Psychological Medicine

cambridge.org/psm

Effectiveness of cognitive remediation in depression: a meta-analysis

Amanda M. Legemaat[1], Maria Semkovska[2], Marlies Brouwer[1], Gert J. Geurtsen[3], Huibert Burger[4], Damiaan Denys[1] and Claudi L. Bockting[1,5]

[1]Department of Psychiatry, Amsterdam University Medical Centers, Academic Medical Center, University of Amsterdam, Amsterdam Neuroscience & Amsterdam Public Health, Meibergdreef 9 1105 AZ, Amsterdam, The Netherlands; [2]Department of Psychology, University of Southern Denmark, Campusvej 55 DK-5230 Odense M, Denmark; [3]Department of Medical Psychology, Amsterdam University Medical Centers, Academic Medical Center, University of Amsterdam, Amsterdam Neuroscience & Amsterdam Public Health, Meibergdreef 9 1105 AZ, Amsterdam, The Netherlands; [4]Department of General Practice and Elderly Care Medicine, University Medical Center Groningen, University of Groningen, Antonius Deusinglaan 1 9713 AV, Groningen, The Netherlands and [5]Centre for Urban Mental Health, University of Amsterdam, Oude Turfmarkt 147 1012 GC, Amsterdam, The Netherlands

Original Article

Cite this article: Legemaat AM, Semkovska M, Brouwer M, Geurtsen GJ, Burger H, Denys D, Bockting CL (2022). Effectiveness of cognitive remediation in depression: a meta-analysis. *Psychological Medicine* **52**, 4146–4161. https://doi.org/10.1017/S0033291721001100

Received: 8 July 2020
Revised: 16 February 2021
Accepted: 5 March 2021
First published online: 14 April 2021

Key words:
Cognitive remediation; cognitive training; depression; major depressive disorder; depressive symptomatology; recurrence; relapse; cognitive functioning; cognitive impairment; daily functioning

Author for correspondence:
Claudi L. Bockting,
E-mail: c.l.bockting@amsterdamumc.nl

Abstract

Background. Preliminary evidence suggests beneficial effects of cognitive remediation in depression. An update of the current evidence is needed. The aim was to systematically assess the effectiveness of cognitive remediation in depression on three outcomes.

Methods. The meta-analysis was pre-registered on PROSPERO (CRD42019124316). PubMed, PsycINFO, Embase and Cochrane Library were searched on 2 February 2019 and 8 November 2020 for peer-reviewed published articles. We included randomized and non-randomized clinical trials comparing cognitive remediation to control conditions in adults with primary depression. Random-effects models were used to calculate Hedges' *g*, and moderators were assessed using mixed-effects subgroup analyses and meta-regression. Main outcome categories were post-treatment depressive symptomatology (DS), cognitive functioning (CF) and daily functioning (DF).

Results. We identified 5221 records and included 21 studies reporting on 24 comparisons, with 438 depressed patients receiving cognitive remediation and 540 patients in a control condition. We found a small effect on DS ($g = 0.28$, 95% CI 0.09–0.46, I^2 40%), a medium effect on CF ($g = 0.60$, 95% CI 0.37–0.83, I^2 44%) and a small effect on DF ($g = 0.22$, 95% CI 0.06–0.39, I^2 3%). There were no significant effects at follow-up. Confounding bias analyses indicated possible overestimation of the DS and DF effects in the original studies.

Conclusions. Cognitive remediation in depression improves CF in the short term. The effects on DS and DF may have been overestimated. Baseline depressive symptom severity should be considered when administering cognitive remediation.

Introduction

Major depressive disorder (MDD) is the most common mental health disorder (Moffitt et al., 2010). It is associated with reduced daily functioning (DF) (Adler et al., 2006; de Jonge et al., 2018; Moffitt et al., 2010; ten Doesschate, Bockting, Koeter, & Schene, 2010) and impaired cognitive functioning (CF) (Ahern & Semkovska, 2017; Keyes, Platt, Kaufman, & McLaughlin, 2017; Rock, Roiser, Riedel, & Blackwell, 2014; Semkovska et al., 2019). Notably, impaired CF is not limited to the acute phase of MDD but persists when MDD has remitted, while the level of CF impairment appears to worsen with repeated episodes (Semkovska et al., 2019). Further, impaired CF associated with MDD has been found to predict the level of DF, independently of mood symptoms (Jaeger, Berns, Uzelac, & Davis-Conway, 2006; McIntyre et al., 2013). Moreover, impaired CF is believed to be an important factor in the maintenance of a vicious cycle of depressive symptomatology (DS), reduced DF and MDD recurrence (Ahern, Bockting, & Semkovska, 2019; Jaeger et al., 2006; Majer et al., 2004). Thus, addressing CF might improve outcomes (Ahern et al., 2019). A promising method in the treatment of MDD and elevated depressive symptoms, which indeed addresses CF, is cognitive remediation (Cella et al., 2020; Motter et al., 2016). This involves drill-and-practice exercises and/or cognitive strategy training. Cognitive remediation aims to improve CF by means of enhancing neuroplasticity (Robertson & Murre, 1999), or to compensate for impaired CF in daily life (Twamley, Vella, Burton, Heaton, & Jeste, 2012). Therapy delivery format is variable, and includes computerized (e.g. online training) and non-computerized (e.g. offline work with a therapist), and individual and group formats.

CAMBRIDGE UNIVERSITY PRESS

A first meta-analysis on the effectiveness of (computerized) cognitive remediation in MDD ($N = 9$; $n = 539$) suggests that it improves DS and CF as well as DF (Motter et al., 2016). However, as the authors acknowledge, the small number of studies and patients included limited their meta-analysis. Given the increasing number of studies (Semkovska, Lambe, Lonargáin, & McLoughlin, 2015; Trapp, Engel, Hajak, Lautenbacher, & Gallhofer, 2016), an updated meta-analysis is warranted. Further, the previous meta-analysis (Motter et al., 2016) could not examine the effect of therapy delivery format, clinical patient characteristics or effects at follow-up. In addition, they did not perform a sensitivity analysis excluding non-randomized studies, or any assessment of risk of bias or certainty, besides risk of publication bias.

The current meta-analysis therefore aimed to update and expand on the previous meta-analysis (Motter et al., 2016) in order to evaluate the effectiveness of cognitive remediation in depression. We primarily aimed to investigate the effects on DS, CF and DF (e.g. work, social functioning, quality of life). Secondarily, the aim was to conduct subgroup- and moderator analyses to assess the influence of therapy delivery format (computerized *v.* non-computerized, group *v.* individual); patients' clinical status (current *v.* remitted depression); control group [placebo *v.* waitlist/treatment as usual (TAU) control group]; baseline depressive symptom severity; and diagnosis (clinical MDD *v.* no formal clinical diagnosis, i.e. depression based on elevated depressive symptoms). We assessed effects at follow-up as well.

Methods

Search strategy and selection criteria

We conducted a meta-analysis in accordance with PRISMA guidelines and its protocol was prospectively registered on PROSPERO (CRD42019124316). Databases PubMed, PsycINFO, Embase and Cochrane Library were searched for relevant studies published from their origin through 2 February 2019. The search was updated on 8 November 2020. The search strategy included key words and MeSH terms related to cognitive remediation and depression (Appendix II). References and citations of included studies, relevant reviews and meta-analyses were searched for additional studies.

To be included, studies needed to be a randomized or non-randomized clinical trial, testing the effectiveness of cognitive remediation, as compared to a non-cognitive remediation control group (e.g. placebo, waitlist, TAU), in current or remitted patients with primary depression, aged ⩾18 years and reporting sufficient statistics to calculate effect sizes. For example, antidepressant medication and cognitive behavioural therapy were considered TAU. Depression was operationalized as an MDD diagnosis confirmed by a clinician, clinical interview or elevated symptoms/disorder based on any instrument aimed at assessing MDD. We utilized tolerant diagnostic criteria in order to remain inclusive and to include studies with a relatively broad range of baseline depressive symptom severity. The rationale for this was to promote the generalizability of the results, and to enable exploring the effect of baseline depressive symptom severity. Statistics were considered sufficient if post-cognitive remediation summary means (M) and standard deviations (S.D.) on either DS, CF or DF were reported. In case of mixed samples (e.g. schizophrenia, bipolar disorder, MDD), we required statistics for the depression subsample. There were no limitations with regard to publication year; we aimed to include all relevant peer-reviewed studies published to this date. Exclusion criteria were coexisting psychotic disorders, brain injuries, other neurological disorders, recent/consecutive electroconvulsive therapy and any form of transcranial stimulation as this might affect cognitive remediation results (Jahshan, Rassovsky, & Green, 2017). Papers written in English, French and Dutch language were included.

After removing duplicates, two authors (AML and MB) independently screened titles and abstracts and selected studies with potential for inclusion. Selected studies were reviewed independently full-text (AML, MS and MB). Any disagreements were resolved through consensus (AML, MS and MB).

Data-analysis

Extracted data for the cognitive remediation and control conditions were: number, gender and age of patients; diagnostic instruments and criteria for depression; current/remitted depression; intervention characteristics; instruments to assess DS, CF and DF; M and S.D. of DS, CF and DF measures post-intervention and at follow-up; DS measures at baseline; time from end of treatment to post-intervention and follow-up assessments; data on quality, including randomization. In case of multiple comparisons within the same study, all relevant comparisons were included in the meta-analysis. In order to justify the weight of the respective comparisons by the true number of participants, participants (n) included in both comparisons were equally divided across the comparisons (i.e. two cognitive remediation samples were each compared to half of the same control sample, and *vice versa* when two relevant control samples were included, they were each compared to half of the same cognitive remediation sample) (Higgins et al., 2020). In case of data overlap, only the most recent study was included to ensure statistical independence.

Outcome measures were divided into three main outcome categories: DS, CF and DF. Measures of cognitive domains by means of objective standardized cognitive tests were considered CF outcomes. Measures of aspects of (satisfaction with) functioning in daily live, e.g. quality of life, administration tasks and social interactions, were categorized as DF outcomes. CF outcomes were further divided into standardized cognitive domains, namely Attention; Processing speed; Motor speed; Working memory; Verbal learning and memory; Visual learning and memory; Executive functioning; Verbal fluency; Global/intellectual functioning (Lezak, Howieson, Bigler, & Tranel, 2012). DF outcomes were divided into subjective and objective. Subjective DF was operationalized as self-reported DF, e.g. a questionnaire on quality of life filled in by a patient. Objective DF was operationalized as clinician-rated DF, e.g. results on an advanced finances task rated by a clinician.

Categories were defined by authors AML and MS. Data were extracted and categorized by AML. Data extractions and categorizations were cross-checked by MS and MB. If any relevant information was found to be missing, the corresponding authors of the respective articles were contacted to request the information and reminded twice.

AML rated the risk of bias and MB cross-checked the ratings using the Cochrane Risk of Bias tool, as recommended by the GRADE system (Guyatt et al., 2011). For each study, seven criteria were scored as low risk of bias (0 points), unclear risk of bias (1 point) or high risk of bias (2 points). A study was rated to have low risk of bias (total points <6) or high risk of bias (total points >6). We assessed the overall certainty of evidence for the three main outcome categories using the GRADE framework.

https://doi.org/10.1017/S0033291721001100 Published online by Cambridge University Press

We used Comprehensive Meta-Analysis Software (version 3) (Borenstein, Hedges, Higgins, & Rothstein, 2013) to calculate effect sizes (Hedges' *g*) based on means and standard deviations, and number of patients in both conditions at the first post-intervention assessment (cognitive remediation compared to control). For follow-up analyses, we used the first follow-up time-point (i.e. any additional assessment after the first post-intervention assessment) as a starting point to assess effects at follow-up. We similarly calculated effect sizes based on means, standard deviations and number of patients in both conditions. For the analyses on DS, CF and DF, the mean of the effect sizes per study on DS, CF or DF outcomes, respectively, was used. For the analyses on CF domains and DF sub-categories, the mean of the effect sizes per study per domain/sub-category was used. For each outcome, a positive effect size indicated greater improvement in the cognitive remediation condition compared to the control condition. Effect sizes were weighted by their inverse variance in order to give more weight to studies with larger sample sizes. To determine statistical significance, two-sided 95% confidence intervals were used. Weighted, mean effect sizes of 0.2–0.49 were considered small; 0.5–0.79 medium; and >0.8 large (Cohen, 1988). The I^2 index was used to quantify heterogeneity. Percentages of <40% were considered small; 30–60% moderate; 50–90% substantial; and 75–100% considerable heterogeneity (Higgins et al., 2020). We used a random-effects model, and mixed (random within and fixed across subgroups) effects model for categorical subgroup analyses, because of the *a priori* assumption that there would be substantial variability between the included studies (Borenstein, Hedges, Higgins, & Rothstein, 2009).

To assess the moderating effect of baseline depressive symptom severity, Montgomery-Asberg Depression Rating Scale (Montgomery & Åsberg, 1979) or Beck Depression Inventory-II (Beck, Steer, & Brown, 1996) scores were transformed to Hamilton Depression Rating Scale (HDRS-17) (Hamilton, 1967) scores (Heo, Murphy, & Meyers, 2007; Vittengl, Clark, Kraft, & Jarrett, 2005). Mean HDRS-17 <8 was considered minimal; HDRS-17 8–15 moderate; and HDRS-17 >15 severe symptoms. Subgroup analyses were performed by clustering studies into contrasting subgroups. To ensure adequate power, a minimum of three studies per subgroup was required. Continuous moderators were analysed by simple meta-regression. Meta-regression analyses were not performed if the number of studies was <10.

Publication bias for the three main outcome categories was assessed by inspecting funnel plots and using Egger's test for their symmetry, and Duval and Tweedie's trim and fill procedure. Sensitivity analyses for the effect on DS, CF and DF were performed excluding outliers defined as individual studies showing an effect size with a 95% confidence interval that did not show any overlap with the 95% confidence interval of the overall, i.e. pooled, effect (Harrer, Cuijpers, Furukawa, & Ebert, 2019); studies with high risk of bias; insufficient sequence generation; small number of patients (*n* in either one of the conditions <5); large number of days from end of treatment to post-intervention assessment (>14 days); and studies with participants without a formal clinical MDD diagnosis.

Results

Study characteristics

We identified 5221 records, and included 21 studies with 438 patients allocated to a cognitive remediation condition and 540 patients allocated to a control condition (Alvarez, Cortés Sotres, León, Estrella, & Sánchez Sosa, 2008; Anguera, Gunning, & Areán, 2017; Bowie et al., 2013; Elgamal, McKinnon, Ramakrishnan, Joffe, & MacQueen, 2007; Hoorelbeke & Koster, 2017; Hoorelbeke, van den Bergh, de Raedt, Wichers, & Koster, 2021; Listunova et al., 2020; Morimoto et al., 2014, 2020; Moshier, Molokotos, Stein, & Otto, 2015; Moshier & Otto, 2017; Naismith et al., 2011; Owens, Koster, & Derakshan, 2013; Pratap et al., 2018; Semkovska & Ahern, 2017; Semkovska et al., 2015; Trapp et al., 2016; Twamley et al., 2019; Wanmaker, Geraerts, & Franken, 2015; Wanmaker, Hopstaken, Asselbergs, Geraerts, & Franken, 2014; Yamaguchi et al., 2017) (Fig. 1).

The total number of comparisons was 24.[†1] Three studies had high risk of bias (Elgamal et al., 2007; Morimoto et al., 2014; Owens et al., 2013) (Appendix I – eTable 1). Results on relevant outcomes were categorized into DS, CF and DF categories and sub-categories (Appendix I – eTable 2). Twenty-one comparisons included DS, 19 included CF and 12 included DF outcomes (see Table 1 for further study details).

Main effects on depressive symptomatology, cognitive and daily functioning

The direction of the effect was favourable and significant for all three outcome categories. There was a small significant effect on DS ($g = 0.28$; 95% CI 0.09–0.46), a medium significant effect on CF ($g = 0.60$; 95% CI 0.37–0.83) and a small significant effect on DF ($g = 0.22$; 95% CI 0.06–0.39). Heterogeneity was moderate for both DS ($I^2 = 40\%$) and CF ($I^2 = 44\%$), and small for DF ($I^2 = 3\%$) (see Table 2 and Fig. 2).

Subgroup analyses

With regard to therapy delivery format, only one study had a full non-computerized format, and only one other study had a full group format. As we required a minimum of three studies per subgroup, we did not perform subgroup analyses based on therapy format. Only two studies included patients with minimal depressive symptom severity at baseline; thus, only subgroups of moderate and severe baseline depressive symptom severity were analysed. There were not enough studies to perform subgroup analyses based on diagnosis for the effects on CF and DF.

Depressive symptomatology

Subgroup analyses showed that there was a significantly larger effect on DS in patients with severe baseline depressive symptoms compared to patients with moderate baseline depressive symptoms: there was no significant effect on DS in patients with moderate baseline symptoms, while in those with severe baseline symptoms, there was a small significant effect ($g = 0.48$). With regard to the effect on DS, difference in effect size between other subgroups did not reach statistical significance (Table 2).

Cognitive functioning

Significant effects for CF domains were: small for Attention ($g = 0.36$), Processing speed ($g = 0.26$) and Verbal learning and memory ($g = 0.47$); and medium for Working memory ($g = 0.54$) (Table 2). There were insufficient studies reporting outcomes on Motor speed ($N = 1$) and Global/intellectual functioning ($N = 2$) to meta-analyse

[†]The notes appear after the main text.

https://doi.org/10.1017/S0033291721001100 Published online by Cambridge University Press

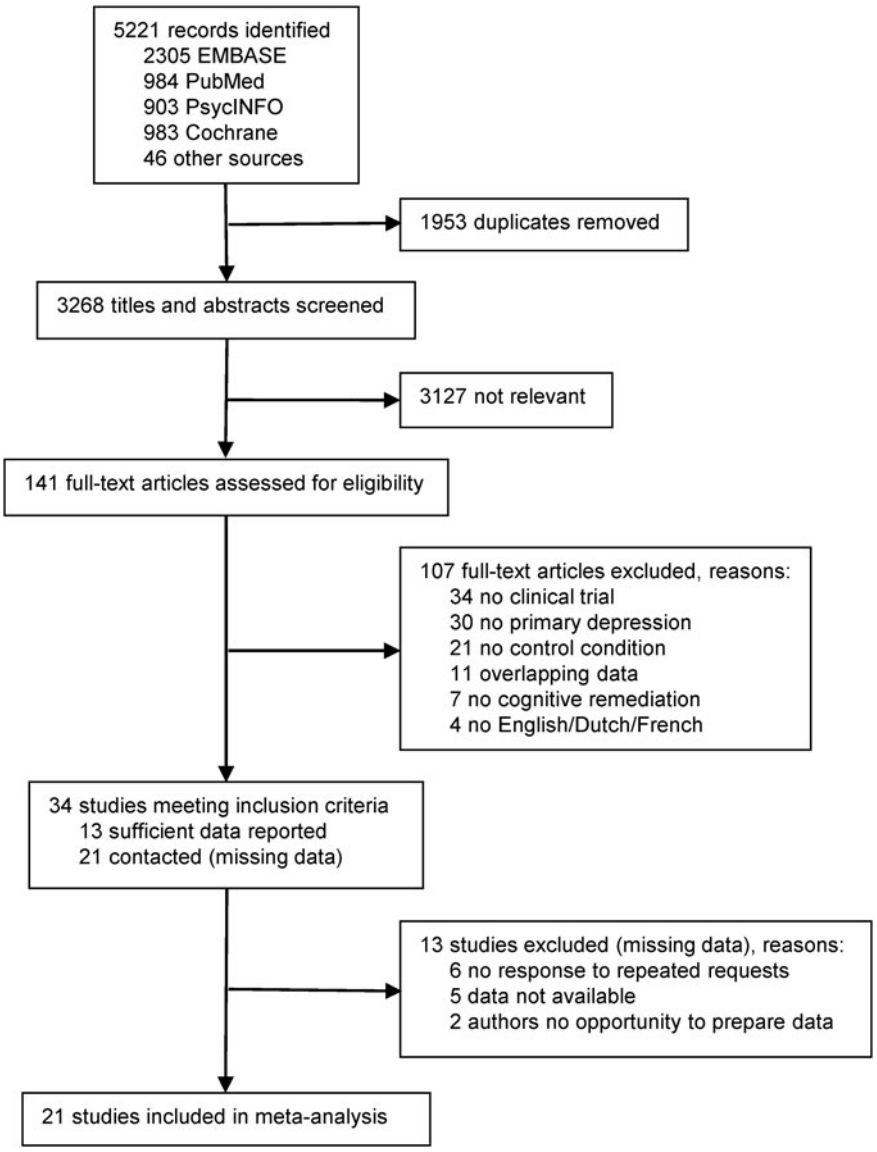

Fig. 1. PRISMA flow diagram.

these effects. There were no significant effects on Visual learning and memory; Executive functioning; or Verbal fluency. Subgroup analyses revealed that the effect on CF was significantly larger in comparison to placebo control groups (large significant effect, $g = 0.84$), than in comparison to Waitlist/TAU control groups (small significant effect, $g = 0.39$). There were no significant differences in effect size for CF between other subgroups.

Daily functioning

There was a small significant effect on subjective DF ($g = 0.22$) and no significant effect on objective DF (Table 2). Subgroup analyses revealed no significant differences in effect sizes for DF between subgroups.

Meta-regression analyses

We performed simple meta-regression analyses on the moderating effects of baseline depressive symptom severity (mean HDRS-17 scores) and post-hoc on age (mean age), gender (percentage female) and cognitive remediation duration (in minutes). There were no significant effects.[2]

Effects at follow-up

As a number of studies provided outcomes at follow-up (i.e. any additional assessment after the first post-intervention assessment, ranging from 1 to 3 months after the first post-intervention assessment), we opted to perform post-hoc analyses on the effects of cognitive remediation *v.* control on DS, CF and DF at follow-up. We took the first follow-up time-point. There were no significant effects of cognitive remediation compared to control at follow-up.[3] As we found no significant durable effects at the first follow-up time point, we did not further analyse effects at follow-up.

Publication bias and certainty of the evidence

Inspection of the funnel plots and Egger's test did not indicate publication bias for DS ($p = 0.67$), CF ($p = 0.40$) or DF ($p = 0.48$). Duval and Tweedie's trim and fill procedure under the

https://doi.org/10.1017/S0033291721001100 Published online by Cambridge University Press

Table 1. Study characteristics

			Cognitive remediation condition				Control condition				
Study	Diagnostic instruments and criteria	HDRS-17 baseline M (severity)	*n*	Mean age (s.d.)	Intervention	Duration	*n*	Mean age (s.d.)	Intervention	Duration	Outcome measures
Alvarez et al. (2008)[a]	DSM-IV; MINI; clinical MDD diagnosis	16.46 (severe)	10	21.0 (2.9)	Alcor cognitive training: series completion task and mental arithmetical operations; adaptive; computerized; individual	64 sessions during 16 weeks; total 960 min	6*	23.8 (2.7)	Waitlist/TAU: antidepressant medication	16 weeks	BDI WAIS VIQ WAIS PIQ
Alvarez et al. (2008)[b]	DSM-IV; MINI; clinical MDD diagnosis	17.09 (severe)	10	23.3 (3.7)	1. Alcor cognitive training: series completion task and mental arithmetical operations; adaptive; computerized; individual 2. Antidepressant medication	1. 64 sessions during 16 weeks; total 960 min 2. 16 weeks	6*	23.8 (2.7)	Waitlist/TAU: antidepressant medication	16 weeks	BDI WAIS VIQ WAIS PIQ
Anguera et al. (2017)	DSM-IV; SCID; PHQ-9; HDRS-17 >24; clinical MDD diagnosis	23.15 (severe)	12	66.9 (6.8)	1. Project: EVOtm cognitive training: guiding a character through an immersive environment, selectively responding to targets; adaptive; computerized; individual 2. Check in with therapist	1. 20 sessions during 4 weeks; total 400 min 2. 8 sessions during 8 weeks	10	69.4 (5.6)	Waitlist/TAU: PST: psychoeducation, practicing PST skills, relapse prevention; therapist delivered; individual	8 sessions during 8 weeks	HDRS-17 PHQ-9 TOVA Clapp's WM task
Bowie et al. (2013)	Clinical MDD diagnosis	19.06 (severe)	11	49.2 (11.8)	1. Scientific Brain Training Pro cognitive training; adaptive; computerized; individual 2. Strategic self-monitoring coaching; non-computerized; group 3. 'Bridging' discussions to facilitate transfer; non-computerized; group 4. Homework sessions; partly computerized; individual 5. Case management and pharmacotherapy services	1–4: 10 sessions + daily homework during 10 weeks; total 3700 min 5. Ongoing	10	42.2 (13.4)	Waitlist/TAU: 1. Waitlist for CR 2. Case management and pharmacotherapy services	1. 10 weeks 2. Ongoing	CPT-IP TMT-A BACS SCT Gold's LNS HVLT TMT-B Stroop CWT inhibition COWAT LIFE-RIFT SSPA Advanced finances task

Elgamal et al. (2007)	DSM-IV; SCID; remitted clinical MDD diagnosis	10.25 (moderate)	12	50.3 (6.4)	1. PSSCOgReHab cognitive training; adaptive; computerized; individual 2. Antidepressant medication	1. 20 sessions during 10 weeks; total 1200 min 2. Ongoing	12	47.4 (6.8)	Waitlist/TAU: antidepressant medication	Ongoing; assessment after 10 weeks	HDRS-17 Ruff's 2&7 SAT WAIS-R DS Forward TMT-A WAIS-R DS Backward CVLT TMT-B WAIS-R Similarities COWAT
Hoorelbeke & Koster (2017)	MINI; remitted clinical MDD diagnosis; stable remission >6 months	6.02 (minimal)	33	46.12 (10.8)	1. Psycho-education session to foster task engagement 2. Cognitive control training: modified PASAT: responding to the sum of the last two digits while hearing a continuous stream of digits; adaptive; computerized; individual	1. Once 2. 10 sessions during 2 weeks; total 143 min	34	47.8 (12.2)	Placebo: 1. Psycho-education session to foster task engagement 2. Low cognitive load training: responding to the last digit while hearing a continuous stream of digits; non-adaptive; computerized; individual	1. Once 2. 10 sessions during 2 weeks	BDI-II RRS RDQ Non-adaptive PASAT WHODAS 2.0 BRIEF-A Global Executive Scale QLDS RS
Hoorelbeke et al. (2021)	MINI; remitted clinical MDD diagnosis	8.9 (moderate)	40	45.14 (14.42)	1. Psycho-education session to foster task engagement 2. Cognitive control training: modified PASAT: responding to the sum of the last two digits while hearing a continuous stream of digits; adaptive; computerized; individual	1. Once 2. 10 sessions during 4 weeks; total 143 min	36	45.6 (11.7)	Placebo: 1. Psycho-education session to foster task engagement 2. Low cognitive load training: responding to the last digit while hearing a continuous stream of digits; non-adaptive; computerized; individual	1. Once 2. 10 sessions during 4 weeks	BDI-II RRS RDQ Non-adaptive PASAT BRIEF-A Global Executive Scale RS
Listunova et al. (2020)[a]	DSM-IV; SCID; MINI; HDRS-24 <20; (partially) remitted clinical MDD diagnosis	13.04 (moderate)	20	45.90 (11.34)	1. Cognitive remediation therapy with CogniPuls; training 6 standard cognitive domains; adaptive; computerized; individual 2. Compensatory transfer sessions; non-computerized 3. Medical and psychotherapeutic TAU	1. 15 sessions during 5 weeks; total 900 min 2. 5 sessions during 5 weeks; total 150 min 3. Ongoing	10**	44.89 (10.32)	Waitlist/TAU: medical and psychotherapeutic TAU	Ongoing; assessment after 5–7 weeks	VTS WAF-A; WAF-G; WAF-S; TMT-A; N-Back-verbal; Figural Memory Test; INHIB; TMT-B; TOL-F Zahlen-Symbol-Test CVLT MINI-ICF self; external SLOF
Listunova et al. (2020)[b]	DSM-IV; SCID; MINI; HDRS-24 <20; (partially) remitted clinical MDD diagnosis	13.06 (moderate)	18	45.33 (15.06)	1. Cognitive remediation therapy with CogniPuls; training 3 most impaired cognitive domains; adaptive; computerized;	1. 15 sessions during 5 weeks; total 900 min 2. 5 sessions	10**	44.89 (10.32)	Waitlist/TAU: medical and psychotherapeutic TAU	Ongoing; assessment after 5–7 weeks	VTS WAF-A; WAF-G; WAF-S; TMT-A; N-Back-verbal; Figural Memory Test; INHIB; TMT-B; TOL-F Zahlen-Symbol-Test CVLT

(*Continued*)

Table 1. (*Continued.*)

Study	Diagnostic instruments and criteria	HDRS-17 baseline M (severity)	Cognitive remediation condition				Control condition				Outcome measures
			n	Mean age (s.d.)	Intervention	Duration	*n*	Mean age (s.d.)	Intervention	Duration	
					individual 2. Compensatory transfer sessions; non-computerized 3. Medical and psychotherapeutic TAU	during 5 weeks; total 150 min 3. Ongoing					MINI-ICF self; external SLOF
Morimoto et al. (2014)	DSM-IV; SCID; MADRS >15/ HDRS-24 >19; clinical MDD diagnosis	20.26 (severe)	10	74.1 (7.8)	3 bottom-up exercises: low-level auditory tone sweep; phonemic discrimination task (both Brain Fitness cognitive training); low-level visual discrimination exercise (Insight cognitive training); 2 top-down exercises: catch the ball; semantic strategy (both newly developed); adaptive; computerized; individual	30 h during 4 weeks; total 1800 min	33	73.1 (7.0)	Waitlist/TAU: escitalopram; check in with therapist	12 sessions during 12 weeks; assessment after 4 weeks	MADRS
Morimoto et al. (2020)	DSM-IV; SCID; MADRS >15; clinical MDD diagnosis	20.35 (severe)	15	74.7 (7.6)	1. Brain HQ: 3 bottom-up exercises: low-level auditory tone sweep; phonemic discrimination task; low-level visual discrimination exercise; 2 top-down exercises: catch the ball; semantic strategy (both newly developed); adaptive; computerized; individual 2. Stable therapeutic dosage of SSRI/SNRI antidepressant	30 h during 4 weeks; total 1800 min	15	72.2 (9.9)	Placebo: 1. Documentary series with questions. Matched for duration, engagement, reward, presentation, contact; adaptive; computerized; individual 2. Stable therapeutic dosage of SSRI/SNRI antidepressant	30 h during 4 weeks; total 1800 min	MADRS WAIS-IV DS Backward CVLT TMT-B Stroop CWT inhibition WHODAS
Moshier et al. (2015)	BDI >16, <35 (*no formal clinical MDD diagnosis*)	17.17 (severe)	16	32.69 (18.0)	Cognitive control training: modified PASAT; attention control intervention: attending to multiple auditory sources; adaptive; computerized; individual	3 sessions during 2 weeks; total 75 min	16	34.6 (16.7)	Placebo: peripheral vision task which does not target brain regions targeted by cognitive control training; adaptive; computerized; individual	3 sessions during 2 weeks	BDI-II CFQ Hot plates repeated knob-checking task

Moshier & Otto (2017)	DSM-IV; SCID; clinical MDD diagnosis	20.95 (severe)	14	37.2 (14.0)	1. Cognitive control training: modified PASAT; attention control intervention: attending to multiple auditory sources; adaptive; computerized; individual 2. Brief behavioural activation therapy for depression	1. 4 sessions during 4 weeks; total 100 min 2. 4 sessions during 4 weeks	12	33.6 (15.8)	1. Placebo: peripheral vision task which does not target brain regions targeted by cognitive control training; adaptive; computerized; individual 2. Brief behavioural activation therapy for depression	1. 4 sessions during 4 weeks 2. 4 sessions during 4 weeks	BDI-II MADRS RRS
Naismith et al. (2011)	HDRS-17 <20; current or remitted clinical MDD diagnosis	8.00 (moderate)	22	64.8 (8.5)	1. Psycho-education on health, cognitive functioning and cognitive strategies 2. Neuropsychological educational approach to remediation (NEAR) cognitive training: exercises and strategy training; verbal 'bridging' groups; adaptive; partly computerized; group 3. Antidepressant medication	1–2: 10 sessions during 10 weeks; total 1200 min 3. Ongoing	19	64.8 (8.5)	Waitlist/TAU: 1. Waitlist for CR 2. Antidepressant medication	1. 10 weeks 2. Ongoing	HDRS-17 TMT-A RAVLT WMS Logical memory TMT-B D-KEFS Stroop CWT inhibition D-KEFS Sorting WHODAS
Owens et al. (2013)	BDI-II >20 (*no formal clinical MDD diagnosis*)	16.27 (severe)	11	27.7 (5.3)	Attention control training: dual *n*-back task: responding when a visual/audio stimulus matches the visual/audio stimulus (*n*) trials back; adaptive; computerized; individual	8 sessions during 2 weeks; total 240 min	11	22.6 (3.4)	Placebo: dual *n*-back task; non-adaptive; computerized; individual	8 sessions during 2 weeks	BDI-II Change detection task
Pratap et al. (2018)[a]	PHQ-9 >5/ PHQ-9 item 10 >2 (*no formal clinical MDD diagnosis*)	..	40***	33.4 (10.9)	Project: EVOtm cognitive training: guiding a character through an immersive environment, selectively responding to targets; adaptive; computerized; individual	20 sessions during 4 weeks; total 400 min	100	33.6 (12.3)	Placebo: Psycho-education app providing health tips e.g. on self-care; non-adaptive; computerized; individual	28 sessions during 4 weeks	PHQ-9 SDS
Pratap et al. (2018)[b]	PHQ-9 >5/ PHQ-9 item 10 >2 (*no formal clinical MDD diagnosis*)	..	40***	33.37 (10.87)	Project: EVOtm cognitive training: guiding a character through an immersive environment, selectively responding to targets; adaptive; computerized; individual	20 sessions during 4 weeks; total 400 min	100	34.9 (12.3)	Waitlist/TAU: Problem Solving Therapy app (iPST): learning 7 steps to create an action plan; non-adaptive; computerized; individual	28 sessions during 4 weeks	PHQ-9 SDS

(*Continued*)

Table 1. (*Continued.*)

			Cognitive remediation condition				Control condition				
Study	Diagnostic instruments and criteria	HDRS-17 baseline M (severity)	*n*	Mean age (s.d.)	Intervention	Duration	*n*	Mean age (s.d.)	Intervention	Duration	Outcome measures
Semkovska et al. (2015)	DSM-IV; SCID; clinical MDD diagnosis	19.4 (severe)	8	42.4 (14.9)	1. RehaCom cognitive training: divided attention 1 and 2, verbal memory, figural memory, shopping and plan a day; adaptive; computerized; individual 2. Hospitalization for MDD	1. 20 sessions during 5 or 10 weeks; total 12 min 2. Ongoing	7	44.4 (13.0)	Placebo: 1. Free online games requiring attention, strategy, remembering ques; adaptive; computerized; individual 2. Hospitalization for MDD	1. 20 sessions during 5 or 10 weeks 2. Ongoing	HDRS-17 BDI-II D2 test WAIS-III DS Forward WAIS-III Digit symbol coding WAIS-III DS Backward WMS Logical memory ROCF D-KEFS Stroop CWT inhibition; Towers; Sorting; 20-questions; Fluency
Semkovska & Ahern (2017)	DSM-IV; SCID; HDRS-17 <7; remitted clinical MDD diagnosis; remission >8 weeks	4.25 (minimal)	11	45.9 (6.7)	RehaCom cognitive training: divided attention 1 and 2, verbal memory, figural memory, shopping and plan a day; adaptive; computerized; individual	20 sessions during 5 weeks; total 1200 min	10	46.9 (9.3)	Placebo: free online games and word games requiring attention, strategy, remembering ques; adaptive; computerized; individual	20 sessions during 5 weeks	HDRS-17 BDI-II D2 test WAIS-III DS Forward WAIS-III Digit symbol coding WAIS-III DS Backward WMS Logical memory ROCF D-KEFS Towers; Sorting; 20-questions; Fluency
Trapp et al. (2016)	DSM-IV; ICD-10; SCID; clinical MDD diagnosis	12.04 (moderate)	21	34.26 (11.6)	1. X-Cog® cognitive training: game-like, controlling characters facing adventurous challenges, instructions include metacognitive strategies, patients were encouraged to apply and develop strategies; adaptive; partly computerized; partly individual 2. Hospitalization for MDD: intensive treatment: CBT, relaxation treatment,	1. 12 sessions during 4 weeks; total 720 min 2. Ongoing	20	36.9 (12.1)	Waitlist/TAU: hospitalization for MDD: intensive treatment: CBT, relaxation treatment, psychotherapeutic, music therapy, physical training, and occupational therapy	Ongoing; assessment after 4 weeks	HDRS-17 BDI-II Degraded CPT WMS Spat. S. Forward WMS DS Forward TMT-A WMS Spat. S. Backward WMS DS Backward WMS Logical memory WMS Visual reproduction TMT-B WCST

					psychotherapeutic, music therapy, physical training and occupational therapy					
Twamley et al. (2019)	DSM-IV; SCID; clinical MDD diagnosis	15.16 (severe)	16	46.5 (10.5)	1. Skills and strategy training to implement skills to compensate for cognitive difficulties; non-computerized; individual 2. Supported employment services	1. 12 sessions during 12 weeks; total 720 min 2. Ongoing	18	43.5 (13.0)	Waitlist/TAU: supported employment enhanced to match the contact time in the CR condition	Ongoing; assessment after 12 weeks	HDRS-17 CPT-IP TMT-A BACS SCT WMS Spat. S. UM LNS HVLT BVMT-R TMT-B WCST NAB Mazes Category fluency Letter fluency ILSS QOLI SSPA UPSA-Brief MIST
Wanmaker et al. (2014)	BDI-II >10 (*no formal clinical MDD diagnosis*)	14.07 (moderate)	34	20.6 (3.9)	Role playing game cognitive training: walking around in a virtual world, completing working memory tasks to defeat enemies; adaptive; computerized; individual	9 sessions during 3 weeks; total 270 min	27	21.0 (3.3)	Placebo: role playing game: walking around in a virtual world, completing working memory tasks with a low difficulty level to defeat enemies; non-adaptive; computerized; individual	9 sessions during 3 weeks	BDI-II RRS Spanboard task Forward
Wanmaker et al. (2015)	DSM-IV; SCID; clinical MDD diagnosis	20.9 (severe)	10	49.2 (12.7)	Role playing game cognitive training: walking around in a virtual world, completing working memory tasks to defeat enemies; adaptive; computerized; individual	9 sessions during 3 weeks; total 270 min	15	47.3 (12.1)	Placebo: role playing game: walking around in a virtual world, completing working memory tasks with a low difficulty level to defeat enemies; non-adaptive; computerized; individual		BDI-II RRS Internal Shift Task DS Forward DS Backward Reading Span
Yamaguchi et al. (2017)	ICD-10; clinical MDD diagnosis	15.00 (moderate)	4	37.8 (4.9)	Cognitive training based on thinking skills for work; 1. CogPack cognitive training; computerized; individual 2. Sessions discussing cognitive skills, activities and	1–2: 24 sessions during 12 weeks; total 1440 min 3. During 12 months	3	34.7 (6.1)	Waitlist/TAU: traditional vocational services: care manager + community employment services	During 12 months	HDRS-17 BACS SCT BACS Token motor BACS Digit sequencing BACS Verbal memory BACS Tower of London

(*Continued*)

Table 1. (*Continued.*)

Study	Diagnostic instruments and criteria	HDRS-17 baseline M (severity)	Cognitive remediation condition				Control condition				Outcome measures
			n	Mean age (S.D.)	Intervention	Duration	*n*	Mean age (S.D.)	Intervention	Duration	
					compensational strategies; non-computerized; group 3. Psychiatric day care or community employment services						BACS Word fluency BACS Letter fluency GAF

MDD, major depressive disorder; TAU, treatment as usual; PST, Problem Solving Therapy; Abbreviations of clinical instruments, in alphabetical order: BACS, Brief Assessment of Cognition in Schizophrenia; BDI, Beck Depression Inventory; BRIEF-A, Behavior Rating Inventory of Executive Function Adult Version; BVMT-R, Brief Visual Memory Test Revised; CBT, Cognitive Behavioral Therapy; CFQ, Cognitive Failures Questionnaire; COWAT, Controlled Oral Word Association Test; CPT-IP, Continuous Performance Test – Identical Pairs; CVLT, California Verbal Learning Test; Degraded CPT, Degraded Continuous Performance Test; D-KEFS, Delis-Kaplan Executive Functioning System; DS, Digit Span; DSM-IV, Diagnostic and Statistical Manual of mental disorders IV; GAF, Global Assessment of Functioning; HDRS-17, Hamilton Depression Rating Scale-17; HDRS-24, Hamilton Depression Rating Scale-24; HVLT, Hopkins Verbal Learning Test; ICD-10, International Classification of Diseases and related health problems-10; ILSS, Independent Living Skills Survey; LIFE-RIFT, Longitudinal Interval Follow-up Evaluation Range of Impaired Functioning Tool; LNS, Letter Number Sequencing Test; MADRS, Montgomery-Asberg Depression Rating Scale; MINI, Mini International Neuropsychiatric Interview; MINI-ICF, Mini – Internal Classification of Functioning, Disability and Health MIST, Memory for Intentions Test; NAB, Neuropsychological Assessment Battery; PASAT, Paced Auditory Serial Addition Test; PHQ-9, Patient Health Questionnaire-9; PIQ, Performance Intelligence Quotient; RAVLT, Rey Auditory Verbal Learning Test; RDQ, Remission of Depression Questionnaire; ROCF, Rey-Osterrieth Complex Figure test; RRS, Ruminative Response Scale; RS, Resilience Scale; Ruff's 2&7 SAT, Ruff's 2&7 Selective Attention Test; R, Revised; SCID, Structured Clinical Interview for the DSM; SCT, Symbol Coding Task; SDS, Sheehan Disability Scale; SLOF, Specific Level of Functioning Scale; Spat; S, Spatial Span; SSPA, Social Skills Performance Assessment; Stroop CWT, Stroop Color Word Test; TMT-A, Trail Making Test part A; TMT-B, Trail Making Test part B; TOVA, Test of Variables of Attention; VIQ, Verbal Intelligence Quotient; VTS, Vienna Test System; WAIS, Wechsler Adult Intelligence Scale; WCST, Wisconsin Card Sorting Test; WHODAS, World Health Organization Disability Assessment Schedule; WM, Working Memory; WMS, Wechsler Memory Scale; QLDS, Quality of Life in Depression Scale; QOLI, Quality Of Life Interview; UM LNS, University of Maryland Letter Number Span; UPSA-Brief, University of California, San Diego, Performance-based Skills Assessment-Brief.

*In Alvarez et al. (2008), the control sample was split into two (a and b) in order to perform analyses using both cognitive remediation samples. The original control sample consisted of 11 patients. **In Listunova et al. (2020), the control sample was split into two (a and b) in order to perform analyses using both cognitive remediation samples. The original control sample consisted of 19 patients. ***In Pratap et al. (2018) the cognitive remediation sample was split into two (a and b) in order to perform analyses using both control samples. The original cognitive remediation sample consisted of 79 patients.

random-effects model indicated no publication bias either for DS, CF or DF. For DS and DF, the pooled effect sizes were downgraded using the GRADE assessment to very low, and for CF to low certainty of evidence (Appendix III).

Sensitivity analyses

Sensitivity analyses excluding two outliers (Hoorelbeke et al., 2021; Morimoto et al., 2020); studies with high risk of bias (Appendix I – eTable 1), one study with small sample size (Yamaguchi et al., 2017), studies with >14 days from end of treatment to post-intervention assessment (Hoorelbeke et al., 2021; Yamaguchi et al., 2017); or studies with participants without a formal clinical MDD diagnosis (Moshier et al., 2015; Owens et al., 2013; Pratap et al., 2018; Wanmaker et al., 2014) did not affect the overall results on the three main outcome categories DS, CF and DF. Further, excluding studies with unclear or insufficient sequence generation (Appendix I – eTable 1) did not affect the results on CF. However, the effect on DS changed from a small significant effect ($g = 0.28$) to a non-significant, smaller effect ($g = 0.18$; 95% CI −0.05 to 0.42; $N = 12$; $n = 204$; $I^2 = 31\%$; 95% CI 0–65%), and the effect on DF changed from a small significant effect ($g = 0.22$) to a non-significant, smaller effect as well ($g = 0.20$; 95% CI −0.03 to 0.43; $N = 8$; $n = 295$; $I^2 = 0\%$; 95% CI 0–68%). Thus, confounding bias might have affected the effects on DS and DF.

Discussion

We performed a meta-analysis to assess the effectiveness of cognitive remediation in depression. Our results indicate a small significant effect on DS, a medium significant effect on CF and a small significant effect on DF. Significant effects for CF domains were small for Attention, Processing speed and Verbal learning and memory, and medium for Working memory. For DF subcategories, there was a small significant effect for Subjective DF. However, we found no indication that these beneficial effects are sustainable as the meta-analysis did not identify any significant effects of cognitive remediation on DS, CF or DF at follow-ups up to 3 months after the post-intervention assessments.

Our findings of small significant effects on DS and DF, and medium significant effect on Working memory are consistent with the only previous meta-analysis on the subject (Motter et al., 2016). However, for Attention, Motter et al. (2016) did find a moderate significant effect whereas we identified a small significant effect. Processing speed outcomes were not meta-analysed separately but merged with Attention outcomes in the previous work (Motter et al., 2016), whereas we have quantified separately the effects of these two cognitive domains. Further, in contrast to our findings, they found no significant effect on Verbal memory. These differences might be explained by a limited number of studies and participants included in the previous meta-analysis, relative to the present meta-analysis.

In subgroup analyses, we found that effects on DS were significantly larger in the subgroup with patients with severe depressive baseline symptoms compared to those with moderate symptoms: there was a small significant effect in patients with severe depressive baseline symptoms, and no significant or sizable effect in patients with moderate symptoms. This is not surprising, since more depressive symptoms mean more room for improvement. This finding emphasizes the importance of taking baseline symptoms into account, as has been argued extensively by others before

https://doi.org/10.1017/S0033291721001100 Published online by Cambridge University Press

Table 2. Main effects and subgroup analyses of cognitive remediation on depressive symptomatology, cognitive and daily functioning

	N	*n*	Hedges' g (95% CI)	*p* value	I^2 (95% CI)	*p* value*
Depressive symptomatology	21	899	0.28 (0.09–0.46)	0.004	40% (0–64%)	
Clinical status						
Current depression	16	670	0.35 (0.13–0.57)	0.002	35% (0–64%)	0.384
Remitted depression	4	188	0.13 (−0.27 to 0.54)	0.522	59% (0–86%)	
Control condition						
Placebo	11	515	0.19 (−0.06 to 0.44)	0.130	33% (0–67%)	0.317
Waitlist/TAU	10	384	0.39 (0.10–0.67)	0.008	47% (0–75%)	
Symptom severity						
Moderate	6	250	−0.03 (−0.36 to 0.30)	0.854	0% (0–75%)	0.022
Severe	11	281	0.48 (0.19–0.76)	0.001	44% (0–72%)	
Diagnosis						
Clinical MDD	16	519	0.32 (0.09–0.55)	0.006	44% (0–69%)	0.482
No clinical diagnosis	5	380	0.18 (−0.16 to 0.51)	0.298	23% (0–69%)	
Cognitive functioning	19	597	0.60 (0.37–0.83)	<0.001	44% (3–67%)	
Clinical status						
Current depression	12	310	0.54 (0.24–0.83)	<0.001	4% (0–60%)	0.358
Remitted depression	6	246	0.76 (0.39–1.14)	<0.001	68% (23–86%)	
Control condition						
Placebo	8	317	0.84 (0.56–1.12)	<0.001	48% (0–77%)	0.025
Waitlist/TAU	11	280	0.39 (0.12–0.66)	0.005	0% (0–60%)	
Symptom severity						
Moderate	8	308	0.54 (0.21–0.88)	0.002	57% (5–80%)	0.932
Severe	9	201	0.52 (0.16–0.88)	0.004	22% (0–63%)	
Cognitive functioning per domain						
Attention	11	322	0.36 (0.07–0.66)	0.016	42% (0–71%)	
Processing speed	10	262	0.26 (0.02–0.50)	0.033	0% (0–62%)	
Working memory	15	463	0.54 (0.22–0.86)	0.001	64% (37–79%)	
Verbal learning and memory	11	292	0.47 (0.08–0.87)	0.019	64% (31–81%)	
Visual learning and memory	7	210	0.12 (−0.17 to 0.41)	0.414	13% (0–75%)	
Executive functioning	11	292	0.23 (−0.00 to 0.46)	0.053	0% (0–60%)	
Verbal fluency	6	122	0.26 (−0.23 to 0.76)	0.301	47% (0–79%)	
Daily functioning	12	646	0.22 (0.06–0.39)	0.008	3% (0–60%)	
Clinical status						
Current depression	7	404	0.24 (0.01–0.48)	0.040	36% (0–73%)	0.889
Remitted depression	4	201	0.27 (−0.03 to 0.57)	0.077	0% (0–85%)	
Control condition						
Placebo	5	345	0.27 (0.03–0.51)	0.025	57% (0–84%)	0.578
Waitlist/TAU	7	301	0.17 (−0.08 to 0.43)	0.180	0% (0–71%)	
Symptom severity						
Moderate	5	182	0.18 (−0.21 to 0.56)	0.368	0% (0–79%)	0.697
Severe	4	117	0.29 (−0.15 to 0.73)	0.198	66% (0–88%)	
Daily functioning per sub-category						
Subjective	11	639	0.22 (0.05–0.39)	0.012	8% (0–63%)	
Objective	4	94	0.05 (−0.36 to 0.45)	0.820	0% (0–85%)	

N, number of comparisons; *n*, number of patients; CI, confidence interval; TAU, treatment as usual.
*This *p* value indicates the between-group difference in the subgroup analyses.

https://doi.org/10.1017/S0033291721001100 Published online by Cambridge University Press

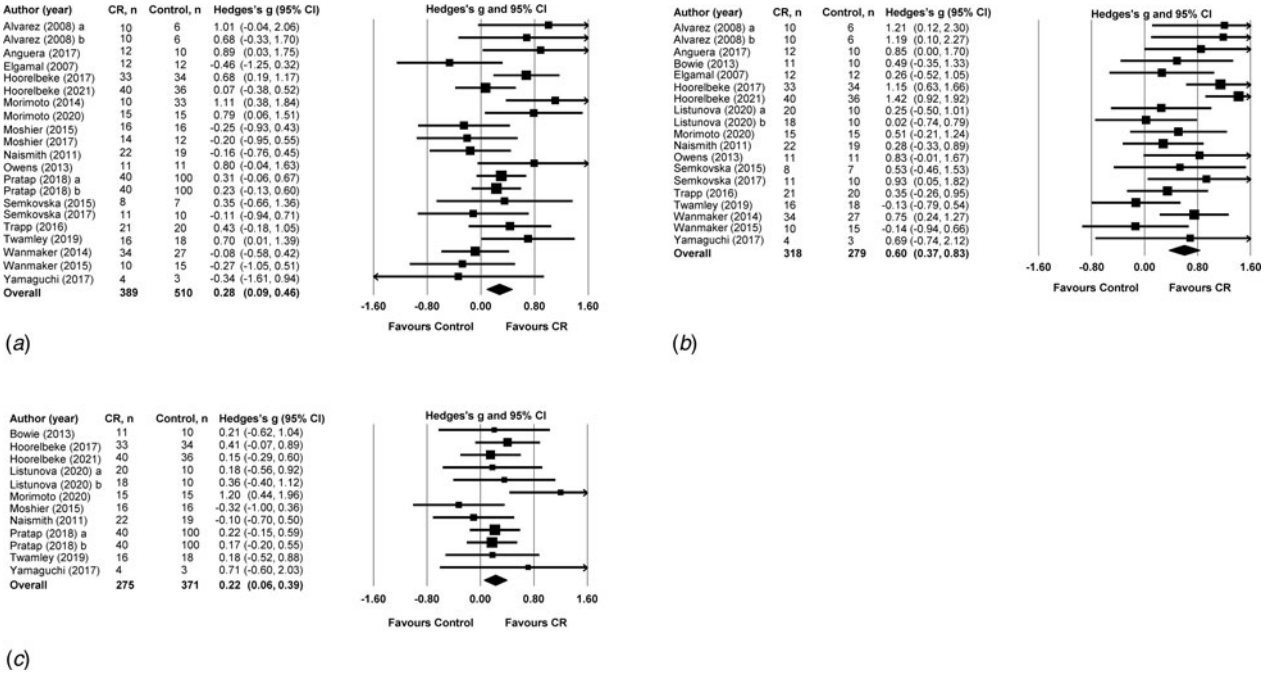

Fig. 2. Forest plots of three main outcomes: (*a*) Forest plot effect on depressive symptomatology, (*b*) Forest plot effect on cognitive functioning, (*c*) Forest plot effect on daily functioning. CR, cognitive remediation; CI, confidence interval.

(e.g. Nunes et al., 2011). Future studies should consider that only in severely depressed individuals, cognitive remediation appears to improve DS. Moreover, this finding suggests that patient characteristics impact on the effectiveness of cognitive remediation. Gaining more knowledge on the association between individual patient characteristics and the effectiveness of cognitive remediation may ultimately lead to personalized cognitive remediation interventions. The potential in this area should be noted.

Although CF improved in comparison to both placebo and waitlist/TAU control conditions, improvement was significantly more pronounced in comparison to placebo than in comparison to waitlist/TAU with large and moderate effect size, respectively. This could be explained by that placebo control conditions were by definition specifically designed in order *not* to improve CF, while this was not the case for waitlist and/or TAU control conditions. We could not demonstrate any significant effect of clinical status (current *v.* remitted depression) or diagnosis (clinical MDD *v.* no clinical diagnosis).

Further, we found no moderating effect of cognitive remediation duration on the effectiveness achieved. Relatively short programmes may be sufficiently effective. However, the absence of a duration effect might have been attributable to the variation in the design of the remediation programmes.

Some limitations of the meta-analysis should be noted. A limitation with regard to the effect on DF is that this outcome measure was very heterogeneous in terms of what instruments were used and what these instruments aimed to measure. Although we tried to categorize DF outcomes as subjective and objective in order to promote homogeneity, there was still a great variety of outcomes included within these categories. Statistical heterogeneity was, however, low ($I^2 = 3\%$). The same could be said for CF, because instruments aimed at assessing various cognitive domains were included though all instruments explicitly aimed to assess CF. Although most studies included had low risk of bias and excluding studies with high risk of bias did not change the results, some studies reported non-random or unclear sequence generation. Our results indicate that confounding in the observational studies may have biased the results for DS and DF: when the analyses on DS and DF were restricted to randomized studies, the effect sizes were lower and no longer significant. According to the GRADE assessment, the pooled effect size for DS and DF was downgraded to very low, and for CF to low certainty of evidence. Both including varying cognitive remediation interventions and studies among patients with a broad range of depression severity likely improves the generalizability of our results. However, the other side of the coin is that such liberal inclusion decreases the specificity with which our results apply to a specific cognitive remediation format and specific population. Notably, interventions were not only diverse qualitatively, but also the quantity (duration) of cognitive remediation varied considerably. Unfortunately, there were not enough studies on cognitive remediation interventions with a fully non-computerized, or group format to perform any subgroup analyses on therapy delivery format as we aimed to. Our findings should be interpreted cautiously, keeping in mind that the vast majority of included studies had a fully computerized and individual format, although some studies combined computerized and non-computerized, and individual and group interventions. There were not enough studies to include a subgroup with minimal depressive symptoms in any of the subgroup analyses on symptom severity at baseline, or to perform subgroup analyses on diagnosis for CF and DF.

Furthermore, cognitive impairment has been shown to increase with the number of depressive episodes (Semkovska et al., 2019), and thus cognitive remediation might be especially relevant for patients with recurrent depression. However, none of the included studies recruited exclusively patients with recurrent depression. Also, only two of the included studies report

https://doi.org/10.1017/S0033291721001100 Published online by Cambridge University Press

evident cognitive impairment at baseline as an inclusion criterion (Listunova et al., 2020; Yamaguchi et al., 2017). Similarly to the larger effect on DS found in patients with severe depressive symptoms at baseline, the effects of cognitive remediation might be more pronounced in patients with evident cognitive impairment. Both these factors might have impacted the current meta-analysis results. It would be relevant for future studies to focus on the effectiveness of cognitive remediation specifically in samples with recurrent depression and/or evident cognitive impairment, and to study whether effects are different compared to samples with single-episode depression and/or no evident cognitive impairment. Further, the number of studies that provided follow-up data was limited. It should also be noted that sample sizes were often small. Future studies should include more participants and thereby increase power.

In conclusion, our findings indicate that cognitive remediation in depression substantially improves CF; more specifically, Attention, Processing speed, Working memory and Verbal learning and memory. We found small significant effects on DS and subjective DF as well. However, these might be overestimations due to confounding bias. Further, our findings indicate that it is important to consider baseline depressive symptom severity: cognitive remediation improved DS in those with severe baseline symptoms but not in those with moderate baseline symptoms. The effects on DS, CF and DF disappeared at follow-up. Given that the endurance of the effects of cognitive remediation is under discussion, it is critical to study how interventions can be innovated or combined with other interventions in order for their effects to last. Development of cognitive remediation protocols that aim for sustainable effects is crucial. More high-quality, well-powered, randomized controlled trials are needed that include long-term follow-ups. The effect of cognitive remediation on DS, DF, as well as optimal therapy delivery format needs to be determined.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291721001100.

Acknowledgements. The authors thank the authors who replied to their e-mail requests for additional data and/or information.

Author contributions. AML, MS, MB, DD and CLB designed the study. AML searched the databases. AML and MB screened records for inclusion and conflicts were dissolved through discussion with MS. Full-text screening was done by AML, MS and MB and conflicts were dissolved trough discussion among them. AML corresponded with the authors in case of missing data/information. AML extracted data supervised by MB and MS. MB and MS cross-checked data extraction. AML analysed and interpreted the data, supervised by MS, MB, GJG, HB and CLB. MB cross-checked the analyses. AML drafted the manuscript. All authors reviewed and revised the manuscript. The final manuscript was approved for submission by all authors.

Financial support. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Conflict of interest. None.

Notes

[1] Three studies reported on multiple relevant comparisons. Two studies reported on two relevant cognitive remediation samples and one control sample, thus both cognitive remediation samples were included, each compared to half of the control sample [Alvarez et al. (2008) a and b; Listunova et al. (2020) a and b]. Another study reported on one cognitive remediation sample and two relevant control samples: both control samples were included, each compared to half of the cognitive remediation sample [Pratap et al. (2018) a and b].

[2] Baseline depressive symptom severity, no effect on DS (coefficient: 0.02; 95% CI −0.02 to 0.07; $p = 0.251$), CF (coefficient: −0.03; 95% CI −0.07 to 0.02; $p = 0.226$) or DF (coefficient: 0.02; 95% CI −0.03 to 0.07; $p = 0.458$); age, no effect on DS (coefficient: 0.01; 95% CI −0.01 to 0.02; $p = 0.300$), CF (coefficient: −0.01; 95% CI −0.02 to 0.01; $p = 0.343$) or DF (coefficient: 0.02; 95% CI −0.00 to 0.04; $p = 0.112$); gender, no effect on DS (coefficient: 0.01; 95% CI −0.01 to 0.03; $p = 0.391$), CF (coefficient: 0.00; 95% CI −0.02 to 0.02; $p = 0.707$) or DF (coefficient: 0.01; 95% CI −0.02 to 0.03; $p = 0.729$); cognitive remediation duration, no effect on DS (coefficient: 0.00; 95% CI −0.00 to 0.00; $p = 0.351$), CF (coefficient: −0.00; 95% CI −0.00 to 0.00; $p = 0.182$) or DF (coefficient: 0.00; 95% CI −0.00 to 0.00; $p = 0.399$).

[3] At follow-up, no effect on DS ($g = 0.15$; 95% CI −0.13 to 0.43; $p = 0.297$; $N = 7$; $n = 454$; $I^2 = 34\%$: 95% CI 0–72%), CF ($g = 0.08$; 95% CI −0.65 to 0.81; $p = 0.836$; $N = 3$; $n = 126$; $I^2 = 68\%$: 95% CI 0–91%) or DF ($g = 0.03$; 95% CI −0.25 to 0.32; $p = 0.813$; $N = 4$; $n = 381$; $I^2 = 27\%$: 95% CI 0–73%).

References

Adler, D. A., McLaughlin, T. J., Rogers, W. H., Chang, H., Lapitsky, L., & Lerner, D. (2006). Job performance deficits due to depression. *American Journal of Psychiatry*, *163*(9), 1569–1576. https://doi.org/10.1176/ajp.2006.163.9.1569.

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. *Clinical Psychology in Europe*, *1*(3), 1–35. https://doi.org/10.32872/cpe.v1i3.34396.

Ahern, E., & Semkovska, M. (2017). Cognitive functioning in the first-episode of major depressive disorder: A systematic review and meta-analysis. *Neuropsychology*, *31*(1), 52–72. https://doi.org/10.1037/neu0000319.

Alvarez, L. M., Cortés Sotres, J. F., León, S. O., Estrella, J., & Sánchez Sosa, J. J. (2008). Computer program in the treatment for major depression and cognitive impairment in university students. *Computers in Human Behavior*, *24*(3), 816–826. https://doi.org/10.1016/j.chb.2007.02.013.

Anguera, J. A., Gunning, F. M., & Areán, P. A. (2017). Improving late life depression and cognitive control through the use of therapeutic video game technology: A proof-of-concept randomized trial. *Depression and Anxiety*, *34*(6), 508–517. https://doi.org/10.1002/da.22588.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). BDI-II, Beck Depression Inventory: Manual. Psychological Corp.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester: John Wiley & Sons, Ltd.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2013). Comprehensive meta-analysis (Version 3) [computer software]. Englewood, N.J.: Biosat Inc.

Bowie, C. R., Gupta, M., Holshausen, K., Jokic, R., Best, M., & Milev, R. (2013). Cognitive remediation for treatment-resistant depression. *The Journal of Nervous and Mental Disease*, *201*(8), 680–685. https://doi.org/10.1097/NMD.0b013e31829c5030.

Cella, M., Price, T., Corboy, H., Onwumere, J., Shergill, S., & Preti, A. (2020). Cognitive remediation for inpatients with psychosis: A systematic review and meta-analysis. *Psychological Medicine*, *50*(7), 1062–1076. https://doi.org/10.1017/S0033291720000872.

Cohen J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

de Jonge, M., Bockting, C. L. H., van Oppen, P., Van, H. L., Peen, J., Kikkert, M. J., & Dekker, J. J. M. (2018). The association between the number of previous episodes and modifiable vulnerability factors in remitted patients with recurrent depression. *PLoS ONE*, *13*(11), e0206495. https://doi.org/10.1371/journal.pone.0206495.

Elgamal, S., McKinnon, M. C., Ramakrishnan, K., Joffe, R. T., & MacQueen, G. (2007). Successful computer-assisted cognitive remediation therapy in patients with unipolar depression: A proof of principle study. *Psychological Medicine*, *37*(9), 1229–1238. https://doi.org/10.1017/S0033291707001110.

Guyatt, G., Oxman, A. D., Akl, E. A., Kunz, R., Vist, G., Brozek, J., … DeBeer, H. (2011). GRADE guidelines: 1. Introduction – GRADE evidence profiles

and summary of findings tables. *Journal of Clinical Epidemiology*, *64*(4), 383–394. https://doi.org/10.1016/j.jclinepi.2010.04.026.

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology*, *6*(4), 278–296. https://doi.org/10.1111/j.2044-8260.1967.tb00530.x.

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2019). Doing meta-analysis in R: A hands-on guide. PROTECT Lab Erlangen. doi: 10.5281/zenodo.2551803.

Heo, M., Murphy, C. F., & Meyers, B. S. (2007). Relationship between the Hamilton depression rating scale and the Montgomery-Åsberg depression rating scale in depressed elderly: A meta-analysis. *The American Journal of Geriatric Psychiatry*, *15*(10), 899–905. https://doi.org/10.1097/JGP.0b013e318098614e.

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2020). Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (updated September 2020). Cochrane. https://training.cochrane.org/handbook.

Hoorelbeke, K., & Koster, E. H. W. (2017). Internet-delivered cognitive control training as a preventive intervention for remitted depressed patients: Evidence from a double-blind randomized controlled trial study. *Journal of Consulting and Clinical Psychology*, *85*(2), 135–146. https://doi.org/10.1037/ccp0000128.

Hoorelbeke, K., van den Bergh, N., de Raedt, R., Wichers, M., & Koster, E. (2021). Preventing recurrence of depression: Long-term effects of a randomized controlled trial on cognitive control training for remitted depressed patients. *Clinical Psychological Science*. Retrieved from http://hdl.handle.net/1854/LU-8683449.

Jaeger, J., Berns, S., Uzelac, S., & Davis-Conway, S. (2006). Neurocognitive deficits and disability in major depressive disorder. *Psychiatry Research*, *145*(1), 39–48. https://doi.org/10.1016/j.psychres.2005.11.011.

Jahshan, C., Rassovsky, Y., & Green, M. F. (2017). Enhancing neuroplasticity to augment cognitive remediation in schizophrenia. *Frontiers in Psychiatry*, *8*, 1–7. https://doi.org/10.3389/fpsyt.2017.00191.

Keyes, K. M., Platt, J., Kaufman, A. S., & McLaughlin, K. A. (2017). Association of fluid intelligence and psychiatric disorders in a population-representative sample of US adolescents. *JAMA Psychiatry*, *74*(2), 179. https://doi.org/10.1001/jamapsychiatry.2016.3723.

Lezak, M., Howieson, D., Bigler, E., & Tranel, D. (2012). *Neuropsychological assessment* (5ht Ed). New York, NY: Oxford University Press.

Listunova, L., Kienzle, J., Bartolovic, M., Jaehn, A., Grützner, T. M., Wolf, R. C., … Roesch-Ely, D. (2020). Cognitive remediation therapy for partially remitted unipolar depression: A single-blind randomized controlled trial. *Journal of Affective Disorders*, *276*, 316–326. https://doi.org/10.1016/j.jad.2020.07.008.

Majer, M., Ising, M., Künzel, H., Binder, E. B., Holsboer, F., Modell, S., & Zihl, J. (2004). Impaired divided attention predicts delayed response and risk to relapse in subjects with depressive disorders. *Psychological Medicine*, *34*(8), 1453–1463. https://doi.org/10.1017/S0033291704002697.

McIntyre, R. S., Cha, D. S., Soczynska, J. K., Woldeyohannes, H. O., Gallaugher, L. A., Kudlow, P., … Baskaran, A. (2013). Cognitive deficits and functional outcomes in major depressive disorder: Determinants, substrates, and treatment interventions. *Depression and Anxiety*, *30*(6), 515–527. https://doi.org/10.1002/da.22063.

Moffitt, T., Caspi, A., Taylor, A., Kokaua, J., Milne, B. J., Polanczyk, G., & Poulton, R. (2010). How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. *Psychological Medicine*, *6*(40), 899–909. https://doi.org/10.1017/S0033291709991036.

Montgomery, S. A., & Åsberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, *134*(4), 382–389. https://doi.org/10.1192/bjp.134.4.382.

Morimoto, S. S., Altizer, R. A., Gunning, F. M., Hu, W., Liu, J., Cote, S. E., … Alexopoulos, G. S. (2020). Targeting cognitive control deficits with neuroplasticity-based computerized cognitive remediation in patients with geriatric major depression: A randomized, double-blind, controlled trial. *The American Journal of Geriatric Psychiatry*, *28*(9), 971–980. https://doi.org/10.1016/j.jagp.2020.05.023.

Morimoto, S. S., Wexler, B. E., Liu, J., Hu, W., Seirup, J., & Alexopoulos, G. S. (2014). Neuroplasticity-based computerized cognitive remediation for treatment-resistant geriatric depression. *Nature Communications*, *5*(1), 4579. https://doi.org/10.1038/ncomms5579.

Moshier, S. J., Molokotos, E. K., Stein, A. T., & Otto, M. W. (2015). Assessing the effects of depressed mood and cognitive control training on memory confidence and accuracy following repeated checking. *International Journal of Cognitive Therapy*, *8*(3), 206–221. https://doi.org/10.1521/ijct_2015_8_02.

Moshier, S. J., & Otto, M. W. (2017). Behavioral activation treatment for major depression: A randomized trial of the efficacy of augmentation with cognitive control training. *Journal of Affective Disorders*, *210*, 265–268. https://doi.org/10.1016/j.jad.2017.01.003.

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. *Journal of Affective Disorders*, *189*, 184–191. https://doi.org/10.1016/j.jad.2015.09.022.

Naismith, S. L., Diamond, K., Carter, P. E., Norrie, L. M., Redoblado-Hodge, M. A., Lewis, S. J. G., & Hickie, I. B. (2011). Enhancing memory in late-life depression: The effects of a combined psychoeducation and cognitive training program. *The American Journal of Geriatric Psychiatry*, *19*(3), 240–248. https://doi.org/10.1097/JGP.0b013e3181dba587.

Nunes, E. V., Pavlicova, M., Hu, M.-C., Campbell, A. N., Miele, G., Hien, D., & Klein, D. F. (2011). Baseline matters: The importance of covariation for baseline severity in the analysis of clinical trials. *The American Journal of Drug and Alcohol Abuse*, *37*(5), 446–452. https://doi.org/10.3109/00952990.2011.596980.

Owens, M., Koster, E. H. W., & Derakshan, N. (2013). Improving attention control in dysphoria through cognitive training: Transfer effects on working memory capacity and filtering efficiency. *Psychophysiology*, *50*(3), 297–307. https://doi.org/10.1111/psyp.12010.

Pratap, A., Renn, B. N., Volponi, J., Mooney, S. D., Gazzaley, A., Arean, P. A., & Anguera, J. A. (2018). Using mobile apps to assess and treat depression in Hispanic and Latino populations: Fully remote randomized clinical trial. *Journal of Medical Internet Research*, *20*(8), e10130. https://doi.org/10.2196/10130.

Robertson, I. H., & Murre, J. M. J. (1999). Rehabilitation of brain damage: Brain plasticity and principles of guided recovery. *Psychological Bulletin*, *125*(5), 544–575. https://doi.org/10.1037/0033-2909.125.5.544.

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: A systematic review and meta-analysis. *Psychological Medicine*, *44*(10), 2029–2040. https://doi.org/10.1017/S0033291713002535.

Semkovska, M., & Ahern, E. (2017). Online neurocognitive remediation therapy to improve cognition in community-living individuals with a history of depression: A pilot study. *Internet Interventions*, *9*, 7–14. https://doi.org/10.1016/j.invent.2017.04.003.

Semkovska, M., Lambe, S., Lonargáin, D. Ó., & McLoughlin, D. M. (2015). Neurocognitive remediation therapy for depression. *The Journal of Nervous and Mental Disease*, *203*(8), 609–616. https://doi.org/10.1097/NMD.0000000000000337.

Semkovska, M., Quinlivan, L., O'Grady, T., Johnson, R., Collins, A., O'Connor, J., … Gload, T. (2019). Cognitive function following a major depressive episode: A systematic review and meta-analysis. *The Lancet Psychiatry*, *6*, 851–861. https://doi.org/10.1016/S2215-0366(19)30291-3.

ten Doesschate, M. C., Bockting, C. L. H., Koeter, M. W. J., & Schene, A. H. (2010). Prediction of recurrence in recurrent depression. *The Journal of Clinical Psychiatry*, *71*(08), 984–991. https://doi.org/10.4088/JCP.08m04858blu.

Trapp, W., Engel, S., Hajak, G., Lautenbacher, S., & Gallhofer, B. (2016). Cognitive remediation for depressed inpatients: Results of a pilot randomized controlled trial. *Australian & New Zealand Journal of Psychiatry*, *50*(1), 46–55. https://doi.org/10.1177/0004867415622271.

Twamley, E. W., Thomas, K. R., Burton, C. Z., Vella, L., Jeste, D. V., Heaton, R. K., & McGurk, S. R. (2019). Compensatory cognitive training for people with severe mental illnesses in supported employment: A randomized controlled trial. *Schizophrenia Research*, *203*, 41–48. https://doi.org/10.1016/j.schres.2017.08.005.

Twamley, E. W., Vella, L., Burton, C. Z., Heaton, R. K., & Jeste, D. V. (2012). Compensatory cognitive training for psychosis. *The Journal of Clinical Psychiatry*, *73*(09), 1212–1219. https://doi.org/10.4088/JCP.12m07686.

Vittengl, J. R., Clark, L. A., Kraft, D., & Jarrett, R. B. (2005). Multiple measures, methods, and moments: A factor-analytic investigation of change in depressive symptoms during acute-phase cognitive therapy for depression. *Psychological Medicine*, *35*(5), 693–704. https://doi.org/10.1017/S0033291704004143.

Wanmaker, S., Geraerts, E., & Franken, I. H. A. (2015). A working memory training to decrease rumination in depressed and anxious individuals: A double-blind randomized controlled trial. *Journal of Affective Disorders*, *175*, 310–319. https://doi.org/10.1016/j.jad.2014.12.027.

Wanmaker, S., Hopstaken, J.F., Asselbergs, J., Geraerts, E., & Franken, I.H.A. (2014). Decreasing dysphoric thoughts by a working memory training: A randomized double-blind placebo-controlled trial. *Journal of Depression and Anxiety*, *03*(04), 1–8. https://doi.org/10.4172/2167-1044.1000165.

Yamaguchi, S., Sato, S., Horio, N., Yoshida, K., Shimodaira, M., Taneda, A., … Ito, J. (2017). Cost-effectiveness of cognitive remediation and supported employment for people with mental illness: A randomized controlled trial. *Psychological Medicine*, *47*(1), 53–65. https://doi.org/10.1017/S0033291716002063.

https://doi.org/10.1017/S0033291721001100 Published online by Cambridge University Press