan et al. (2006)	WC	WC	WC	WC	AA	NA	AA	WC	12
Lautenbacher et al. (2001)	WC	AA	AA	AA	AA	NA	AA	AA	8
Lucas et al. (1991)	WC	WC	WC	WC	WC	NA	AA	AA	12
Ludewig et al. (2003)	PA	AA	WC	AA	AA	NA	PA	AA	6

Purcell et al. (1998)	WC	WC	WC	WC	WC	AA	AA	WC	14

i. Eligibility criteria are specified, ii. Comparison group is matched, iii. Diagnosis using appropriate criteria and measure, iv. Neuropsychological Measures are robust, v. Sample size adequate for all groups, vi. Levels of uptake are reported, vii. Results – appropriate outputs provided, viii. Appropriate statistical techniques used.

WC=well covered, AA=adequately addressed, PA=poorly addressed, NA=not addressed.

Table 3 Verbal memory in panic disorder compared to healthy controls

Verbal Memory Test	study/authors	PD v HC	d
CVLT			
Immediate free recall trial 5	Asmundson et al. (1994)	↓	1.50
Immediate free recall trial 1	Asmundson et al. (1994)	-	1.11
Short delay free recall	Asmundson et al. (1994)	↓	1.33
Total free recall	Asmundson et al. (1994)	↓	1.11
	Deckersbach et al. (2011)	-	0.65
	Gladsjo et al. (1998)	-	0.30
Short delay cued recall	Asmundson et al. (1994)	-	0.83
Retention	Asmundson et al. (1994)	-	-0.29
	Deckersbach et al. (2011)	-	0.05
Response inhibition	Asmundson et al. (1994)	-	0.43
Response discrimination	Asmundson et al. (1994)	-	0.43
	Deckersbach et al. (2011)	-	0.31
	Gladsjo et al. (1998)	-	0.10
Word Lists			
AVLT delayed recall	Galderisi et al. (2008)	-	-0.06
Warrington RMT	Gladsjo et al. (1998)	-	0.14
Words remembered short term	Gordeev et al. (2008)	↓	10.27
Words remembered			
Free recall	Airaksinen et al. (2005)	↓	0.40
Cued recall	Airaksinen et al. (2005)	↓	0.35
Selective Reminding Test (SRT)			
Long term recall	Boldrini et al. (2005)	-	0.37
	Lucas et al. (1991)	↓	0.61
Long term storage	Boldrini et al. (2005)	-	1.98
	Lucas et al. (1991)	↓	0.57
Intrusions	Boldrini et al. (2005)	-	-0.05

Delayed recall	Boldrini et al. (2005)	-	1.30
	Lucas et al. (1991)	↓	0.90
Trials to criterion	Lucas et al. (1991)	-	0.44
Logical Memory (WMS)			
Delayed recall	Lucas et al. (1991)	-	0.59
Immediate recall	Deckersbach et al. (2011)	-	-0.18
Paired associate learning (WMS)		-	
Immediate recall	Lucas et al. (1991)	-	6.95
Delayed recall	Lucas et al. (1991)	-	0.28

↓Indicates significantly worse performance than HC

↑indicates significantly better performance than HC

- indicates no significant difference compared to HC

Table 4 Visual memory in panic disorder compared to healthy controls

Visual Memory Test	study/authors	PD v HC	d
Benton Visual Retention Test			
Form F (BVRT-F)	Asmundson et al. (1994)	-	0.16
Number of Errors	Deckersbach et al. (2011)	↓	0.74
Number of Errors	Lucas et al. (1991)	↓	1.06
RCFT			
immediate recall	Deckersbach et al. (2011)	↓	0.79
percent recall	Boldrini et al. (2005)	-	0.27
percent recall	Deckersbach et al. (2011)	-	-0.64
Visual Selective Reminding test (VSRT)			
long term storage	Gladsjo et al. (1998)	-	0.16
long term storage	Lucas et al. (1991)	↓	0.83
total recalled	Gladsjo et al. (1998)	-	0.21
long term retrieval	Gladsjo et al. (1998)	-	0.21
long term retrieval	Lucas et al. (1991)	↓	0.87
Delayed recall	Lucas et al. (1991)	↓	1.05
Trials to criterion	Lucas et al. (1991)	↓	0.74
Warrington RMT Faces	Gladsjo et al. (1998)	-	-0.33
Continuous Visual Memory test (CVMT)			
total recalled	Gladsjo et al. (1998)	-	-0.06
d-Prime	Gladsjo et al. (1998)	-	-0.13
Numbers remembered, short term	Gordeev et al. (2008)	↓	11.53
Visual Reproduction (WMS) - delayed recall	Lucas et al. (1991)	↓	0.88
CANTAB			
Spatial Recognition Memory	Kaplan et al. (2006)	-	-0.13
Spatial Recognition Memory	Purcell et al. (1998)	-	0.37

pattern recognition memory	Kaplan et al. (2006)	-	-0.12
pattern recognition memory	Purcell et al. (1998)	-	0.06
Delayed match to sample	Kaplan et al. (2006)	-	0.13
Delayed match to sample	Purcell et al. (1998)	-	0.18

↓Indicates significantly worse performance than HC

↑indicates significantly better performance than HC

- indicates no significant difference compared to HC

Table 5 Working memory in panic disorder compared to healthy controls

Cognitive function	Test	study/authors	PD v HC	d
Working Memory	Digit span	Boldrini et al. (2005)	-	1.95
		Deckersbach et al. (2011)	-	-0.28
		Gladsjo et al. (1998)	-	0.53
		Lucas et al. (1991)	-	-0.04
	Corsi Block Tapping Task (CBT)			
	span	Boldrini et al. (2005)	-	2.43
	supraspan	Boldrini et al. (2005)	↓	3.78
	accuracy index	Galderisi et al. (2008)	↑	-0.51
	Hebb Digit Recurring test - accuracy index	Galderisi et al. (2008)	-	0.43
	WMS-R			
	spatial span forward	Deckersbach et al. (2011)	-	-0.05
	spatial span backward	Deckersbach et al. (2011)	-	-0.49
	CANTAB			
	Spatial Span	Kaplan et al. (2006)	-	0.11
	Spatial Span	Purcell et al. (1998)	-	0.60
Attention	Mental control task (WMS-R)	Lucas et al. (1991)	-	0.21
	Continuous performance test	Galderisi et al. (2008)	-	-0.48
	TMT A	Airaksinen et al. (2005)	-	0.08
		Asmundson et al. (1994)	-	0.01
		Gladsjo et al. (1998)	-	-0.44
	Digit symbol (WAIS-R)	Galderisi et al. (2008)	-	0.47
		Gladsjo et al. (1998)	-	0.65
	Digit vigilance - time	Gladsjo et al. (1998)	-	-0.03
	Digit vigilance - errors	Gladsjo et al. (1998)	-	0.09
	CANTAB			

	Rapid Visual Information Processing	Kaplan et al. (2006)	-	0.12
Perception	Digit Cancellation Test (DCT)	Asmundson et al. (1994)	-	-0.06
	Signal Detection (from Weiner-Test-System)	Lautenbacher et al. (2001)	-	0.48
	munsterberg test	Gordeev (2008)	↓	29.47
	Schulte tables	Gordeev (2008)	↓	21.70
	Visual Field Neglect task (from TAP)	Lautenbacher et al. (2001)	↓	0.31
	Facial Recognition test (BFRT) short form	Boldrini et al. (2005)	-	0.16
Visuospatial	Block design (WAIS-R)	Asmundson et al. (1994)	↓	0.90
ability		Gladsjo et al. (1998)	-	0.26
	Mental rotation test	Deckersbach et al. (2011)	-	-0.29
	Spatial orientation and learning – virtual water maze analog	Gorini et al. (2010)	↓	0.55
	RCFT			
	Copy	Boldrini et al. (2005)	-	2.32
	Copy	Deckersbach et al. (2011)	-	-0.32

↓Indicates statistically significantly worse performance than HC

↑indicates statistically significantly better performance than HC

- indicates no significant difference compared to HC

Table 6 Executive function in panic disorder compared to healthy controls

Cognitive function	Test	study/authors	PD v HC	d
Set shifting	TMT B	Airaksinen et al. (2005)	↓	0.37
		Asmundson et al. (1994)	-	0.31
		Gladsjo et al. (1998)	-	-0.34
	WCST			
	categories	Boldrini et al. (2005)	-	2.53
	total errors	Boldrini et al. (2005)	-	2.77
	perseverative errors	Boldrini et al. (2005)	-	1.87
	non-p errors	Boldrini et al. (2005)	-	1.95
	perseverative responses	Boldrini et al. (2005)	-	1.63
	null sorts	Boldrini et al. (2005)	-	0.04
	CANTAB			
	Intradimensional-Extradimensional Shift			
	Stages	Kaplan et al. (2006)	-	0.11
	Adjusted errors	Kaplan et al. (2006)	-	0.20
	IDS trial score	Purcell et al. (1998)	-	-0.11
	EDS trial score	Purcell et al. (1998)	-	0.40
	Spatial working memory	Kaplan et al. (2006)	-	-0.04
	Spatial working memory	Purcell et al. (1998)	-	-0.13
Planning and orgainsing	Tower of London	Purcell et al. (1998)	-	0.16
	RCFT organisation	Deckersbach et al. (2011)	↓	1.42
verbal fluency	FAS in 60 sec	Airaksinen et al. (2005)	-	-0.08

	FAS in 60 sec	Boldrini et al. (2005)	-	2.38
	letter fluency	Gladsjo et al. (1998)	-	0.64
	COWAT	Deckersbach et al. (2011)	-	0.18
	category fluency	Gladsjo et al. (1998)	-	0.82
Decision				
making	Iowa Gambling Task	Cavedini et al. (2002)	-	u/a
	Cambridge Gambling Task	Kaplan et al. (2006)	-	u/a
	Two-Choice Prediction Task	Ludewig et al. (2003)	-	u/a

↓Indicates significantly worse performance than HC, ↑indicates significantly better performance than HC,

- indicates no significant difference compared to HC

u/a indicates sufficient data unavailable to make calculation

Table 7 Weighted mean effect sizes by domain

Domain	n	Range of ES	Median ES	Weighted mean ES	95% CI lower	95% CI upper	Q	Critical Q
Verbal Memory	8	-0.29-10.27	0.43	0.55	0.44	0.66	226.52	14.07
Visual Memory	8	0.64-11.53	0.21	0.31	0.20	0.42	267.01	14.07
Working memory	7	-0.51-3.78	0.27	0.29	0.10	0.47	61.04	12.6
Attention	6	-0.48-0.65	0.09	0.08	-0.07	0.23	0.54	11.07
Perception	4	-0.06-29.47	0.40	0.64	0.38	0.90	473.53	7.81
visuospatial	5	-0.32-2.32	0.41	0.38	0.13	0.63	28.49	9.49
set shifting	6	-.034-2.77	0.34	0.31	0.17	0.45	66.62	11.07
Planning and organising	2	0.16-1.42	0.79	0.29	0.07	0.51	11.16	3.84
Verbal Fluency	4	0.08-2.38	0.64	0.43	0.20	0.67	26.68	7.81
Total/Mean	9		0.29	0.35	0.30	0.41	966.82	15.51

Range of ES = range of effect sizes produced by each study on tasks in this domain, CI=confidence interval, and Q=homogeneity statistic.

Table 8 Memory performance of panic disorder compared to other anxiety disorders

Cognitive function	Test	study/authors	d PD v OCD	SocP	GAD	SpP
Verbal Memory	**CVLT**					
	Immediate free recall trial 5	Asmundson et al. (1994)		0.45 ns		
	Immediate free recall trial 1	Asmundson et al. (1994)		-0.36 ns		
	Short delay free recall	Asmundson et al. (1994)		0.72 ns		
	Total free recall	Asmundson et al. (1994)		0.05 ns		
	Short delay cued recall	Asmundson et al. (1994)		0.41 ns		
	Retention	Asmundson et al. (1994)		-0.44 ns		
	Response inhibition	Asmundson et al. (1994)		0.60 ns		
	Response discrimination	Asmundson et al. (1994)		0.31 ns		
	Word Lists					
	Words remembered					
	Free recall	Airaksinen et al. (2005)	0.20 ns	0.08 ns	0.11 ns	0.37 ns
	Cued recall	Airaksinen et al. (2005)	-0.02 ns	-0.04 ns	0	0.21 ns
	Selective Reminding Test (SRT)					

	Long term recall	Boldrini et al. (2005)	-1.88 ↑[a]	
	Long term storage	Boldrini et al. (2005)	0.30 ns	
	Intrusions	Boldrini et al. (2005)	-4.09 ↑	
	Delayed recall	Boldrini et al. (2005)	-0.41 ns	
Visual Memory	**Benton Visual Retention Test**			
	Form F (BVRT-F)	Asmundson et al. (1994)		0.46 ns
	RCFT			
	percent recall	Boldrini et al. (2005)	-0.13 ns	
	CANTAB			
	Spatial Recognition Memory	Purcell et al. (1998)	-0.57 ↑	
	pattern recognition memory	Purcell et al. (1998)	-0.64 ↑	
	Delayed match to sample	Purcell et al. (1998)	-0.28 ns	

↓Indicates statistically significantly worse performance than comparison group

↑ Indicates statistically significantly better performance than comparison group

ns Indicates no significant difference compared to comparison group

OCD = Obsessive compulsive Disorder, SoP= Social Phobia, GAD = Generalised Anxiety Disorder, SpP = Specific Phobia

a = not significant after Bonferroni corrections

Table 9 Working memory, attention, perception, visuospatial ability and executive function performance of panic disorder compared to other anxiety disorders

Cognitive function	Test	study/authors	d PD v OCD	SocP	GAD	SpP
Working Memory	Digit span	Boldrini et al. (2005)	1.62 ↓			
	Corsi Block Tapping Task (CBT)					
	span	Boldrini et al. (2005)	-2.45 ↑			
	supraspan	Boldrini et al. (2005)	0.38 ns			
	Spatial Span	Purcell et al. (1998)	0.08 ns			
Attention	TMT A	Airaksinen et al. (2005)	-0.11 ns	0.09 ns	-0.47 ns	0.28 ns
		Asmundson et al. (1994)		-0.20 ns		
	Digit Cancellation Test (DCT)	Asmundson et al. (1994)		-0.69 ↑		
Perception	Facial Recognition test (BFRT)	Boldrini et al. (2005)	0			
Visuospatial	Block design (WAIS-R)	Asmundson et al. (1994)		-0.04 ns		
Ability	**RCFT**					

	Copy	Boldrini et al. (2005)	-2.98 ↑			
Set shifting	TMT B	Airaksinen et al. (2005)	0.07 ns	0.31 ns	0.22 ns	0.17 ns
		Asmundson et al. (1994)		-0.35 ns		
	WCST					
	categories	Boldrini et al. (2005)	0.08 ns			
	total errors	Boldrini et al. (2005)	0.93 ↓			
	perseverative errors	Boldrini et al. (2005)	-0.55 ns			
	non-p errors	Boldrini et al. (2005)	0.95 ↓			
	perseverative responses	Boldrini et al. (2005)	-0.55 ns			
	null sorts	Boldrini et al. (2005)	-3.14 ↑			
	IDS trial score	Purcell et al. (1998)	-0.50 ns			
	EDS trial score	Purcell et al. (1998)	-0.14 ns			
	Spatial working memory	Purcell et al. (1998)	-0.05 ns			
Planning and						
Organising	Tower of London	Purcell et al. (1998)	-0.43 ↑			
Verbal fluency	FAS in 60 sec	Airaksinen et al. (2005)	-0.35 ns	-0.36 ns	0.08 ns	0.18 ns
	FAS in 60 sec	Boldrini et al. (2005)	-0.63 ns			
Decision making	Iowa Gambling Task	Cavedini et al. (2002)	↓			

↓Indicates statistically significantly worse performance than comparison group

↑Indicates statistically significantly better performance than comparison group

ns Indicates no significant difference compared to comparison group

OCD = Obsessive compulsive Disorder, SocP= Social Phobia, GAD = Generalised Anxiety Disorder, SpP = Specific Phobia

Table 10 Neuropsychological performance in PD compared to depression

Cognitive function	Test	study/authors	PD v PD+MDD	PD v MDD	d
Visual Memory	**CANTAB**				
	Spatial Recognition Memory	Kaplan et al. (2006)	-		0.15
	Pattern Recognition Memory	Kaplan et al. (2006)	-		0.04
	Delayed Match to Sample	Kaplan et al. (2006)	-		-0.03
	Spatial Recognition Memory	Purcell et al. (1998)		-	-0.03
	Pattern Recognition Memory	Purcell et al. (1998)		-	0.27
	Delayed Match to Sample	Purcell et al. (1998)		-	0.03
Working Memory	**CANTAB**				
	Spatial Span	Kaplan et al. (2006)	-		0.23
	Spatial Span	Purcell et al. (1998)		↓	0.66
Attention	**CANTAB** Rapid Visual				
	Information Processing	Kaplan et al. (2006)	-		0.06
Perception	Signal Detection	Lautenbacher et al. (2001)		-	-0.07
	Visual Field Neglect task	Lautenbacher et al. (2001)		-	0.30
Set Shifting	**CANTAB** Intradimensional-				
	Extradimensional Shift				
	Stages	Kaplan et al. (2006)	-		0.10
	Adjusted errors	Kaplan et al. (2006)	-		-0.12
	IDS trial score	Purcell et al. (1998)		↑	-0.73
	EDS trial score	Purcell et al. (1998)		-	-0.33
	Spatial working memory	Purcell et al. (1998)		-	-0.11

Planning and organising	Tower of London	Purcell et al. (1998)		-	0.01
Decision making	Cambridge Gambling Task	Kaplan et al. (2006)	-		u/a
	Two-Choice Prediction Task	Ludewig et al. (2003)		-	u/a

PD+MDD = patient group diagnosed with both PD and Major Depressive Disorder

↓Indicates statistically significantly worse performance in PD only than MDD or PD+MDD group

↑Indicates statistically significantly better performance in PD only than MDD or PD+MDD group

- Indicates no significant difference compared to MDD or PD+MDD group