Is suboptimal effort an issue? A systematic review on neuropsychological performance validity in major depressive disorder

Legemaat, Amanda M.; Haagedoorn, Marcella A.S.; Burger, Huibert; Denys, Damiaan; Bockting, Claudi L.; Geurtsen, Gert J.

Published in:
Journal of Affective Disorders

DOI:
10.1016/j.jad.2022.12.043

Publication date:
2023

Journal of Affective Disorders 323 (2023) 731–740

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Review Article

Is suboptimal effort an issue? A systematic review on neuropsychological performance validity in major depressive disorder

Amanda M. Legemaat [a], Marcella A.S. Haagedoorn [c], Huibert Burger [e], Damiaan Denys [a], Claudi L. Bockting [a,d], Gert J. Geurtsen [b,*]

[a] Department of Psychiatry Amsterdam University Medical Centers, Location AMC, University of Amsterdam, Amsterdam Neuroscience & Amsterdam Public Health, Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands
[b] Department of Medical Psychology Amsterdam University Medical Centers, Location AMC, University of Amsterdam, Amsterdam Neuroscience & Amsterdam Public Health, Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands
[c] Department of Geriatric Psychiatry, Mental Health Care North-Holland North, Maelsonstraat 1, 1624 NP Hoorn, the Netherlands
[d] Centre for Urban Mental Health, University of Amsterdam, Oude Turfmarkt 147, 1012 GC Amsterdam, the Netherlands
[e] Department of General Practice and Elderly Care Medicine, University Medical Center Groningen, University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, the Netherlands

ARTICLE INFO

Keywords:
Performance validity
Symptom validity
Major depressive disorder
Depression
Cognitive impairment
Neuropsychological assessment

ABSTRACT

Background: In Major Depressive Disorder (MDD), emotion- and motivation related symptoms may affect effort during neuropsychological testing. Performance Validity Tests (PVT's) are therefore essential, but are rarely mentioned in research on cognitive functioning in MDD. We aimed to assess the proportion of MDD patients with demonstrated valid performance and determine cognitive functioning in patients with valid performance. This is the first systematic review on neuropsychological performance validity in MDD.

Methods: Databases PubMed, PsycINFO, Embase, and Cochrane Library were searched for studies reporting on PVT results of adult MDD patients. We meta-analyzed the proportion of MDD patients with PVT scores indicative of valid performance.

Results: Seven studies with a total of 409 MDD patients fulfilled inclusion criteria. Six studies reported the exact proportion of patients with PVT scores indicative of valid performance, which ranged from 60 to 100 % with a proportion estimate of 94 %. Four studies reported on cognitive functioning in MDD patients with valid performance. Two out of these studies found memory impairment in a minority of MDD patients and two out of these studies found no cognitive impairment.

Limitations: Small number of studies and small sample sizes.

Conclusions: A surprisingly small number of studies reported on PVT in MDD. About 94 % of MDD patients in studies using PVT's had valid neuropsychological test performance. Concessive information regarding cognitive functioning in MDD patients with valid performance was lacking. Neuropsychological performance validity should be taken into account since this may alter conclusions regarding cognitive functioning.

1. Introduction

During the past decade, numerous studies have been published on cognitive functioning in Major Depressive Disorder (MDD) and systematic reviews and meta-analyses concluded that MDD is associated with cognitive impairment (Ahern and Semkovska, 2017; Rock et al., 2014; Snyder, 2013). In this area of research, as well as in subsequent research on etiology and treatment of MDD-related cognitive

Abbreviations: MDD, Major Depressive Disorder; PVT, Performance Validity Test; DSM, Diagnostic and Statistical Manual of Mental Disorders; ICD, International Classification of Diseases; HDRS, Hamilton Depression Rating Scale; MADRS, Montgomery–Åsberg Depression Rating Scale; BDI, Beck Depression Inventory; GDS, Geriatric Depression Scale; ASTM, Amsterdam Short Term Memory Test; TOMM, Test of Memory Malingering; VSVT, Victoria Symptom Validity Test; WMT, Green's Word Memory Test; Rey DCT, Rey Dot Counting Test.

* Corresponding author at: Department of Medical Psychology, Amsterdam University Medical Centers, Location Academic Medical Center, Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands.
E-mail address: g.j.geurtsen@amsterdamumc.nl (G.J. Geurtsen).

https://doi.org/10.1016/j.jad.2022.12.043
Received 26 January 2022; Received in revised form 25 November 2022; Accepted 11 December 2022
Available online 14 December 2022

impairment (Jamieson et al., 2019; Legemaat et al., 2021; Listunova et al., 2018; Lundbeck, 2018; McIntyre et al., 2014; Mahableshwarkar et al., 2015; Miskowiak et al., 2016; Motter et al., 2016; Thérond et al., 2021), conclusions are based on the presupposition that the assessment of cognitive functioning in MDD is valid.

Cognitive functioning in domains such as memory, executive functioning, attention and processing speed is determined by neuropsychological assessment. Test scores that clearly deviate from normative data indicate cognitive impairment. But test scores are only representative of cognitive functioning, if patients perform to the best of their cognitive ability: i.e. provide adequate effort. MDD patients may, however, not provide adequate effort (Roiser and Sahakian, 2013). Emotion and motivation related symptoms of depression may prevent them to perform to the best of their cognitive ability. For example, depressed mood, loss of interest, and fatigue can hinder a patient to direct attention to relevant information. Rumination and negative beliefs (e.g. 'I will not be able to do this') associated with MDD can further influence focus and diligence.

As a consequence, MDD patients may not be able to put forth adequate effort required for a valid determination of cognitive functioning. Therefore, impaired neuropsychological test scores do not necessarily equal cognitive impairment (Benitez et al., 2011; Joormann and Gotlib, 2008; Moritz et al., 2017; Scheurich et al., 2008; Whitmer and Gotlib, 2012). While it may not make much difference from a patients perspective whether impaired neuropsychological test scores are a result of emotion and motivation related factors or actual cognitive impairment, it may be crucial to understand the exact nature of the problem: both for prognosis as for the development and provision of adequate treatment options.

As clinicians cannot reliably differentiate between valid and invalid neuropsychological test performance based on their clinical impression (Dandachi-FitzGerald et al., 2017), Performance Validity Tests (PVT's) have been developed (Bush et al., 2005; Greher and Wodushek, 2017; Wodushek and Greher, 2017). PVT's can determine whether neuropsychological test scores are representative of cognitive functioning, or are affected by inadequate effort (Greher and Wodushek, 2017). PVT's appear to be similar to other neuropsychological tests, but require a minimal amount of effort in order to obtain a score above a standardized cut-off. Even patients with a neurodegenerative disease, for example mild cognitive impairment or even early dementia, have been shown to pass the standardized cut-off scores (Rienstra et al., 2013). Patients can either pass or fail a PVT. Passing a PVT indicates that neuropsychological test scores are valid, while failing indicates they are invalid. PVT's are considered the gold standard to assess the validity of neuropsychological test performance in clinical and research practice. Moreover, performance validity testing is deemed a crucial aspect of neuropsychological assessment and a prerequisite for diagnosing cognitive impairment (Bush et al., 2005; Heilbronner et al., 2009; McCrea et al., 2008).

Notably, PVT's are rarely employed in research on cognitive functioning in MDD. Neuropsychological test performance is generally assumed to be valid in these studies (Ahern and Semkovska, 2017; Rock et al., 2014; Snyder, 2013). Without the utilization of PVT's, this remains, however, just that – an assumption. This methodological gap in the majority of research on cognitive functioning in MDD, raises questions about the validity of findings. Without the utilization of PVT's, it remains uncertain whether findings are affected by performance invalidity. Remarkably, Gass and Patten (2020) as well as Benitez et al. (2011) report *no* cognitive impairment in depressed patients with demonstrated valid test performance. Further, other studies that administered a PVT showed that in part of their MDD sample, neuropsychological test performance was invalid (Lee et al., 2000; Yanez et al., 2006). However, there is conflicting evidence as another study found 100 % valid performance in their sample of MDD patients (Rees et al., 2001).

Determining whether neuropsychological test performance is valid, is essential to understand the nature and extent of cognitive impairment in MDD and thus required in order to develop and provide adequate treatment. A systematic overview of the literature on this topic is lacking. This review therefore aims to assess the evidence on neuropsychological performance validity in MDD. More specifically, the primary aim is to review the literature on the proportion of MDD patients with demonstrated valid test performance. The secondary aim is to assess cognitive functioning in MDD patients with valid test performance, and if and how cognitive functioning differs from MDD patients with invalid test performance.

2. Methods

2.1. Search strategy and selection criteria

We conducted a systematic review following the PRISMA guidelines. Databases PubMed, PsycINFO, Embase, and Cochrane Library were searched using key words and MeSH terms related to neuropsychological performance validity, PVT's, and MDD (Supplementary material, Appendix I). The search was conducted on the 4th of June 2021 from the beginning of records. References and citations of included studies were hand-searched for additional relevant studies.

Included were original studies concerning adults (≥18 year); with a primary diagnosis of MDD according to the Diagnostic and Statistical Manual of Mental Disorders (DSM) or International Classification of Diseases (ICD) criteria, or at least moderate depressive symptoms according to the Hamilton Depression Rating Scale-17 (HDRS-17 ≥ 16 (Bech et al., 1986; Hamilton, 1960)), the Montgomery–Åsberg Depression Rating Scale (MADRS ≥20 (Mittmann et al., 1997; Montgomery and Åsberg, 1979)), the Beck Depression Inventory-II (BDI-II ≥ 20 (Beck et al., 1996)), or any other scale that aims to assess MDD symptoms; and reporting outcomes of at least one validated PVT. We searched for peer-reviewed publications in English or Dutch language only. Excluded were studies on MDD patients with comorbid neurological disease (e.g. dementia, Parkinson, cerebrovascular accident), primary chronic physical disease (e.g. fibromyalgia), or psychotic disorder.

2.2. Screening procedure

After removing duplicates, records were screened in a two-stage procedure. First, AML and GJG independently screened titles and abstracts and selected studies with potential for inclusion. Second, selected articles were assessed full-text for eligibility by AML and judgements were cross-checked by GJG. Any disagreements were resolved through consensus.

2.3. Data extraction

Extracted data included sample characteristics and major in- and exclusion criteria, information regarding the PVT's administered, and PVT results. If applicable, information regarding neuropsychological tests administered, neuropsychological test results, characteristics of unaffected controls, and other relevant findings were extracted as well. When publications did not report on data relevant to the current review, study authors were contacted and reminded once to provide the missing information.

2.4. Meta-analysis

We meta-analyzed the proportion of MDD patients with PVT scores indicative of valid performance across studies. Only studies that provided the exact number of patients with PVT scores indicative of valid performance, were included in the meta-analysis as well as the range description of the proportion with valid performance.

As the weights of studies with proportions close to zero or one are overestimated in inverse variance pooling methods, we first stabilized

the variances by applying the Freeman-Tukey double arcsine transformation (Freeman and Tukey, 1950). Next, a pooled estimate of the transformed values and their variances was obtained using the random effects model according to DerSimonian and Laird (1986). We used a random effects model in view of the a priori assumed high heterogeneity of studies. Finally, the pooled value with 95 % confidence interval was transformed back to the proportion scale. We calculated the I^2 statistic as an indicator of heterogeneity among studies.

3. Results

3.1. Study selection and data collection process

The systematic search identified 2590 records. After removing duplicates, 1584 titles and abstracts were screened and 28 articles were retrieved for full-text assessment. Two additional studies were identified through snowballing. The screening procedure resulted in the inclusion of a total of seven studies (Fig. 1) (Beblo et al., 2020; Benitez et al., 2011; Egeland et al., 2003; Lee et al., 2000; Rees et al., 2001; Rowland et al., 2017; Yanez et al., 2006).

3.2. Study characteristics

All studies reported on patients with diagnosed MDD. Sample size ranged from 22 to 127 and mean age ranged from 35.1 to 68.7 years. The included studies had a total of 409 MDD patients. Five studies included only patients with MDD according to DSM or ICD criteria (Beblo et al., 2020; Egeland et al., 2003; Lee et al., 2000; Rees et al., 2001; Rowland et al., 2017) and four of these studies confirmed MDD with a semi-structured interview (Beblo et al., 2020; Egeland et al., 2003; Lee et al., 2000; Rowland et al., 2017). One study included a sample with MDD according to DSM criteria and a second sample with MDD according to the Geriatric Depression Rating Scale (GDS, Benitez et al., 2011). Another study included patients with a clinical MDD diagnosis and confirmed severity with the BDI-II (Yanez et al., 2006). None of the studies mentioned circumstances beyond MDD diagnosis that indicated initial performance validity concerns. In one study, any initial validity concerns were an explicit exclusion criterion (Rowland et al., 2017) and these were operationalized as bias according to the Slick criteria (A. Presence of a substantial external incentive for exaggeration/fabrication of symptoms; B. Evidence from neuropsychological testing of exaggeration/fabrication of cognitive impairment; C. Evidence from self-report involving inconsistencies/discrepancies suggesting a deliberate attempt to exaggerate/fabricate cognitive impairment; D. criteria B/C are not accounted for by psychiatric, neurological, or developmental factors, Slick et al., 1999). Two studies included exclusively veterans (Benitez et al., 2011; Rowland et al., 2017).

PVT's used were the Amsterdam Short Term Memory Test (ASTM, Beblo et al., 2020); the Test of Memory Malingering (TOMM, Benitez et al., 2011; Rees et al., 2001; Yanez et al., 2006); the Victoria Symptom Validity Test (VSVT, Egeland et al., 2003); the Green's Word Memory Test (WMT, Rowland et al., 2017); the Rey 15-item Memorization (Lee et al., 2000); and the Rey Dot Counting Test (Rey DCT, Lee et al., 2000). One study used two PVT's (Lee et al., 2000), all other studies used one PVT. Six studies reported the exact number of MDD patients with PVT scores indicative of valid performance (Beblo et al., 2020; Egeland et al.,

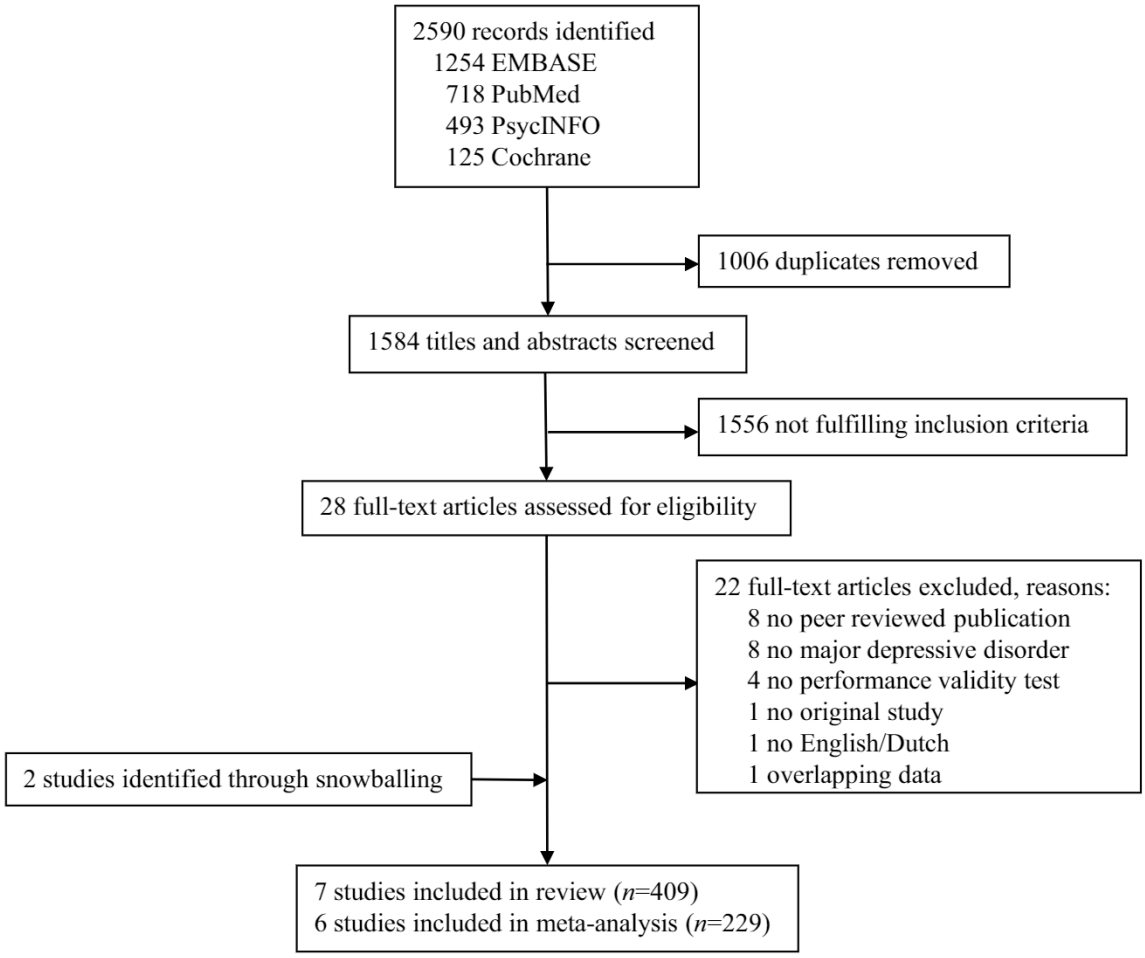

Fig. 1. PRISMA flow diagram.

2003; Lee et al., 2000; Rees et al., 2001; Rowland et al., 2017; Yanez et al., 2006). One study did not report the exact number of patients with PVT scores indicative of valid performance and these data were not available upon request (Benitez et al., 2011). Therefore, this study was excluded from the range description and meta-analysis of the proportion of MDD patients with valid performance. Further, six studies mentioned additional neuropsychological assessment but only three provided information on cognitive functioning of MDD patients with valid performance (Benitez et al., 2011; Egeland et al., 2003; Rees et al., 2001). Authors of one study provided these data upon request (Beblo et al., 2020). Six studies compared PVT or neuropsychological test results of the MDD sample with an unaffected control sample (Beblo et al., 2020; Benitez et al., 2011; Egeland et al., 2003; Rees et al., 2001; Rowland et al., 2017; Yanez et al., 2006). An overview of the characteristics of the included studies and main findings are presented in Table 1.

3.3. Main findings

3.3.1. Performance validity assessment

In the six studies ($n = 229$) that reported the exact number of MDD patients with PVT scores indicative of valid performance the proportion with valid performance ranged from 60 to 100 %. Meta-analysis of the proportion with PVT scores indicative of valid performance across these six studies, resulted in a proportion estimate of 94 % (0.94, *CI* 0.79–1.00) with valid performance. Heterogeneity was large ($I^2 = 90$ %).

In three of the included studies ($n = 98$), all MDD patients passed the performance validity assessment (Beblo et al., 2020; Egeland et al., 2003; Rees et al., 2001). In Rees et al. (2001), all MDD patients exceeded the criterion for valid performance, including those who received ECT treatment during, or before the study ($n = 10/26$). Further, Rees et al. (2001) found no significant effects of level of depression on PVT performance.

In the other four of the included studies ($n = 311$), part of the MDD patients failed at least one PVT (Benitez et al., 2011; Lee et al., 2000; Rowland et al., 2017; Yanez et al., 2006). In Benitez et al. (2011) 63 % of a sample of MDD patients passed their performance validity criteria. Notably, these criteria required the absence of indicators of data invalidity based on clinical judgement *in addition* to TOMM scores indicative of valid performance. In a seconds sample in the same study, 73 % of patients with mild, and 49 % of patients with severe depression based on GDS scores passed these same performance validity criteria. In comparison, 93 % and 74 % of unaffected controls in the first and second sample, respectively, passed the performance validity criteria. Lee et al. (2000) found that 89 up to 95 % of MDD patients passed the Rey 15-item Memorization, depending on the criterion used. All of the patients passed the second PVT administered, the Rey DCT. Further, Lee et al. (2000) found no significant differences in PVT scores based on depression severity level, or any significant correlation between HDRS and PVT scores. In Rowland et al. (2017) 60 % of MDD patients passed the WMT. They found that a significant higher proportion of psychiatric patients (including MDD patients) failed the PVT, even in the absence of initial performance validity concerns (Slick et al., 1999). Moreover, in another publication based on the same sample, depression diagnosis was a significant predictor of PVT failure (Shura et al., 2017). In Yanez et al. (2006), respectively 90 up to 95 % of MDD patients passed the TOMM, depending on whether the criterion was set at the commonly used Trial 2, or the Retention trial. In contrast, in the unaffected control sample all participants passed the TOMM.

3.3.2. Neuropsychological assessment

Four studies ($n = 278$) reported on cognitive functioning in MDD patients who passed performance validity criteria (Egeland et al., 2003; Beblo et al., 2020; Benitez et al., 2011; Rees et al., 2001). In one of these four studies, only part of the patients passed performance validity criteria. The authors reported that outcomes changed based on whether they included in their analyses the total sample, or only the sub-sample that passed performance validity criteria (Benitez et al., 2011). In the other three studies, all patients passed performance validity criteria. None of the studies included in this review compared cognitive functioning in MDD patients with valid versus invalid performance.

Benitez et al. (2011) found that significant differences in neuropsychological test scores (*Repeatable Battery for the Assessment of Neuropsychological Status* [RBANS]) between MDD patients and unaffected controls, disappeared when only participants who passed performance validity criteria were included. In the valid neuropsychological assessment sample, cognitive impairments were no longer present. Further, they found no significant differences in neuropsychological test scores between participants with GDS scores in the normal, mild, and severe range, irrespective of including all participants or only those who passed performance validity criteria.

Egeland et al. (2003) reported that in their 100 % valid sample, 4 up to 33 % of MDD patients – depending on the specific neuropsychological test – performed below the lenient criterion for impairment of the 16 % percentile compared to normative data. On group level, the MDD sample performed significantly worse on two out of thirteen single test scores (*Paced Auditory Serial Addition Test* [PASAT] and *Rey Complex Figure Test-Long Term Memory* [RCFT-LTM]) and on a composite measure of delayed memory. Further, 28 % of the MDD sample scored below the 16th percentile on three or more memory tests. With regard to impairment pattern, the authors found distinct impairment in working memory and unguided retrieval, and relatively intact acquisition and recognition memory.

These findings are in line with Beblo et al. (2020), who reported that they found no significant differences on memory recognition between MDD patients with 100 % valid performance and unaffected controls, while they did find unaffected controls significantly outperformed the MDD patients on delayed memory recall. Further, when impairment was operationalized as scores below the 16th percentile compared to normative data, 14 % of MDD patients scored at impairment level on recognition, and 45 % on delayed memory recall. When the more common criterion for impairment of scores below the 5th percentile was utilized, no MDD patients scored at impairment level on recognition, and 14 % scored at impairment level on delayed memory recall (Beblo et al., 2020). In Rees et al. (2001) the total MDD sample had demonstrated valid test performance and none scored at impairment level on a commonly used screener for cognitive impairment (Mini-Mental State Examination [MMSE]).

4. Discussion

4.1. Main findings

This review aimed to assess the evidence on neuropsychological performance validity in MDD. We reviewed the literature on the proportion of MDD patients with demonstrated valid test performance and cognitive functioning in MDD patients with valid versus invalid test performance. Seven studies were included in the descriptive review. Six studies were included in the meta-analysis and range description as they reported on the exact proportion of MDD patients with demonstrated valid test performance, which ranged from 60 up to 100 %. Meta-analytic calculation of the proportion of MDD patients with valid test performance across these six studies, resulted in a proportion estimate of 94 %. The study not included in the range description and meta-analysis, did not report the exact number of patients with valid test performance according to PVT (Benitez et al., 2011). In this study, additional validity criteria based on clinical judgement, resulted in lower passing rates relative to the other studies; the authors reported a proportion fulfilling their validity criteria of 49 % in patients with severe depression up to 73 % in patients with mild depression. It is however questionable whether clinical judgement should be utilized to determine performance validity as it has been shown that clinicians cannot reliably differentiate between

Table 1
Study characteristics and main findings.

Study	Sample characteristics		Performance validity assessment			Neuropsychological assessment				Other	
	Population base, major in- and exclusion criteria	*n* Age *M* (*SD*) Ed. years *M* (*SD*)/ Grade % Other statistics	PVT	Criterion valid	*n* valid (%)	NP tests	Criterion impair-ment	*n* impairment (%)		Unaffected controls	Main relevant findings
								Valid sample	Invalid sample		
Beblo et al., 2020	Patients treated in specialized ward for MDD In: 18–65 years; DSM-IV and ICD-10 MDD confirmed with Mini-DIPS Ex: severe mental/physical disorder; pregnancy	$n = 22$ Age 42.5 (12.4) Ed. years 11.1 (2.0) $n = 11$ comorbid anxiety disorder $n = 20$ uses antidepressants	ASTM	≥85 correct	22 (100 %)	WMS-IV LM Delayed FR Recognition WMS-IV LM Delayed FR Recognition	≤16th percentile compared to normative data <5th percentile compared to normative data	10 (45 %) 3 (14 %) 3 (14 %) 0 (0 %)	N.a.	No Axis I disorders; 18–65 years $n = 28$ Age 42.0 (16.2) Ed. years 11.9 (1.5)	MDD vs UC: 1. No sig. difference on PVT 2. MDD sig. lower scores on Delayed FR (MDD M = 17.1 versus UC M = 23.6, $p = 0.001$) 3. No sig. difference on Recognition ($p = 0.119$) 4. No sig. correlation Delayed FR; Recogniton and ASTM in MDD ($r = 0.26$, $p = 0.232$; $r = 0.28$, $p = 0.202$) or UC ($r = 0.11$, $p = 0.572$; $r = 0.03$, $p = 0.901$)
Benitez et al., 2011	Veterans, geriatric memory clinic patients Sample 1: In: DSM-IV-TR MDD Ex: diagnosed dementia/cognitive disorder Sample 2: Mild depression: GDS > 9 Severe depression: GDS > 19 Ex: diagnosed with dementia	 Sample 1: $n = 40$ Age 63.7 (8.1) Ed. years 12.7 (3.0) Sample 2: Mild depression: $n = 95$ Age 68.7 (9.2) Ed. years 12.8 (3.3) Severe depression: $n = 45$ Age 63.3 (9.4) Ed. years 12.5 (3.0)	TOMM	Both ≥45 correct on 2nd and retention trial and no other indicators of data invalidity[a]	 Sample 1: 25 (63 %) Sample 2: Mild dep.: 69 (73 %) Severe dep.: 22 (49 %)	RBANS Imm. Memory Visuosp./ Constr. Language Attention Del. Memory Total MMSE	RBANS mean index scores ≤ 'borderline'	0 (0%) 0 (0 %) 0 (0 %) 0 (0 %) 0 (0 %) 0 (0 %) N.r	N.r.	Veterans, geriatric memory clinic patients Sample 1: No psychiatric or cognitive disorder $n = 58$ Age 70.2 (8.9) Ed. years 13.6 (3.1) Sample 2: Normal: GDS ≤ 9; no dementia $n = 168$ Age 73.4 (8.1) Ed. years 12.2 (3.4)	Sample 1 MDD vs UC: 1. No sig. difference (F (6, 71) = 2.12; $p = 0.062$) on RBANS when only the valid sample was included; MDD and UC mean index scores were both low average to average 2. MDD sig. lower scores on 4/5 RBANS indexes when total MDD sample was included (F (1, 96) = 11.52–24.91; $p \leq 0.001$–0.001); MDD mean index scores were borderline to low average, UC mean index scores were low average to average 3. 54/58 (93 %) of UC's met criteria for valid performance Sample 2: 1. No sig. differences on RBANS between normal, mild, and severe GDS, irrespective of including total (F (12, 580) = 1.64, $p = 0.077$) or only the valid sample (F (12, 410) = 0.99, $p = 0.456$) 2. 125/168 (74 %) of UC's (normal GDS) met criteria for valid performance
Egeland et al., 2003	Out- and inpatients between 19 and 51 years In: recurrent DSM-IV MDD confirmed with SCID-I Ex: head trauma, neurological/ developmental disorder, alcohol/ substance abuse, medical disease	$n = 50$ Age 35.1 (8.7) Ed. years 13.9 (2.9) $n = 5$ psychotic symptoms	VSVT	≥13 correct on the first 16 items presented with 5 s. intervals	50 (100 %)	PASAT DS Backward DS Forward RCFT LTM RCFT Rec. RMT Faces CVLT Total	<16th percentile compared to UC control group	33 % 22 % 16 % 43 % 21 % 20 % 26 % 26 %	N.a.	$n = 50$ Age 32.9 (9.0) Ed. years 13.9 (2.5)	MDD vs UC: 1. MDD sig. lower scores on PASAT (MDD $M = 79.3$, UC $M = 95$, $p \leq 0.001$), RCFT-LTM (MDD $M = 20.5$, UC $M = 24.5$, $p \leq 0.001$) and C. LDFR (MDD $M = -0.56$, UC $M = 0.0$, $p \leq 0.001$) 2. 28 % of MDD sample scored

(*continued on next page*)

Table 1 (*continued*)

Study	Sample characteristics		Performance validity assessment			Neuropsychological assessment				Other	
	Population base, major in- and exclusion criteria	*n* Age *M* (*SD*) Ed. years *M* (*SD*)/ Grade % Other statistics	PVT	Criterion valid	*n* valid (%)	NP tests	Criterion impair-ment	*n* impairment (%)		Unaffected controls	Main relevant findings
								Valid sample	Invalid sample		
	likely to affect nervous system, history of hypomania	$n = 46$ uses antidepressants				CVLT LDFR CVLT LDCR CVLT Discr. RMT Words C. LDFR, *z* Rec. mem., *z* IQ-estimate		28 % 4 % 24 % 32 % 21 % 31 %			<16th percentile on ≥3 memory tests
Lee et al., 2000	Middle-aged and elderly individuals recruited for a research project on late life depression; no participant was in litigation or attempting to obtain disability payments In: DSM-III-R MDD confirmed with SCID, Ex: psychotic symptoms; history of (hypo)mania; drug/alcohol abuse; stroke; epilepsy; Parkinson's disease; hemiparesis/hemisensory deficits	$n = 64$ Age 59.9 (7.9) Ed. years 15.0 (2.7)	Rey 15-item Memori-zation Rey DCT	≥9 total or spatial correct or ≥8 adjusted spatial correct M ungrouped time > M grouped time; ≤3 errors; ungrouped time ≤ 180″; grouped time ≤ 130″	61 (95 %) 57 (89 %) 64 (100 %)	N.a.	N.a.	N.a.	N.a.	N.a.	1. No sig. differences in PVT scores between mild, moderate and severely depressed groups ($p \geq 0.29$) 2. No sig. correlations between HDRS and PVT scores ($p \geq 0.20$)
Rees et al., 2001	Inpatients at affective disorder's unit In: DSM-IV MDD Ex: psychotic symptoms; significantly impaired attention (no exclusions based on these criteria)	$n = 26$ Age 40.4 (11.2) Ed. years 14.9 (2.8) $n = 26$ uses antidepressants $n = 4$ current ECT $n = 10$ ECT in the past (including those receiving current ECT)	TOMM	≥45 correct on 2nd/retention trial	26 (100 %)	MMSE	<25	0 (0 %)	N.a.	Previously collected data (Rees, 1996; Tombaugh, 1997) $n = 26$	MDD vs UC: 1. Exact same group means on Trial 1, Trial 2 and retention trial 2. No sig. differences in PVT scores 3. No sig. effect on PVT of level of depression ($F(2,23) = 0.04$, $p > 0.05$) 4. No effect on 2nd/retention trial of ECT (all 4 ECT patients had a score of 100 % by trial 2)
Rowland et al., 2017	Veterans, repository participants In: DSM-IV MDD confirmed with SCID Ex: combat exposure prior to 1985; neuropsychological evaluation in prior 6 months; psychosis; substance abuse/ dependence; pre-deployment/ non-military PTSD; TBI greater than mild; never-deployed; missing WMT scores; any apparent incentives to perform poorly/bias according to Slick criteria (Slick et al., 1999)	$n = 47$ Age n.r. Ed. years n.r. $n = xx$ comorbid PTSD $n = xx$ comorbid mTBI	WMT	>82.5 % on IR, DR, and CNS	28 (60 %)	UPSIT FTT GP WAIS-III Block DSC SS LNS SIM RCFT CPT-II TMT Stroop CWT PASAT COWAT WCST	N.r.	N.r.	N.r.	Veterans, repository participants; same exclusion criteria as MDD sample $n = 188$ had no MDD; $n = 115$ had not any psychiatric diagnosis according to SCID, DSM-IV	MDD vs UC: 1. Sig. higher proportion of psychiatric patients (including MDD) failed PVT ($\chi^2 = 6.56$, $p = 0.018$, $\varphi = 0.192$) In another publication, based on the same sample (Shura et al., 2017): 1. In valid sample: Self-reported depressive symptoms were sig. associated with slower processing speed on 7/10 measures ($p < 0.001$–0.031) 2. Depression diagnosis was a sig. predictor of PVT failure ($\chi^2 =$

(*continued on next page*)

Table 1 (*continued*)

Study	Sample characteristics		Performance validity assessment			Neuropsychological assessment				Other	
	Population base, major in- and exclusion criteria	*n* Age *M* (*SD*) Ed. years *M* (*SD*)/ Grade % Other statistics	PVT	Criterion valid	*n* valid (%)	NP tests	Criterion impair-ment	*n* impairment (%)		Unaffected controls	Main relevant findings
								Valid sample	Invalid sample		
						CVLT-II BVMT-R					18.611, $p \leq 0.001$, $ES = 0.289$) 3. Depression diagnosis was no sig. predictor of any other neuropsychological measure ($p \geq 0.060$) 4. PVT failure was sig. associated with slower processing speed on 8/10 measures ($p < 0.001$–0.034)
Yanez et al., 2006	Outpatient at a clinician's office who finished a Social Security Disability evaluation In: BDI-II ≥ 30	$n = 20$ Age 39.1 (8.9) Ed. years 11.2 (1.6) $n = 20$ MDD diagnosis $n = 19$ use antidepressants $n = 13$ previously hospitalized for depression	TOMM	≥45 correct on 2nd trial on retention trial	18 (90 %) 19 (95 %)	21-item test WMS-III LM	N.r.	N.r.	N.r.	Recruited friends and family of MDD sample; none took psychotropic medication/ received psychiatric care; BDI-II ≤ 20 $n = 20$ Age 41.7 (10.8) Ed. years 13.3 (2.8)	MDD vs UC: 1. No sig. differences between MDD and UC on PVT (F (1, 38) = 3.95, $p < 0.06$) 2. No UC's had an PVT scores below cut-off

Abbreviations: Ed. = Education; PVT = Performance Validity Test; NP=NeuroPsychological; MDD = Major Depressive Disorder; dep. = depression; PTSD=Post Traumatic Stress Disorder; TBI = Traumatic Brain Injury; N.a. = Not applicable; N.r. = Not reported; sig. = significant. Clinical instruments, in alphabetical order: ASTM = Amsterdam Sort Term Memory test; BVMT = Brief Visuospatial memory Test; C. LDFR = Composite measure of CVLT-LDFR and RCFT-LTM; COWAT = Controlled Oral Word Association Test; CPT = Conners' Continuous Performance Test; CVLT = California Verbal Learning Test*; Del. = Delayed;* Discr. = correctly discriminated responses recognition trial, Discriminability; DS=Digit Span; DSC=Digit-Symbol Coding; DSM = Diagnostic and Statistical Manual of Mental Disorders; FR = Free Recall; FTT = Finger Tapping Test; GDS = Geriatric Depression Scale; GP = Grooved Pegboard test; ICD=International Classification of Diseases*; Imm. = Immediate;* IQ-estimate = Intelligence Quotient estimated from Similarities and Picture Completion subtests from WAIS-R; LDCR = Long Delay Cued Recall; LDFR = Long Delay Free Recall; LM = Logical Memory; LNS = Letter Number Sequencing; *LTM = Long Term Memory;* Mini-DIPS = Mini Diagnostic Interview for Mental Disorders; MMSE = Mini-Mental State Examination; *PASAT = Paced Auditory Serial Addition Test*; R = Revised; RBANS = *Repeatable Battery for the Assessment of Neuropsychological Status;* RCFT = *Rey Complex Figure Test;* Rec. = Recognition; Rec. mem. = mean z score of CVLT-discriminability and RMT Faces; Rey DCT = Rey Dot Counting Test; RMT = Recognition Memory Test; SCID = Structured Clinical Interview for the DSM; SIM = Similarities; SS=Symbol Search; Stroop CWT = Stroop Color and Word Test; TMT = Trail Making Test; TOMM = Test of Memory Malingering; TR = Text Revision; UC=Unaffected Controls; UPSIT = University of Pittsburg Smell Identification Test; *Visuosp./Constr. = Visuospatial/Constructive;* VSVT = Victoria Symptom Validity Test; WAIS=Wechsler Adult Intelligence Scale; WCST-64 = Wisconsin Card Sort Test-64; WMS=Wechsler Memory Scale; WMT = Green's Word Memory Test.

[a] Indicators of data invalidity used: unusual error patterns and/or impairment on testing deemed disproportionate to functional level or inconsistent with knowledge of brain functioning.

valid and invalid neuropsychological test performance (Dandachi-Fitz-Gerald et al., 2017). With regard to the impact of severity of depressive symptoms, Benitez et al. (2011) found that a higher percentage of patients with mild depression passed their performance validity criteria, in comparison to patients with severe depression. Lee et al. (2000) and Rees et al. (2001), in contrast, found no significant effects of level of depression on PVT scores. Benitez et al. (2011) found no significant differences in neuropsychological test performance between patients with different severity levels of depression. Other reviews and meta-analyses on cognitive impairment in MDD, found mixed results regarding whether higher depression severity is associated with greater cognitive impairment (Ahern and Semkovska, 2017; McDermott and Ebmeier, 2009; Porter et al., 2007).

This review shows that neuropsychological performance validity in MDD varies across as well as within samples. Even studies using the exact same PVT and criterion for valid performance, found different results. However, some PVT's appeared to yield higher passing rates than others. The ASTM, TOMM, VSVT, Rey 15-item test, and Rey DCT demonstrated valid performance in 87 up to 100 % of MDD patients, while the WMT demonstrated valid performance in 60 up to 67 %. With regard to sensitivity, numerous studies have shown that the TOMM was inferior to other PVT's, a.o. to the ASTM, VSVT and WMT, to detect performance invalidity (Bolan et al., 2002; Gervais et al., 2004; Greiffenstein et al., 2008; van Hout et al., 2003; Tan et al., 2002). Compared to other PVT's, the WMT appears to be the most sensitive (Tan et al., 2002). On the other hand, with regard to specificity, it has been shown that the TOMM was superior to a.o. the ASTM and WMT to detect performance invalidity (Gervais et al., 2004; Merten et al., 2007). The ASTM has been shown to be superior to the Rey 15-item test and Rey DCT in terms of discriminative value (Schmand et al., 1999). Reports of sensitivity and specificity in non-demented adults in the three most commonly used PVT's are, respectively, 75–100 % and 89–100 % for the ASTM (Schmand et al., 1999); 80–100 % and 94–100 % for the TOMM (Greve et al., 2006; Rees et al., 1998; Tan et al., 2002; Tombaugh, 1996, 1997); 99–100 % and 99–100 % for the WMT (Brockhaus and Merten, 2004; Green, 2005; Green et al., 2002; Merten et al., 2004; Tan et al., 2002), which suggests that the WMT has the best sensitivity and specificity overall. Further, one of the studies included in the current review used multiple PVT's within the same sample, and reported results varied depending on the specific PVT administered. This is in line with the evidence that the use of multiple PVT's improves sensitivity and specificity to detect performance invalidity (Lippa, 2018; Wodushek and Greher, 2017; Greher and Wodushek, 2017).

With regard to cognitive functioning in MDD patients with valid test performance, we found mixed results. Two small studies ($n = 72$) found cognitive impairment, mainly in working memory and unguided retrieval (Beblo et al., 2020; Egeland et al., 2003). Two other studies found no cognitive impairment in MDD patients with valid test performance ($n = 161$) (Benitez et al., 2011; Rees et al., 2001). Further, one of these studies found that there was no longer evidence of cognitive impairment in MDD patients after invalid performers were excluded. Hence, whether or not performance validity was taken into account led to opposite conclusions regarding cognitive impairment (Benitez et al., 2011). None of the studies included in this review compared cognitive functioning between MDD patients with valid versus invalid performance.

Given the abundant literature on cognitive functioning in MDD, the small number of studies that assessed performance validity by administering a PVT and thus qualified for inclusion in the current review, is striking. That one of the included studies showed that conclusions may be completely different when performance validity is taken into account (Benitez et al., 2011), makes it all the more problematic that the lion's share of literature on cognitive functioning in MDD did not. Notably, performance validity may be important to take into account in psychiatric disorders in general. In Schizophrenia, performance invalidity and negative symptoms have been shown to predict findings of cognitive impairment (Strauss et al., 2015). Moreover, a recent meta-analysis on performance validity in psychotic disorders indicates that about 18 % of patients with a psychotic disorder has invalid performance; and that there is an association between PVT failure and neuropsychological test results (Ruiz et al., 2020). A limited number of studies have taken performance validity into account in other psychiatric disorders and systematic reviews and meta-analyses are lacking.

The issue of effort in neuropsychological test performance in MDD relates to the interplay of 'cold' and 'hot' cognition. Cold cognition refers to cognitive functioning that is independent of affect and emotions (Roiser and Sahakian, 2013). Neuropsychological tests aim to assess cold cognition. Hot cognition on the other hand, refers to cognitive functioning that is dependent of affect and emotions. This includes the impact of emotion and motivation related symptoms of depression on cognitive functioning. Illustrations of hot cognition are the superior ability of MDD patients versus unaffected controls to remember negative information (Gaddy and Ingram, 2014), dysfunctional attitudes, and negative schema's that are typical during MDD. Recently it has been proposed to integrate aspects of hot cognition – including emotion and motivation related aspects – and cold cognition into a conceptual framework of MDD (Ahern et al., 2019). This framework considers cold and hot cognition as complementary processes contributing to depressive symptomatology, reinforcing and sustaining each other. Taking the issue of effort into account in the assessment of cognitive functioning in MDD, does justice to the concept of interplay between cold and hot cognition: aspects of hot cognition may affect effort and thus lead to invalid performance on neuropsychological tests aiming to assess cold cognition. In this context it should be taken into account that invalid neuropsychological test performance might be an objectification of a problem in itself, contributing to the burden of disease in MDD: if a proportion of MDD patients is not able to put forth adequate effort during neuropsychological assessment to measure cognitive functions validly, they might not be able to put forth adequate effort in daily life to use their cognitive functions effectively. It is, however, a rather crucial difference for the target of treatment whether there is actual cognitive impairment (cold), or (only) hindrance to use cognitive functions effectively due to emotion related factors (hot).

4.2. Limitations

Major limitations of this review are the small number of studies that could be included and the small sample sizes of these studies. Further, as the included studies had different aims, they did not always report all data relevant to this review nor were these data always available upon request: we received data from authors of only one of the four studies where data were requested from (Beblo et al., 2020).

Lack of data reported was especially an issue with regard to the second aim of this review: to assess cognitive functioning in MDD patients with valid test performance, and if and how this differs from patients with invalid test performance. Even when neuropsychological assessment took place, a number of studies did not report assessment outcomes for the valid and/or invalid MDD sample. Only one study reported separate outcomes of analyses comparing MDD patients and unaffected controls, including all participants or only participants who met performance validity criteria (Benitez et al., 2011). None of the included studies compared cognitive functioning in MDD patients with valid versus invalid test performance.

Due to these limitations it is not possible to draw conclusions about the total MDD population based on this review. Thus, the prevalence of valid test performance in the total MDD population remains unknown as well as the prevalence of cognitive impairment.

4.3. Conclusions and recommendations for future research

About 94 % of MDD patients included in studies using PVT's had valid performance and a noteworthy proportion of 6 % did not. Further,

although very limited evidence was available, there are indications that taking performance validity into account can significantly affect study outcomes and alter conclusions regarding cognitive impairment in MDD. Based on the current evidence, *not* taking neuropsychological performance validity into account when studying cognition in MDD, leads to major uncertainty about outcomes in clinical as well as research context. Researching the issue of suboptimal effort is essential to come to a comprehensive understanding of cognitive impairment in MDD patients and determine the relevant targets for treatment. We urge future research on this topic, to add – preferably multiple – performance validity tests into their research design, and analyze valid neuropsychological test results separately. Specific MDD state-related aspects of performance validity can be explored in a design with multiple assessments before, during, and after a depressive episode.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

AML, MASH, HB, CLB, DD and GJG designed the study. AML searched the databases. AML and GJG screened records for inclusion. Full-text screening was done by AML and cross-checked by GJG. AML corresponded with the authors in case of missing data/information. AML extracted data. MASH cross-checked data extraction. AML and HB meta-analyzed the data. AML, MASH, HB and GJG drafted the manuscript. All authors reviewed and revised the manuscript. The final manuscript was approved for submission by all authors.

Conflict of interest

We have no conflicts of interest to declare.

Acknowledgments

We thank the authors who replied to out e-mail requests for additional data and/or information.

Appendix. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jad.2022.12.043.

References

Ahern, E., Semkovska, M., 2017. Cognitive functioning in the first-episode of major depressive disorder: a systematic review and meta-analysis. Neuropsychology 31, 52–72. https://doi.org/10.1037/neu0000319.

Ahern, E., Bockting, C.L.H., Semkovska, M., 2019. A hot-cold cognitive model of depression: integrating the neuropsychological approach into the cognitive theory framework. Clin. Psychol. Eur. 1, 1–35. https://doi.org/10.32872/cpe.v1i3.34396.

Beblo, T., Driessen, M., Dehn, L., 2020. Memory deficits in patients with major depression: yes, they are trying hard enough! Expert. Rev. Neurother. 20, 517–522. https://doi.org/10.1080/14737175.2020.1754799.

Bech, P., Kastrup, M., Rafaelsen, O.J., 1986. Mini-compendium of rating scales for states of anxiety, depression, mania, schizophrenia with corresponding DSM-III syndromes. Acta Psychiat. Scand. Suppl. 326, 1–3.

Beck, A.T., Steer, R.A., Ball, R., Ranieri, W., 1996. Comparison of Beck depression inventories-IA and -II in psychiatric outpatients. J. Pers. Assess. 67, 588–597. https://doi.org/10.1207/s15327752jpa6703_13.

Benitez, A., Horner, M.D., Bachman, D., 2011. Intact cognition in depressed elderly veterans providing adequate effort. Arch. Clin. Neuropsychol. 26, 184–193. https://doi.org/10.1093/arclin/acr001.

Bolan, B., Foster, J.K., Schmand, B., Bolan, S., 2002. A comparison of three tests to detect feigned amnesia: the effects of feedback and the measurement of response latency. J. Clin. Exp. Neuropsychol. 24, 154–167. https://doi.org/10.1076/jcen.24.2.154.1000.

Brockhaus, R., Merten, T., 2004. Neuropsychologische diagnostic suboptimalen leistungsverhaltens mit dem word memory test. Nervenarzt 75, 882–887. https://doi.org/10.1007/s00115-003-1592-4.

Bush, S.S., Ruff, R.M., Troster, A.I., Barth, J.T., Koffler, S.P., Pliskin, N.H., Reynolds, C.R., Silver, C.H., 2005. Symptom validity assessment: Practice issues and medical necessity NAN policy & planning committee. Arch. Clin. Neuropsychol. 20, 419–426. https://doi.org/10.1016/j.acn.2005.02.002.

Dandachi-FitzGerald, B., Merckelbach, H., Ponds, R.W., 2017. Neuropsychologists' ability to predict distorted symptom presentation. J. Clin. Exp. Neuropsychol. 39, 257–264. https://doi.org/10.1080/13803395.2016.1223278.

DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. Control. Clin. Trials 7, 177–188. https://doi.org/10.1016/0197-2456(86)90046-2.

Egeland, J., Sundet, K., Rund, B.R., Asbjørnsen, A., Hugdahl, K., Landrø, N.I., Lund, A., Roness, A., Stordal, K.I., 2003. Sensitivity and specificity of memory dysfunction in schizophrenia: a comparison with major depression. J. Clin. Exp. Neuropsychol. 25, 79–93. https://doi.org/10.1076/jcen.25.1.79.13630.

Freeman, M.F., Tukey, J.W., 1950. Transformations related to the angular and the square root. Ann. Math. Stats. 21, 607–611. https://doi.org/10.1214/aoms/1177729756.

Gaddy, M.A., Ingram, R.E., 2014. A meta-analytic review of mood-congruent implicit memory in depressed mood. Clin. Psychol. Rev. 34, 402–416. https://doi.org/10.1016/j.cpr.2014.06.001.

Gass, C.S., Patten, B., 2020. Depressive symptoms, memory complaints, and memory test performance. J. Clin. Exp. Neuropsychol. 42, 602–610. https://doi.org/10.1080/13803395.2020.1782848.

Gervais, R.O., Rohling, M.L., Green, O., Ford, W., 2004. A comparison of WMT, CARB, and TOMM failure rates in non-head injury disability claimants. Arch. Clin. Neuropsychol. 19, 475–487. https://doi.org/10.1016/j.acn.2003.05.001.

Green, P., 2005. Word Memory Test for Windows: User's Manual and Program. Green's Publishing, Edmonton.

Green, P., Lees-Haley, P.R., Allen, L.M., 2002. The Word MMemory Test and the validity of neuropsychological test scores. J. Forensic Neuropsychol. 2, 97–124. https://doi.org/10.1300/J151v02n03_05.

Greher, M.R., Wodushek, T.R., 2017. Performance validity testing in neuropsychology: scientific basis and clinical application-a brief review. J. Psychiatr. Pract. 23, 134–140. https://doi.org/10.1097/PRA.0000000000000218.

Greiffenstein, M.F., Greve, K.W., Bianchini, K.J., Baker, W.J., 2008. Test of memory malingering and word memory test: a new comparison of failure concordance rates. Arch. Clin. Neuropsychol. 23, 801–807. https://doi.org/10.1016/j.acn.2008.07.005.

Greve, K.W., Bianchini, K.J., Doane, G.M., 2006. Classification accuracy of the Test of Memory Malingering in Traumatic Brain Injury: results of a known-group analysis. J. Clin. Exp. Neuropsychol. 28, 1176–1190. https://doi.org/10.1080/13803390500263550.

Hamilton, M., 1960. A rating scale for depression. J. Neurol. Neurosurg. Psychiatry 23, 56–62. https://doi.org/10.1136/jnnp.23.1.56.

Heilbronner, R.L., Sweet, J.J., Morgan, J.E., Larrabee, G.J., Millis, S.R., Conference, P., 2009. American Academy of Clinical Neuropsychology Consensus Conference Statement on the neuropsychological assessment of effort, response bias, and malingering. Clin. Neuropsychol. 23, 1093–1129. https://doi.org/10.1080/13854040903155063.

Jamieson, A., Goodwill, A.M., Termine, M., Campbell, S., Szoeke, C., 2019. Depression related cerebral pathology and its relationship with cognitive functioning: A systematic review. J. Affect. Disord. 250, 410–418. https://doi.org/10.1016/j.jad.2019.03.042.

Joormann, J., Gotlib, I.H., 2008. Updating the contents of working memory in depression: interference from irrelevant negative material. J. Abnorm. Psychol. 117, 182–192. https://doi.org/10.1037/0021-843X.117.1.182.

Lee, A., Boone, K.B., Lesser, I., Wohl, M., Wilkins, S., Parks, C., 2000. Performance of older depressed patients on two cognitive malingering tests: false positive rates for the Rey 15-item memorization and dot counting tests. Clin. Neuropsychol. 14, 303–308. https://doi.org/10.1076/1385-4046(200008)14:3;1-P;FT303.

Legemaat, A.M., Semkovska, M., Brouwer, M., Geurtsen, G.J., Burger, H., Denys, D., Bockting, C.L., 2021. Effectiveness of cognitive remediation in depression: a meta-analysis. Psychol. Med. 14, 1–16. https://doi.org/10.1017/S0033291721001100.

Lippa, S.M., 2018. Performance validity testing in neuropsychology: a clinical guide, critical review, and update on a rapidly evolving literature. Clin. Neuropsychol. 32, 391–421. https://doi.org/10.1080/13854046.2017.1406146.

Listunova, L., Roth, C., Bartolovic, M., Kienzle, J., Bach, C., Weisbrod, M., Roesch-Ely, D., 2018. Cognitive impairment along the course of depression: non-pharmacological treatment options. Psychopathology 51, 295–305. https://doi.org/10.1159/000492620.

Lundbeck, 2018. FDA updates Trintellix® (vortioxetine) label to include data showing improvement in processing speed, an important aspect of cognitive function in acute Major Depressive Disorder (MDD). In: https://www.globenewswire.com/news-release/2018/05/02/1495434/0/en/FDA-updates-Trintellix-vortioxetine-label-to-include-data-showing-improvement-in-processing-speed-an-important-aspect-of-cognitive-function-in-acute-Major-Depressive-Disorder-MDD.html accessed 21 December 2021.

Mahableshwarkar, A.R., Zajecka, J., Jacobson, W., Chen, Y., Keefe, R.S., 2015. A randomized, placebo-controlled, active-reference, double-blind, flexible-doze study of the efficacy of vortioxetine on cognitive function in major depressive disorder. Neuropsychopharmacology 40, 2025–2037. https://doi.org/10.1038/npp.2015.52.

McCrea, M., Pliskin, N., Barth, J., Cox, D., Fink, J., French, L., Hammeke, T., Hess, D., Hopewell, A., Orme, D., Powell, M., Ruff, R., Schrock, B., Terryberry-Spohr, L., Vanderploeg, R., Yoash-Gantz, R., 2008. Official position of the military TBI task force on the role of neuropsychology and rehabilitation psychology in the evaluation, management, and research of military veterans with traumatic brain injury. Clin. Neuropsychol. 22, 10–26. https://doi.org/10.1080/13854040701760981.

McDermott, L.M., Ebmeier, K.P., 2009. A meta-analysis of depression severity and cognitive function. J. Affect. Disord. 119, 1–8. https://doi.org/10.1016/j.jad.2009.04.022.

McIntyre, R.S., Lophaven, S., Olsen, C.K., 2014. A randomized, double-blind, placebo controlled study of vortioxetine on cognitive function in depressed adults. Int. J. Neuropsychopharmacol. 17, 1557–1567. https://doi.org/10.1017/S1461145714000546.

Merten, T., Henry, M., Hilsabeck, R., 2004. Symptomvalidierungstests in der neuropsychologischen diagnostik: Eine analogstudie. Z. Neuropsychol. 15, 81–90. https://doi.org/10.1024/1016-264X.15.2.81.

Merten, T., Bossink, L., Schmand, B., 2007. On the limits of effort testing: symptom validity tests and severity on cognitive symptoms in nonlitigant patients. J. Clin. Exp. Neuropsychol. 29, 308–318. https://doi.org/10.1080/13803390600693607.

Miskowiak, K.W., Ott, C.V., Petersen, J.Z., Kessing, L.V., 2016. Systematic review of randomized controlled trials of candidate treatments for cognitive impairment in depression and methodological challenges in the field. Eur. Neuropsychopharmacol. 26, 1845–1867. https://doi.org/10.1016/j.euroneuro.2016.09.641.

Mittmann, N., Mitter, S., Borden, E.K., Herrmann, N., Naranjo, C.A., Shear, N.H., 1997. Montgomery-Åsberg severity gradations. Am. J. Psychiatry 154, 1320–1321. https://doi.org/10.1176/ajp.154.9.1320b.

Montgomery, S.A., Åsberg, M., 1979. A new depression scale designed to be sensitive to change. Br. J. Psychiatry 134, 382–389. https://doi.org/10.1192/bjp.134.4.382.

Moritz, S., Stöckert, K., Hauschildt, M., Lill, H., Jelinek, L., Beblo, T., Diedrich, S., Arlt, S., 2017. Are we exaggerating neuropsychological impairment in depression? Reopening a closed chapter. Expert. Rev. Neurother. 17, 839–846. https://doi.org/10.1080/14737175.2017.1347040.

Motter, J.N., Pimontel, M.A., Rindskopf, D., Devanand, D.P., Doraiswamy, P.M., Sneed, J.R., 2016. Computerized cognitive training and functional recovery in major depressive disorder: a meta-analysis. J. Affect. Disord. 1, 184–191. https://doi.org/10.1016/j.jad.2015.09.022.

Porter, R.J., Bourke, C., Gallagher, P., 2007. Neuropsychological impairment in major depression: Its nature, origin and clinical significance. Aust. N. Z. J. Psychiatry. 41, 115–128. https://doi.org/10.1080/00048670601109881.

Rees, L.M., 1996. The Test of Memory Malingering (TOMM): Simulation Studies and Clinical Validation. Unpublished Doctoral Dissertation. Ottawa, ON: Carleton University.

Rees, L.M., Tombaugh, T.N., Gansler, D.A., Moczynski, N.P., 1998. Five validation experiments of the test of memory malingering (TOMM). Psychol. Assess. 10, 10–20. https://doi.org/10.1037/1040-3590.10.1.10.

Rees, L.M., Tombaugh, T.N., Boulay, L., 2001. Depression and the test of memory malingering. Arch. Clin. Neuropsychol. 16, 501–506. https://doi.org/10.1016/S0887-6177(00)00064-0.

Rienstra, A., Groot, P.F., Spaan, P.E., Majoie, C.B., Nederveen, A.J., Walstra, G.J., de Jonghe, J.F., van Gool, W.A., Olabarriaga, S.D., Korkhov, V.V., Schmand, B., 2013. Symptom validity testing in memory clinics: hippocampal-memory associations and relevance for diagnosing mild cognitive impairment. J. Clin. Exp. Neuropsychol. 35, 59–70. https://doi.org/10.1080/13803395.2012.751361.

Rock, P.L., Roiser, J.P., Riedel, W.J., Blackwell, A.D., 2014. Cognitive impairment in depression: a systematic review and meta-analysis. Psychol. Med. 44, 2029–2040. https://doi.org/10.1017/S0033291713002535.

Roiser, J.P., Sahakian, B.J., 2013. Hot and cold cognition in depression. CNS Spectr. 18, 139–149. https://doi.org/10.1017/S1092852913000072.

Rowland, J.A., Miskey, H.M., Brearly, T.W., Martindale, S.L., Shura, R.D., 2017. Word memory test performance across cognitive domains, psychiatric presentations, and mild traumatic brain injury. Arch. Clin. Neuropsychol. 32, 306–315. https://doi.org/10.1093/arclin/acw107.

Ruiz, I., Raugh, I.M., Bartolomeo, L.A., Strauss, G.P., 2020. A meta-analysis of neuropsychological effort test performance in psychotic disorders. Neuropsychol. Rev. 30, 407–424. https://doi.org/10.1007/s11065-020-09448-2.

Scheurich, A., Fellgiebel, A., Schermuly, I., Bauer, S., Wölfges, R., Müller, M.J., 2008. Experimental evidence for a motivational origin of cognitive impairment in major depression. Psychol. Med. 38, 237–246. https://doi.org/10.1017/S0033291707002206.

Schmand, B., De Sterke, S., Lindeboom, J., 1999. AKTG/ Amsterdamse Korte Termijn geheugen Test. Pearson Assessment and Information B.V, Amsterdam.

Shura, R.D., Denning, J.H., Miskey, H.M., Rowland, J.A., 2017. Symptom and performance validity with veterans assessed for attention-deficit/hyperactivity disorder (ADHD). Psychol. Assess. 29, 1458–1465. https://doi.org/10.1037/pas0000436.

Slick, D.J., Sherman, E.M., Iverson, G.L., 1999. Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. Clin. Neuropsychol. 13, 545–561. https://doi.org/10.1076/1385-4046(199911)13:04;1-Y;FT545.

Snyder, H.R., 2013. Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. Psychol. Bull. 139, 81–132. https://doi.org/10.1037/a0028727.

Strauss, G.P., Morra, L.F., Sullivan, S.K., Gold, J.M., 2015. The role of low cognitive effort and negative symptoms in neuropsychological impairment in schizophrenia. Neuropsychology 29, 282–291. https://doi.org/10.1037/neu0000113.

Tan, J.E., Slick, D.J., Strauss, E., Hultsch, D.F., 2002. How'd they do it? Malingering strategies on symptom validity tests. Clin. Neuropsychol. 4, 495–505. https://doi.org/10.1076/clin.16.4.495.13909.

Thérond, A., Pezzoli, P., Abbas, M., Howard, A., Bowie, C.R., Guimond, S., 2021. The efficacy of cognitive remediation in depression: a systematic literature review and meta-analysis. J. Affect. Disord. 284, 238–246. https://doi.org/10.1016/j.jad.2021.02.009.

Tombaugh, T.N., 1996. Test of Memory Malingering TOMM (Manual). Multi-Health System, New York.

Tombaugh, T.N., 1997. The test of memory-malingering (TOMM: Normative data from cognitively intact and cognitively impaired individuals). Psychol. Assess. 9, 260–268. https://doi.org/10.1037/1040-3590.9.3.260.

van Hout, M.S.E., Schmand, B., Wekking, E.M., Hageman, G., Deelman, B.G., 2003. Suboptimal performance on neuropsychological tests in patients with suspected chronic toxic encephalopathy. Neurotoxicology. 24, 547–551. https://doi.org/10.1016/S0161-813X(03)00054-8.

Whitmer, A.J., Gotlib, I.H., 2012. Switching and backward inhibition in major depressive disorder: the role of rumination. J. Abnorm. Psychol. 121, 570–578. https://doi.org/10.1037/a0027474.

Wodushek, T.R., Greher, M.R., 2017. Performance validity testing in neuropsychology: methods for measurement development and maximizing diagnostic accuracy. J. Psychiatr. Pract. 23, 214–220. https://doi.org/10.1097/PRA.0000000000000233.

Yanez, Y.T., Fremouw, W., Tennant, J., Strunk, J., Coker, K., 2006. Effects of severe depression on TOMM performance among disability-seeking outpatients. Arch. Clin. Neuropsychol. 21, 161–165. https://doi.org/10.1016/j.acn.2005.07.009.