Neuropsychology

Are Cognitive Interventions Effective in Alzheimer's Disease? A Controlled Meta-Analysis of the Effects of Bias

Javier Oltra-Cucarella, Rubén Pérez-Elvira, Raul Espert, and Anita Sohn McCormick

Online First Publication, April 7, 2016. http://dx.doi.org/10.1037/neu0000283

Neuropsychology

0894-4105/16/$12.00 http://dx.doi.org/10.1037/neu0000283

Are Cognitive Interventions Effective in Alzheimer's Disease? A Controlled Meta-Analysis of the Effects of Bias

Javier Oltra-Cucarella
Hospital General Universitario Santa María del Rosell, Cartagena, Murcia, Spain

Rubén Pérez-Elvira
NEPSA Rehabilitación Neurológica, Salamanca, Spain

Raul Espert
University of Valencia

Anita Sohn McCormick
Texas A&M University

Objective: There is limited evidence about the efficacy of cognitive interventions for Alzheimer's disease (AD). However, aside from the methodological quality of the studies analyzed, the methodology used in previous meta-analyses is itself a risk of bias as different types of effect sizes (*ES*s) were calculated and combined. This study aimed at examining the results of nonpharmacological interventions for AD with an adequate control of statistical methods and to demonstrate a different approach to meta-analysis. ***Method:*** ESs were calculated with the independent groups pre/post design. Average *ES*s for separate outcomes were calculated and moderator analyses were performed so as to offer an overview of the effects of bias. ***Results:*** Eighty-seven outcomes from 19 studies ($n = 812$) were meta-analyzed. *ES*s were small on average for cognitive and functional outcomes after intervention. Moderator analyses showed no effect of control of bias, although *ES*s were different from zero only in some circumstances (e.g., memory outcomes in randomized studies). Cognitive interventions showed no more efficacy than placebo interventions, and functional *ES*s were consistently low across conditions. ***Conclusions:*** cognitive interventions delivered may not be effective in AD probably due to the fact that the assumptions behind the cognitive interventions might be inadequate. Future directions include a change in the type of intervention as well as the use of outcomes other than standardized tests. Additional studies with larger sample sizes and different designs are needed to increase the power of both primary studies and meta-analyses.

Keywords: Alzheimer's Disease, nonpharmacological interventions, memory, dementia, activities of daily functioning

Alzheimer's disease (AD) is a neurodegenerative condition resulting in significant neuronal loss and a decline in the function of various neurotransmitter systems (Casey, Antimisiaris & O'Brien, 2010). These brain changes lead to cognitive dysfunction, neuropsychiatric symptoms, and difficulties with activities of daily living (ADLs; Burns & Iliffe, 2009). Memory loss is commonly the first symptom in AD manifesting insidiously as an inability to form new memories (Burns & Iliffe, 2009). Dementia due to AD is formally diagnosed when cognitive or behavioral symptoms represent a significant decline from previous levels of functioning and interfere significantly with performance on ADLs such as working or personal responsibilities (McKhann et al., 2011).

Given that the symptoms of AD result from brain changes and neurotransmitter deficiencies, pharmacological treatments that attempt to ameliorate the cognitive symptoms have been investigated (Casey et al., 2010; Citron, 2010). However, pharmacological treatment has not been shown to have a clinically relevant effect despite statistically significant effects on clinical trials (Lanctôt, Rajaram, & Herrmann, 2009). Thus, nonpharmacological treatments tailored to address cognitive, behavioral, and psychological symptoms have received much attention in recent years (Brodaty & Arasaratnam, 2012; Gardette, Coley, & Andrieu, 2010).

Cognitive intervention is an alternative to pharmacological treatment for people with AD. There are three types of cognitive intervention: cognitive training, cognitive stimulation, and cogni-

Javier Oltra-Cucarella, Unit of Neuropsychology, Hospital General Universitario Santa María del Rosell, Cartagena, Murcia, Spain; Rubén Pérez-Elvira, NEPSA Rehabilitación Neurológica, Salamanca, Spain; Raul Espert, Department of Psychobiology, University of Valencia; Anita Sohn McCormick, Department of Educational Psychology, Texas A&M University.

We kindly thank Professor Scott Morris, Professor Elizabeth Tipton, and two anonymous reviewers for their suggestions with statistical analyses, effect sizes, and robust variance estimation and Dr. Giorgio Arcara, Dr. Rosa Manenti, Dr. Zoë Hoare, Dr. Elisabetta Farina, and Dr. Orestes Forlenza for their help in gathering the data necessary to perform this meta-analysis. We would like to acknowledge the assistance of Mary Beth Spitznagel and the International Neuropsychological Society's International Liaison Committee's Research Editing and Consulting Program. The authors disclose no conflicts of interest.

Correspondence concerning this article should be addressed to Javier Oltra-Cucarella, Unit of Neuropsychology, Hospital General Universitario Santa María del Rosell, Service of Neurology, Cartagena, Murcia, Spain. E-mail: javi.oltra@cop.es

tive rehabilitation. Cognitive stimulation involves engagement in group activities such as conversation aimed at enhancing cognitive and social functioning (Clare, 2003). Cognitive training involves the use of guided standardized tasks through which people with AD are taught theory-based strategies and abilities to improve their cognitive functioning. This type of intervention can be conducted both individually as well as in a group and can also include computer-based training (Ballard, Khan, Clack, & Corbett, 2011; Kurz et al., 2012), and varying difficulty levels are offered based on the severity of cognitive impairment. Cognitive rehabilitation (also referred to as cognitive remediation), unlike cognitive training and cognitive stimulation programs, aims to develop and implement cognitive strategies applied to specific ADLs through the use of task-specific techniques and external supports (Grandmaison & Simard, 2003; Kurz et al., 2012). Through cognitive rehabilitation programs, people with AD can be taught face-name associations, routes, the use of an external memory aid, or even ameliorate behavioral difficulties (Clare, 2003).

The efficacy of cognitive interventions for AD has been examined recently through three excellent meta-analyses (Clare, Woods, Moniz Cook, Orrel, & Spector, 2003; Olazarán et al., 2010; Sitzer, Twamley, & Jeste, 2006) and two successive updates (Bahar-Fuchs, Clare, & Woods, 2013; Clare & Woods, 2008). The conclusions of these three meta-analyses are quite different and even contradictory. After reviewing 11 randomized controlled trials (RCTs), Bahar-Fuchs et al. (2013) concluded that "no positive or adverse effects for cognitive training were detected" (p. 25). Sitzer et al. (2006) concluded that cognitive training produced large effects on verbal memory after reviewing 19 controlled trials. After reviewing 179 studies, Olazarán et al. (2010) concluded that there is a Grade B (consistent low-quality RCTs) recommendation for cognitive training and cognitive stimulation to improve cognition (e.g., memory) in people with dementia. So why are the results from these studies, apparently investigating the same construct, so different?

One explanation might be the methodology used to conduct the meta-analysis and calculate *ES*s. Sitzer et al. (2006) and Bahar-Fuchs et al. (2013) calculated *ES*s using change scores and change standard deviations, but then calculated *ES*s with raw scores for part of the outcomes analyzed. In contrast, Olazarán et al. (2010) calculated *ES*s using change scores and odds ratios. When *ES*s are calculated using change scores, both pre- and posttest standard deviations are included in the formula (Borenstein, Hedges, Higgins, & Rothstein, 2009). However, when *ES*s are calculated using raw scores, only pretest standard deviations are used, so combining *ES*s derived from both methods is incompatible. Combining pre- and posttest standard deviations will produce comparable *ES*s when the variability of scores is constant across time periods (Morris & DeShon, 2002), which is unlikely when a treatment group is compared to a group with different treatments or no treatment at all and a Group $\times$ Time interaction is expected.

Another explanation for the varying findings is the correlation between *ES*s within each study used to calculate the average *ES*, which is twofold. On the one hand, the pre/post correlation of each individual measure is needed to calculate the repeated measures *ES*. On the other hand, studies analyzing the effect of a treatment include both multiple groups and multiple outcomes, which means that *ES*s reported are not independent. Bahar-Fuchs et al. (2013) assumed that the correlation between pre- and posttest scores was zero, whereas Sitzer et al. (2006) and Olazarán et al. (2010) did not report how they controlled for the pre/post correlation. Statistical analyses exist that allow to analyzing the effects of correlational coefficients on *ES*s and to modeling for covariance structures (Tanner-Smith & Tipton, 2014), so analyzing dependent *ES*s as if they were independent may lead to incorrect conclusions.

In summary, the three meta-analyses relied on incompatible statistical measures to calculate and compare *ES*s, which makes their results questionable. As Cooper (2010) stated, combining different outcome measures may obscure "important distinctions among the outcomes and might have been misleading" (p. 150). Furthermore, Higgins and Green (2008) added that the use of a meta-analysis may not be meaningful if the treatments are so different that an effect estimate cannot be interpreted in any specific context.

This study, then, has three main goals. First, it will provide a review of the literature regarding the efficacy of cognitive interventions to improve cognition in AD. Special attention will be given to the verbal memory domain as memory impairment is the initial and most prominent cognitive deficit in AD, and the amnestic-type is the most common presentation of AD dementia symptoms (McKhann et al., 2011). Second, our study will conduct a meta-analysis of these interventions with adequate statistical methods to reduce the impact of bias. Finally, and as the main goal of this study, we hope to provide a different approach to meta-analysis by analyzing *ES*s according to specific characteristics of each study design that might potentially affect results in research studies (Higgins & Green, 2008): *randomization, independence of evaluators*, and *control condition*. Even if outcomes from primary studies are scaled in the same metric (e.g., raw score metric) in a meta-analysis, different *ES*s would be expected between random and nonrandom studies, between studies with or without blind assessors or between studies with different control groups (e.g., cognitive plus pharma vs. placebo). For example, higher *ES*s are expected when a cognitive intervention is compared to a drug-only group because "comparing two active interventions is likely to reduce effect size" (Olazarán et al., 2010, p. 172) and because drug therapy has limited effects on cognition (Birks & Harvey, 2006; Raina et al., 2008).

Based on previous reviews, we expected to find low to medium *ES*s on cognitive and functional outcomes and hypothesized that studies without control for potential sources of bias would provide higher *ES*s compared to controlled studies.

Method and Materials

Procedure

Electronic databases such as PUBMED, PsycINFO, Google Scholar, and ScienceDirect were explored from February to April 2014 and again in March 2015 using a combination of the search terms *verbal memory*, *non-pharmacological*, and *cognitive* with the terms *rehabilitation*, *stimulation*, *training*, *treatment*, and *intervention*, with *Alzheimer's disease* in the title and abstract. There was no limitation on date of publication. A global search showed 34,845 results. Forty additional papers were found through a references-specific search. Figure 1 provides a flow diagram of the procedure based on the PRISMA Statement (Moher, Liberati, Tetzlaff, Altman, & the PRISMA Group, 2009).

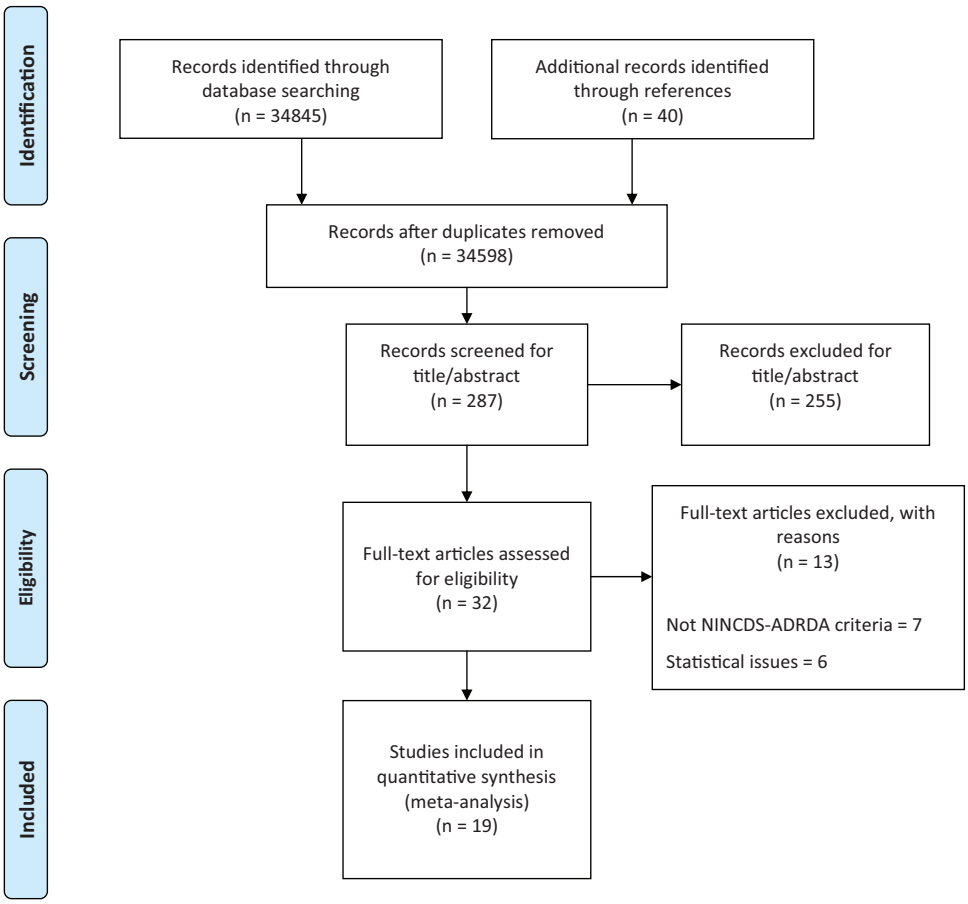

Figure 1. Flow diagram. See the online article for the color version of this figure.

Inclusion Criteria

To be included in the meta-analysis, studies had to (a) have a population with AD in the experimental group; (b) have at least one outcome, either primary or secondary, related to verbal memory and/or cognition; (c) be cohort studies where the control group receives the same treatment, an alternate treatment, or no treatment at all compared to the experimental group; (d) be published in English, because excluding studies published in languages other than English has little effect on treatment effect estimates (Jüni, Holenstein, Sterne, Bartlett, & Egger, 2002; Morrison et al., 2012); (e) include participants meeting a probable or possible diagnosis of AD according to the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria (McKhann et al., 1984; McKhann et al., 2011); and (f) report pre- and posttest means and standard deviations for both treatment and control groups. Each paper was evaluated for inclusion independently by the first two authors using the criteria stated above. In case of a discrepancy, the third author reviewed the paper and decided for inclusion. This happened only for the study by Cipriani, Bianchetti, and Trabucchi (2006), which included only participants with mild cognitive impairment (MCI) in the control group. It was finally included because the experimental group was comprised of people with AD.

Methodological Quality Assessment

An evaluation checklist designed to include the main points regarding quality of research methodology from the document "Strengthening the Reporting of Observational Epidemiological Studies: STROBE Statement" (Vandenbroucke et al., 2007; see Appendix A), along with the specific outcomes for each study, were used to perform separate meta-analyses with a reliable and standardized research methodology. As in the case of the criteria for inclusion, the methodological quality of each paper was evaluated independently by the first two authors, and discrepancy was resolved by the third author.

Correlational Meta-Analysis

Statistical Analysis

Average *r ES* calculation. Because the formula for the *ES* variance includes the pre/post correlation, we contacted the authors and requested the pre/post correlation on each outcome from the primary studies. We meta-analyzed the available correlations using the Hedges and colleagues' random effects methodology described by Field (2001) (see Appendix B) separately for the experimental and control groups. We could only meta-analyze correlations for the outcomes from Clare et al. (2010; $k = 2$), Viola et al. (2011;

$k = 2$), Bergamaschi et al. (2013; $k = 4$), and Cotelli et al. (2014; $k = 13$). Results showed a mean *ES* $r = .70$ (95% CI [.62, .77], weighted variance = .013) for the experimental groups and a mean *ES* $r = .76$ (95% CI [.69, .83], weighted variance = .011) for the control groups, with a high and significant heterogeneity among *ES*s in both the experimental (Hedges' $Q = 10.09$, $df = 3$, $p = .018$, $I^2 = 70.3\%$) and control groups (Hedges' $Q = 10.98$, $df = 3$, $p < .012$, $I^2 = 72.7\%$).

Sensitivity analysis. Based on the findings of the correlational meta-analysis, we then performed a sensitivity analysis using correlational values in ranges of .10 to investigate their impact on the average *ES*s. As shown in Appendix C, the average *ES* remained constant across different correlational values, with a slightly difference of .03. The average *ES* using the real correlations compared to the average *ES* using the average *r* changed from 0.29 to 0.27. Based on these results, we calculated the variance of *ES*s from both the experimental and control groups using the average correlational *ES* for each group, respectively.

Statistical Analysis

Effect size calculation. *ES*s were calculated with the formula for independent groups pre/posttest (IGPP) based on raw scores (Morris, 2008; Morris & DeShon, 2002), which includes the standard deviation of the pretest as an unbiased estimate of the variance. *ES*s for each independent group were corrected using the bias function formula and weighted by the reciprocal of its sampling variance (Morris & DeShon, 2002), which was calculated with the average *r* separately for the experimental and control groups. All these statistics were calculated using Microsoft Excel 2007. Although there are different options for calculating *ES*s in pre/post designs, the IGPP design was selected for three reasons: (a) it provides separate *ES*s for experimental and control groups, so it is possible to observe *ES*s in groups with different control interventions; (b) it includes the pre/post correlations of each separate group for the variance calculations, not the pooled pre/post correlation across groups; and (c) the observed variance of the IGPP design does not differ from the theoretical variance more than 3% under most conditions (Morris, 2008).

To compare our results with the three previous meta-analyses, we calculated the average *ES* for the following:

- The Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975)
- Memory outcomes: word-list tasks (Rey Auditory Verbal Learning Test, Rey, 1964; Hopkins Verbal Learning Test, Benedict, Schretlen, Groninger, & Brandt, 1998; selective reminding test, Kessler, Denzler, & Moskowitsch, 1988), a story recall task (brief story recall, Novelli et al., 1986; logical memory test, Wechsler, 1987; prose memory, Capitani, Della Sala, Laiacona, Marchetti, & Spinnler, 1994), and a face–name association task and a memory battery (Rivermead Behavioral Memory Test, Wilson, Baddeley, & Cockburn, 2003). Word-list tasks and story recall tasks were analyzed separately where possible.
- Basic ADLs (Katz Index for Activities of Daily Living, Katz, Ford, Moskowitz, Jackson, & Jaffe, 1963; Bayer Activities of Daily Living Scale, Hindmarch, Lehfeld, de Jongh, & Erzigkeit, 1998; Functional Living Skills Assessment, Farina et al., 1999).
- The Instrumental ADL Questionnaire (IADL; Lawton & Brody, 1969).
- Functional outcomes: including all the functional outcomes when separate analyses for ADL and IADL could not be computed (IADL, Katz Index for Activities of Daily Living, Bayer Activities of Daily Living Scale, Functional Living Skills Assessment, Advance Activities of Daily Living Questionnaire, Reuben, Laliberte, Hiris, & Mor, 1990; and the Rapid Disability Rating Scale–2, Linn & Linn, 1982). No reference to the functional scales were obtained from Giordano et al. (2010) and Cotelli et al. (2014).

Source of bias. Regarding randomization, all but Farina et al. (2002), Hofmann et al. (2003), Cipriani et al. (2006), Farina et al. (2006), Giordano et al. (2010), and Viola et al. (2011) randomly assigned participants to groups.

Regarding independence of evaluators, all but Heiss et al. (1993; Heiss et al., 1994); Hofmann et al. (2003); Cipriani et al. (2006); Requena, Maestú, Campo, Fernández, and Ortiz (2006); Tárraga et al. (2006); and Giordano et al. (2010) reported that cognitive assessors were blind to group allocation.

For the control condition bias, three comparison groups were differentiated: (a) a structured cognitive intervention, (b) a placebo control condition, and (c) a pharma control condition with or without cognitive intervention or no treatment at all. Farina et al. (2002); Hofmann et al. (2003); Loewenstein, Acevedo, Czaja, and Duara (2004); Cipriani et al. (2006); Requena et al. (2006); Tárraga et al. (2006); Jelcic et al. (2012); Bergamaschi et al. (2013); and Cotelli et al. (2014) included a control intervention with exercises tapping cognitive abilities. The studies with a placebo control condition included interventions such as social support (Heiss et al., 1993; Heiss et al., 1994; Davis, Massman, & Doody, 2001), educational information (Cahn-Weiner, Malloy, Rebok, & Ott, 2003), recreational activities (Farina et al., 2006; Bergamaschi et al., 2013), semistructured intervention on current affairs and relevant events (Galante, Venturini, & Fiaccadori, 2007), muscle relaxation (Clare et al., 2010), creative work (Jelcic et al., 2012), or motor exercises (Cotelli et al., 2014). Requena et al. (2006), Tárraga et al. (2006), Clare et al. (2010), Giordano et al. (2010), Viola et al. (2011), and Fernández-Calvo et al. (2015) included a control group receiving only pharmacotherapy or no therapy at all.

Statistical Analysis

As in correlational meta-analysis, homogeneity of *ES*s was calculated using Hedges' Q (Morris, 2008) and the I^2 statistic (Higgins & Thompson, 2002). In Hedges' Q, the null hypothesis is that all the studies have a common *ES*. The I^2 provides a measure of the total variation in *ES*s that is due to the heterogeneity between studies. Percentages for low, moderate, and high heterogeneity are 20%, 50%, and 70% respectively. Whenever I^2 yielded a negative result, it was expressed as 0% (Higgins, Thompson, Deeks, & Altman, 2003).

As the same participants provided different *ES*s within each study, the robust variance estimator (RVE) for correlated *ES*s was applied (Hedges, Tipton, & Johnson, 2010; Tanner-Smith & Tipton, 2014). The RVE provides an *ES* after the statistical dependency among *ES*s is accounted for. It provides both an unconditional *ES* and a multivariable metaregression model to test the

influence of moderators on *ES*s. We run the RVE for correlated effects using the SPSS macro developed by Tanner-Smith and Tipton (2014). This estimation provides a beta coefficient that is equal to the unconditional mean *ES* across studies, and tests the hypothesis of $b \neq 0$ with a t test with "number of studies – parameters" degrees of freedom (in this case, $k - 1$). To calculate the RVE, the assumed correlation between all pairs of *ES* within studies must be specified. Tanner-Smith and Tipton (2014) suggest to use an average correlation found in one or more studies, although the cost of choosing an inappropriate value is negligible because it does not affect the precision of the confidence intervals, t test or statistical inferences. We, therefore, used the average correlation within the experimental groups ($r = .70$). The same macro was used to test the effect of moderators. As different designs and different populations were used in each study, the random effects model was applied (Borenstein, Hedges, Higgins, & Rothstein, 2009), which assumes that each *ES* is a sample of a distribution of possible *ES* and includes the variance within and the variance between studies in its calculation.

We calculated the power of each unconditional *ES* using the formula described by Hedges and Pigott (2001) for the hypothesis $ES = 0$. This formula uses the z statistic and the normal cumulative distribution. For a two-tailed test, Power $= 1 - \Phi(z_{\alpha/2} - Z^*) + \Phi(-z_{a/2} - Z^*)$. The symbol Φ represents the standard normal cumulative distribution function and Z^* is calculated with the formula $ES/\sqrt{Var^*}$ with the random effects variance in the denominator. The results of the unconditional *ES*s will be presented except when the robust variance estimator showed that a significantly different from zero *ES* turned out to be nonsignificant or vice versa. Power of tests from the RVE could not be computed because, to date, there has not been any work on computing power with RVE (Elizabeth Tipton, personal communication, October 10, 2015). Power was computed for 23 out of 35 tests.

To test whether the moderators were highly collinear with one another, we created a binary variable for each study (0 = no, 1 = yes) within each of the moderator variables and produced a correlation matrix of the moderator variables at study level ($n = 19$) so as to give an overview of the intercorrelation between moderators (e.g., correlation between random and nonrandom studies is −1; if all the random studies had a placebo control group, then the correlation between these two moderators would be 1). Results are presented in Appendix D.

The significance level was set at $\alpha < .01$ for two reasons: first, because of the high number of tests for statistical significance; second, because of the small number of studies included for calculating *ES*s and for estimating the metaregression coefficient. Tanner-Smith and Tipton (2014) suggest using an alpha level $p < .01$ or $p < .001$ when using the robust variance estimation with 10–40 studies.

Results

Characteristics and Quality of Studies

Nineteen studies with 812 participants ($N_{\text{treatment}} = 356$; $N_{\text{Control}} = 456$) were meta-analyzed. For a detailed description of each study, its participants, duration and the type of measures used as well as the outcomes obtained, see Table 1.

Table 2 provides a summary of the characteristics of each study defined as potential biases. The median number of biases was 7 ($M = 6.16 \pm 2.22$), with a range from 2 to 9. The most frequent potential bias was the absence of a double-blind design, followed by the lack of exclusion criteria and the absence of a detailed explanation of the interventions. Nine studies compared a treatment intervention with a cognitive stimulation control condition, 10 studies compared a treatment intervention with a placebo control condition, and six studies compared a treatment intervention with cognitive plus pharma, pharma-only control condition, or no treatment at all.

Effect Size Comparison

Eighty-seven measures were identified in the 19 studies (see Figure 2). Sixty-seven (77%) were cognitive measures and 20 (23%) were functional measures. The average *ES* for all research outcomes together was small and significantly different from zero (*ES* = 0.30, var = .00, 99% CI [0.17, 0.44], $z = 5.77$, $p = .000$) with high heterogeneity between *ES*s across studies ($I^2 = 76\%$). This average *ES* was taken as a general measure of treatment effects to give an overview of the variability of *ES*s. Separate *ES*s for experimental and control groups are shown in Appendix E.

The MMSE (Table 3) showed a small *ES* that was not significantly different from zero ($k = 15$; *ES* = 0.41, *SE* = .16, 99% CI [−0.06, 0.88], $t = 2.58$, $p = .022$), whereas the Alzheimer's Disease Assessment Scale–Cognitive (ADAS-Cog) subscale showed a high and nonsignificant *ES* ($k = 4$; *ES* = 0.77, *SE* = .15, 99% CI [−0.09, 1.63], $t = 5.21$, $p = .014$). The memory domain (including all the memory outcomes) showed a small but nondifferent from zero *ES*s ($k = 14$; *ES* = 0.23, *SE* = .08, 99% CI [−0.02, 0.47], $t = 2.81$, $p = .015$). When analyzed separately, the story recall tasks showed a small and different from zero *ES* ($k = 6$; *ES* = 0.29, var = .01, 99% CI [0.00, 0.58], $z = 2.62$, $p = .009$), whereas *ES* from the list learning tasks *ES* ($k = 5$; *ES* = 0.18, var = .01, 99% CI [−0.11, 0.48], $z = 1.62$, $p = .105$) and the Rivermead Behavioral Memory Test ($k = 6$; *ES* = 0.10, var = .01, 99% CI [−0.16, 0.36], $z = .99$, $p = .321$) were not different from zero.

Regarding functional outcomes (Table 3), *ES* was negligible and not significantly different from zero ($k = 11$; *ES* = 0.15, var = .01, 99% CI [−0.05, 0.36], $z = 1.94$, $p = .053$). When calculated separately, neither ADL ($k = 8$; *ES* = 0.16, var = .01, 99% CI [−0.11, 0.43], $z = 1.49$, $p = .136$) nor IADL ($k = 7$; *ES* = 0.09, var = .01, 99% CI [−0.17, 0.35], $z = .90$, $p < .370$) *ES*s were different from zero.

Moderator Analyses

Tables 4 and 5 show *ES*s when studies were combined based on control for randomization, independence of evaluators, and control intervention respectively.

Randomization. The MMSE *ES* was small-to-medium and not different from zero in both randomized ($k = 9$; *ES* = 0.59, *SE* = .26, 99% CI [−0.27, 1.46], $t = 2.31$, $p = .050$) and nonrandomized studies ($k = 6$; *ES* = 0.25, var = .02, 99% CI [−0.10, 0.59], $z = 1.86$, $p = .063$). *ES* from memory outcomes was small and different from zero in randomized studies ($k = 11$; *ES* = 0.29, var = .01, 99% CI [0.09, 0.49], $z = 3.72$, $p < .001$)

Table 1
Studies Included in the Meta-Analysis

k	Study	*n* EG	*n* CG	Medication	Treatment intervention	Control intervention	Duration	Follow-up	Primary outcome	Secondary outcome	Findings
1	Heiss et al. (1993)	10 AD	10 AD	Pyritinol (P)	Cognitive training (CT)	Placebo; social support	6 months	—	MMSE	Verbal selective reminding	A tendency for improvement in CT + PS group
		10 AD		Phosphatydilserine (PS)	CT + P						
		10 AD			CT + PS						
2	Heiss et al. (1994)	18 AD	17 AD	Pyritinol (P)	CT	Placebo; social support	1 hr twice a week for 24 weeks	—	MMSE	Verbal selective reminding	A tendency for improvement in CT + PS group
		17 AD		Phosphatydilserine (PS)	CT + P						
		18 AD			CT + PS						
3	Davis et al. (2001)	19 AD	18 AD	Donepezil	5 individual sessions	Placebo; unstructured conversation	1 session per week for 5 weeks	—	Face–name recall task	MMSE	Significant improvements in face–name recall across sessions
				Tacrine	Face–name training				(no raw data provided)	LM (WMS-III)	No differences in dementia severity or verbal memory
				Fluoxetine Sertraline	Spaced retrieval						
4	Farina et al. (2002)	10 AD	10 AD	Not specified	30 sessions Procedural memory training	Cognitive stimulation; stimulation of spared cognitive functions	6 sessions per week for 5 weeks	3 months	RBMT-Screening score	MMSE	FLSA improvements in both groups; no change in RBMT, ADL, nor IADL
									RBMT-Profile score	ADL	
										IADL RMBPC FLSA	
5	Cahn-Weiner et al. (2003)	17 AD	17 AD	Cholinesterase inhibitors	6 individual sessions	Placebo; educational information	1 sessions per week for 6 weeks	8 weeks	HVLT-R	ADL	No main effects of group or time, nor interaction
										EMQ	

(*table continues*)

Table 1 (*continued*)

k	Study	*n* EG	*n* CG	Medication	Treatment intervention	Control intervention	Duration	Follow-up	Primary outcome	Secondary outcome	Findings
6	Hofmann et al. (2003)	9 AD	10 HC	None	12 sessions of interactive computer-based training	Cognitive stimulation	3 times per week for 4 weeks	3 weeks	MMSE	—	No significant Group × Time interactions
7	Loewenstein et al. (2004)	25 AD	19 AD	Donepezil	24 individual sessions	Mental stimulation	2 sessions per week for 12–16 weeks	3 months	Face–name association (FNA)	Logical Memory (LM; WMS-III)	Increase in both short- and long-term in TG in FNA; no differences in LM
				Rivastigmine Galantamine						RMBPC Bayer ADLS	
8	Cipriani et al. (2006)	10 AD	10 MCI	Donepezil	32 individual sessions	Cognitive stimulation	4 sessions per week for 8 weeks	—	Neuropsychological training software (NPT)	MMSE	Improvements in NPT and MMSE after treatment in AD group
				Rivastigmine						RBMT Screening score AADL	No differences in RBMT
9	Farina et al. (2006)	16 AD	16 AD	Cholinesterase inhibitors	15 group sessions of "cognitive-specific" activities	Placebo; recreational activities	3 sessions per week for 4 weeks, 2 sessions per week for 1 week, and 1 session per week for 1 week	6 months	MMSE	RBMT-P	No differences in the measures used except for the FLSA
				Antidepressants Neuroleptics						RBMT-S FLSA ADL IADL RMBPC	
10	Requena et al. (2006)	20 AD	10 AD	Donepezil	208 group sessions	Cognitive stimulation	4 sessions per week for 1 year	—	MMSE	ADAS-Cog	No differences between pharma + cognitive and cognitive alone in MMSE, but pharma + cognitive improved in ADAS-Cog

(*table continues*)

Table 1 (*continued*)

k	Study	*n* EG	*n* CG	Medication	Treatment intervention	Control intervention	Duration	Follow-up	Primary outcome	Secondary outcome	Findings
			18 AD			Pharma alone					
			18 AD								
11	Tárraga et al. (2006)	15 AD	16 AD	ChEIs	Interactive multimedia Internet-based	240 IPP + AChEI	IMIS: 3 sessions per week for 24 weeks	—	ADAS-Cog	MMSE	Both IMIS and IPP improved in ADAS-Cog
			12 AD		System (IMIS) + Integrated Psychostimulation Program (IPP)	Pharma only	IPP: 2 daily sessions, 5 days per week for 24 weeks			RBMT-Story immediate recall	Differences in MMSE between groups
					72 individual IMIS sessions + 120 IPP					Kurz test	No differences in Kurz test
12	Galante et al. (2007)	7 AD	4 AD	AChEI	12 individual 60-min sessions of cognitive training with cognitive software	Placebo; semistructured interview on current and past events	3 times per week for 4 weeks	3 months and 9 months	MMSE	Prose memory	No significant Group × Time interaction
										IADL BADL	
13	Clare et al. (2010)	20 AD	23 AD	Donepezil	8 individual sessions. Goal oriented	Placebo; muscle relaxation and breathing exercises	1 session per week for 8 weeks	6 months	The Canadian Occupational Performance Measure (COPM)	Rivermead Behavioral Memory Test, 2nd Ed.	COPM increases in TG; no differences in RBMT-II
			19 AD	Reminyl Rivastigmine		Pharma only					
14	Giordano et al. (2010)	62 AD	38 AD	Donepezil	Reality orientation therapy (ROT)	Pharma only	6 sessions per week for 3 weeks	2 months (only MMSE)	MMSE	ADL	Significant improvements in MMSE and ADAS-Cog in TG

(*table continues*)

Table 1 (*continued*)

k	Study	*n* EG	*n* CG	Medication	Treatment intervention	Control intervention	Duration	Follow-up	Primary outcome	Secondary outcome	Findings
					15 group (5–6) sessions + 3 individual sessions				ADAS-Cog	IADL	No differences in ADL or IADL
										PPT	PPT higher in TG after intervention
15	Viola et al. (2011)	25 AD	16 AD	AChI or memantine	24 group sessions of cognitive training, computer-based intervention or occupational therapy	Pharma only	2 sessions per week for 12 weeks	—	MMSE	Short Cognitive Test (SKT)	Significant worsening on the SKT only in the CG
16	Jelcic et al. (2012)	20 AD	20 AD	None	Lexical-semantic rehabilitation	Placebo; unstructured cognitive stimulation	2 sessions per week for 12 weeks	6 months	MMSE	IADL	Significant improvement in MMSE, short-, and long-term story recall and long-term RAVLT
					24 group (4) sessions				Brief story recall RAVLT		No differences within/between in IADL
17	Bergamaschi et al. (2013)	16 AD	16 AD	Donepezil	100 group sessions	Placebo; nonspecific cognitive activity	5 sessions per week for 5 months	—	MMSE	Story recall–immediate	Significant Group × Time interaction in the MMSE, MODA, ADL, and IADL, but not on story recall
									Milan Overall Dementia Assessment (MODA)	ADL IADL	
18	Cotelli et al. (2014)	12 AD	12 AD	Donepezil	20 individualized sessions of anodal tDCS plus computerized memory training	20 individualized tDCS + computerized memory training	5 sessions per week for 2 weeks	3 months	FNA (no statistics reported)	MMSE	No effects of memory training

(*table continues*)

Table 1 (*continued*)

k	Study	*n* EG	*n* CG	Medication	Treatment intervention	Control intervention	Duration	Follow-up	Primary outcome	Secondary outcome	Findings
			12 AD	Rivastigmine		Placebo; 20 anodal tDCS + motor training		6 months		ADL	
										IADL	
										RBMT Story recall–Imm	
										RBMT Story recall–Del.	
										RAVLT Imm	
										RAVLT Del	
19	Fernández-Calvo et al. (2015)	25 AD	30 AD	Donepezil	48 individualized multicomponent cognitive stimulation program	Pharma only; waiting list	3 sessions per week for 16 weeks	—	MMSE	RDRS-2	Less cognitive and functional decline in the EG
				Rivastigmine					ADAS-Cog		

Note. EG = experimental group; CG = control group; RBMT = Rivermead Behavioral Memory Test; MMSE = Mini-Mental State Examination; ADL = Activities of Daily Living; IADL = Instrumental Activities of Daily Living; AADL = Advance Activities of Daily Living; RDRS-2 = Rapid Disability Rating Scale-2; LM = Logical Memory subtest; PPT = Physical Performance Test; RAVLT = Rey Auditory Verbal Learning Test; ADAS-Cog = Alzheimer's Disease Assessment Scale–Cognitive subscale; RMBPC = Revised Memory and Behavioral Problems Checklist; EMQ = Everyday Memory Questionnaire; tDCS = transcranial direct current stimulation.

but not in nonrandomized studies ($k = 3$; $ES = -0.02$, var = .02, 99% CI [−0.39, 0.34], $z = -0.16$, $p = .869$). Functional *ES* were not different from zero in randomized ($k = 7$; $ES = 0.22$, var = .01, 99% CI [−0.07, 0.50], $z = 1.94$, $p = .052$) or nonrandomized studies ($k = 4$; $ES = 0.07$, var = .01, 99% CI [−0.22, 0.25], $z = .62$, $p = .533$). Moderator analyses showed no significant differences between randomized and nonrandomized studies in memory *ES* (all $p > .01$).

Independence of evaluators. The MMSE *ES* was small and not different from zero in studies with ($k = 8$; $ES = 0.43$, var = .05, 99% CI [−0.02, 0.88], $z = 1.87$, $p = .062$) or studies without independent evaluators ($k = 7$; $ES = 0.33$, $SE = .18$, 99% CI [−0.32, 0.99], $t = 1.88$, $p = .109$). *ES* from memory outcomes was small and not different from zero in studies with ($k = 10$; $ES = 0.26$, $SE = .10$, 99% CI [−0.02, 0.54], $t = 2.53$, $p = .032$) and in studies without independent evaluators ($k = 4$; $ES = 0.17$, var = .02, 99% CI [−0.15, 0.49], $z = 1.37$, $p = .169$). Functional *ES* was not different from zero in studies with ($k = 9$; $ES = 0.19$, var = .01, 99% CI [−0.04, 0.43], $z = 2.13$, $p = .033$) or without independent evaluators ($k = 2$; $ES = 0.00$, var = .03, 99% CI [−0.42, 0.42], $z = .01$, $p = .993$). Moderator analyses showed no significant differences between studies with or without independent evaluators in any of the variables (all $p > .01$).

Comparison conditions. The MMSE *ES* was not different from zero when the control group received cognitive stimulation ($k = 8$; $ES = 0.09$, var = .03, 99% CI [−0.38, 0.55], $z = .48$, $p = .628$), a placebo intervention ($k = 8$; $ES = 0.17$, var = .07, 99% CI [−0.49, 0.83], $z = .67$, $p = .505$), or only pharma ($k = 4$; $ES = 0.69$, $SE = .25$, 99% CI [−0.78, 2.16], $t = 2.74$, $p = .071$). *ES* from memory outcomes was small and different from zero in studies with a placebo control condition ($k = 10$; $ES = 0.22$, var = .01, 99% CI [0.02, 0.42], $z = 2.79$, $p = .005$), but not in studies with a cognitive control condition ($k = 7$; $ES = 0.37$, var = .04, 99% CI [−0.15, 0.89], $z = 1.85$, $p = .063$). It was not possible to calculate *ES* from memory outcomes in studies with pharma-only control condition due to a lack of outcomes (one outcome in one study). *ES* from functional outcomes was not different from zero in studies with a cognitive control condition ($k = 4$; $ES = -0.04$, var = .01, 99% CI [−0.35, 0.28], $z = -.29$, $p = .766$), a placebo control condition ($k = 6$; $ES = 0.24$, var = .01, 99% CI [−0.05, 0.53], $z = 2.16$, $p = .030$), or pharma-only control condition ($k = 2$; $ES = 0.19$, var = .05, 99% CI [−0.38, 0.77], $z = .87$, $p = .382$). Metaregression analyses showed no effect of the control condition in any of the variables analyzed (all $p > .01$).

Conclusions

This meta-analysis intended to extend the results provided by previous meta-analyses regarding the effectiveness of cognitive interventions for AD by analyzing *ES*s with a different methodology: calculating *ES*s in a common metric, calculating the average *ES* using metaregression for correlated *ES*s and analyzing potential sources of biases as moderator variables.

When the *ES*s of all the selected studies were combined to get an average *ES*, the results showed low efficacy and high heterogeneity on *ES*s. The effects of interventions on cognitive and functional outcomes were negligible to low. Mean *ES*s for measures of general cognitive functioning were higher than those reported by Bahar-Fuchs et al. (2013; $ES = 0.10$) and similar to

Table 2
Characteristics of the Studies Included in the Meta-Analysis

Source of bias	Study reported in Table 1																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	%
Groups	4	4	3	2	4	3	2	2	3	3	2	2	3	2	2	2	2	2	2	
Randomization	Y	Y	Y	N	Y	N	Y	N	N	Y	Y	Y	Y	N	N	Y	Y	Y	Y	31.58
Double blind	N	N	Y	N	Y	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	89.47
Independence of assessors	N	N	Y	Y	Y	N	Y	N	Y	N	N	Y	Y	N	Y	Y	Y	Y	Y	36.84
Control intervention	Y	Y	Y	Y	N	Y	N	Y	N	Y	Y	N	N	N	N	Y	Y	Y	N	42.10
Placebo intervention	Y	Y	N	N	Y	N	Y	N	Y	N	N	Y	Y	N	N	N	N	Y	N	57.89
Inclusion criteria	N	N	N	Y	N	N	N	Y	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	42.10
Exclusion criteria	N	N	N	Y	N	N	N	N	Y	Y	N	N	Y	N	N	Y	N	N	N	73.68
Control of selection bias	N	Y	N	Y	Y	N	Y	N	N	N	N	N	Y	Y	Y	Y	Y	Y	Y	42.10
Follow-up assessment	N	N	N	Y	Y	Y	Y	N	Y	Y	N	Y	Y	Y	N	N	Y	Y	N	42.10
Standardized assessment	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	0
Functional assessment	N	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y	Y	N	Y	Y	Y	N	47.37
Detailed explanation of intervention	N	N	Y	N	Y	Y	N	N	N	N	N	N	N	N	N	Y	Y	Y	Y	63.16
Attrition causes	N	Y	N	Y	Y	N	N	N	Y	Y	Y	Y	Y	N	Y	N	N	N	N	52.63
Σ bias	9	7	7	4	3	9	7	9	5	7	8	5	3	8	8	4	5	2	7	

Note. Y = yes; N = no (bias).

those reported by Olazarán et al. (2010; *ES* = 0.30–0.59) and Sitzer et al. (2006; *ES* = 0.37–0.40), although not significantly different from zero.

Several factors could explain this difference in *ES*s. First, Bahar-Fuchs et al. (2013) included only RCT, whereas our meta-analysis includes studies with different designs, which increases heterogeneity. Second, Bahar-Fuchs et al. reported *ES*s for mild to moderate AD and vascular dementia, whereas we analyzed studies of participants with AD only, except for one study where the participants had mixed vascular + AD dementia (Clare et al., 2010) and one with participants with AD and MCI (Loewenstein et al., 2004).

Regarding the ADLs, we found a higher *ES* than the one reported by Bahar-Fuchs et al. (2013; *ES* = 0.00) and smaller than the one reported by Olazarán et al. (2010; *ES* = 0.37–0.41) and Sitzer et al. (2006; *ES* = 0.32–0.75). As heterogeneity among functional *ES*s was low to moderate (I^2 = 0.00–0.50), and mean *ES*s for functional outcomes were consistent across methodological designs (*ES* = −0.04–0.29), our results suggest that the effect of cognitive stimulation on functional outcomes for AD is low, in line with other findings regarding the absence of transfer to untrained situations. As dementia is diagnosed when cognitive impairment interferes with ADLs, our results highlight the need to develop cognitive interventions that provide persons with AD with strategies to cope with ADLs and further suggest that the kind of cognitive interventions published to date may be ineffective to improve memory, cognition, and functionality.

Several factors could explain our results. First, measures of general cognitive outcome are too broad and do not focus on the outcome of interest (e.g., the MMSE for assessing verbal memory). A measure of general cognitive functioning, thus, would have limited usefulness as it is likely unrelated—or only tangentially related—to what has been trained during the intervention. For example, Loewenstein et al. (2004) and Davis et al. (2001) trained their participants in a face–name association task, whereas Giordano et al. (2010) trained their participants using reality orientation therapy. It is unlikely that a measure of general cognitive functioning can reliably identify any improvement in memory regarding face–name association, even though the participant identifies correctly all the items included in the intervention task. In the case where reality orientation therapy is effective and the participants improve their time orientation, it is likely that some improvements in measures such as the MMSE are evident, but its functional correlate with ADLs is uncertain. Thus, we suggest that the efficacy of interventions be measured through the achievement of personalized goals rather than through the use of general tests such as the MMSE or the ADAS-Cog subscale to truly impact the needs of people with AD dementia. Some examples of this type of intervention were carried out by Clare et al. (2010) and Zanetti et al. (2001). Clare et al. (2010) performed an intervention focused on relevant goals to the individual and found that participants who received personalized intervention rated higher on both performance and satisfaction after intervention, whereas control groups did not show any significant change. Zanetti et al. (2001) trained their participants in basic and instrumental ADLs and analyzed the impact of the intervention through the decrease in the time needed to perform the task.

A possible explanation for the absence of significant functional improvements is that none of the studies in the meta-analysis included a treatment group trained by an occupational therapist in specific ADLs. As occupational therapists are skilled in improving ADLs through meaningful activity training programs including the preferences, interests, and the strengths of people with dementia (National Collaborating Centre for Mental Health, 2007), any rehabilitation program seeking to improve functionality and independence must include these health workers within the professional staff. The only work that included occupational therapists was the one by Fernández-Calvo et al. (2015), but their specific intervention on ADLs was not detailed. If the rationale for a memory intervention is that any improvement in cognition will transfer to untrained situations such as ADLs, then probably nonsignificant results are likely to be reported because of a lack of transfer to untrained situations (McDaniel & Bugg, 2012).

Another explanation could be the limited sensitivity of functional scales and questionnaires in general to small although significant changes for individual participants. The use of group

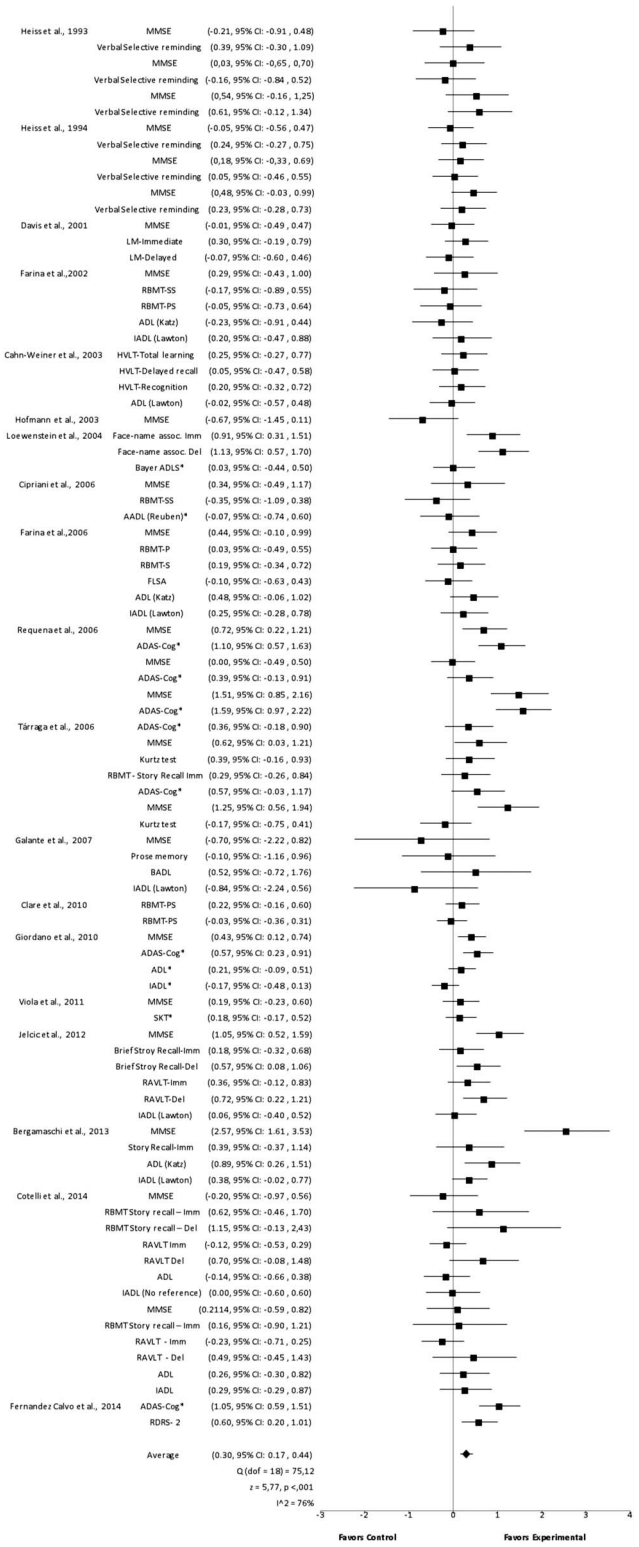

Figure 2. Forest plot with all outcomes and studies.

means instead of single-subject data could mask clinically significant changes that cannot be observed in standardized tests and questionnaires. The use of data from a group where no improvements are found will suggest that the intervention is not effective (which is the case in this work); however, analyzing the data from individuals within a group may provide insights about the individual's characteristics related to the efficacy of such intervention.

Given the importance of functional ADLs, we consider that future works should compare a cognitive intervention only to an intervention combining cognitive + functional activities focusing on training participants in ADLs (e.g., preparing meals) and significant relevant goals (e.g., finding personal belongings, dressing), and measure their improvement using outcomes such as the time required to perform each task, the number of errors on each task, the number and type of external help needed, or the satisfaction derived from performing those tasks (Clare, Evans, Parkinson, Woods, & Linden 2011; Clare et al., 2010; Zanetti et al., 2001). It is important to note that we do not suggest that interventions must be carried out according to the predefined outcome measure. Rather, outcomes specifically related to trained situations should be selected to measure the efficacy of any given intervention before even initiating it. Our hypothesis is that *ES*s will become medium to large for variables related to trained situations and will remain low for measures such as the MMSE or the IADL questionnaire.

Differences in *ES* Due to Methodological Bias

Contrary to our expectations, we could not find differences in *ES*s for cognitive or functional outcomes when participants are randomly assigned to groups, participants are assessed by personnel blinded to group allocation or when a treatment intervention is compared to another cognitive intervention, a placebo intervention or an alternate intervention (cognitive plus pharma, pharma-only, or no treatment at all). Olazarán et al. (2010) found that the effects of multicomponent interventions were similar to the effects obtained by drug therapy and suggested that these interventions should be complementary. Nevertheless, one would expect a higher *ES* when the cognitive intervention is compared to a drug-only group (Birks & Harvey, 2006; Olazarán et al., 2010; Raina et al., 2008). According to these findings, there would be no reason for recommending cognitive interventions or other kind of interventions such as social support or mental stimulation over pharmacotherapy, as *ES*s in the latter case were not different from zero. However, it cannot be concluded that there are no beneficial effects of cognitive intervention over pharmacotherapy alone or in combination with cognitive stimulation because the small number of studies and outcomes reduced the power to find significant effects.

Contrary to our expectations, studies that did not control for bias did not show higher *ES*s. Overall, average *ES*s with and without control for bias showed that the effects of the interventions on measures of general cognitive functioning were similar to those on functional outcomes, which supports the conclusions of others who question the efficacy of cognitive training as it has been performed in previous research (Bahar-Fuchs et al., 2013; McDaniel & Bugg, 2012). These results are in line with other meta-analyses assessing cognitive interventions in other populations with neurologic diseases such as multiple sclerosis (Magalhães et al., 2014) or brain injury (Rohling, Faust, Beverly, & Demakis, 2009), which found no significant effects of interventions on memory.

Table 3
Effect Size Comparisons According to Cognitive and Functional Outcome Criteria

	Cognitive						Functional		
Statistics	MMSE	ADAS-Cog	WL tasks	SR tasks	RBMT	Memory	Functional	ADL	IADL
Study as reported in Table 1	1–4, 6, 8–12, 14–18	10, 11, 14, 19	1, 2, 5, 16, 18	3, 11, 12, 16–18	4, 8, 9, 11, 13, 18	1–5, 7–9, 11–13, 16–18	4, 5, 7–9, 12, 14, 16–19	4, 5, 7, 9, 12, 14, 17, 18	4, 9, 12, 14, 16–18
Sample size	$N_T = 392$ $N_C = 348$	$N_T = 122$ $N_C = 152$	$N_T = 77$ $N_C = 143$	$N_T = 89$ $N_C = 98$	$N_T = 85$ $N_C = 118$	$N_T = 215$ $N_C = 294$	$N_T = 220$ $N_C = 204$	$N_T = 165$ $N_C = 144$	$N_T = 143$ $N_C = 128$
Number of outcomes	23	7	15	10	11	34	19	10	8
Q (df)	27.56 (14)	5.72 (3)	8.28 (4)	5.39 (5)	6.31 (5)	23.18 (13)	13.90 (10)	8.25 (7)	5.76 (6)
I^2	49%	48%	52%	7%	21%	44%	28%	15%	0%

Note. MMSE = Mini-Mental State Examination; ADAS-Cog = Alzheimer's Disease Assessment Scale–Cognitive subscale; WL = word list task; SR = story recall task; RBMT = Rivermead Behavioral Memory Test; ADL = activities of daily living; IADL = Instrumental ADL. N_T = sample size in the treatment groups; N_C = sample size in the control groups.

Table 4
Effect Size Comparison According to Randomization and Independence of Assessors

	Randomization						Independence of assessors					
	MMSE		Memory		Functional		MMSE		Memory		Functional	
Statistics	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N
Study as reported in Table 1	1–3, 10–12, 16–18	4, 6, 8, 9, 14, 15	1–3, 5, 7, 11, 13, 16–18	4, 8, 9	5, 7, 12, 16–19	4, 8, 9, 14	3, 4, 9, 12, 15–18	1, 2, 6, 8, 10, 11, 14	3–5, 7, 9, 12, 13, 16–18	1, 2, 8, 11	4, 5, 7, 9, 12, 16–19	8, 14
Sample size	$N_T = 122$ $N_C = 220$	$N_T = 132$ $N_C = 100$	$N_T = 62$ $N_C = 78$	$N_T = 36$ $N_C = 36$	$N_T = 105$ $N_C = 130$	$N_T = 98$ $N_C = 74$	$N_T = 125$ $N_C = 124$	$N_T = 144$ $N_C = 196$	$N_T = 162$ $N_C = 186$	$N_T = 53$ $N_C = 108$	$N_T = 148$ $N_C = 156$	$N_T = 72$ $N_C = 48$
Number of outcomes	17	6	29	5	12	8	9	14	26	8	17	3
Q (df)	18.93 (8)	5.50 (5)	19.05 (10)	1.61 (2)	8.75 (6)	4.20 (3)	11.82 (7)	13.89 (6)	17.80 (9)	4.18 (3)	11.64 (8)	1.20 (1)
I^2	58%	9%	48%	0%	31%	29%	41%	57%	49%	28%	31%	16%

Note. MMSE = Mini-Mental State Examination; Y = yes; N = no; N_T = sample size in the treatment groups; N_C = sample size in the control groups.

Table 5
Effect Size Comparison With the Control Group Included

Control group	Study as reported in Table 1	Outcomes	*Q*	*df*	I^2
MMSE					
Cognitive stimulation	4, 6, 8, 10, 11, 18	6	5.11	5	2%
Placebo	1–3, 9, 12, 16–18	10	12.29	7	43%
Pharma	10, 11, 14, 15	5	4.56	3	34%
Functional					
Cognitive stimulation	4, 7, 8, 18	6	1.07	3	0%
Placebo	5, 9, 12, 16–18	11	7.54	5	34%
Pharma	14, 19	3	1.99	1	50%
Memory					
Cognitive stimulation	4, 7, 8, 11, 18	10	7.06	4	43%
Placebo	1–3, 5, 9, 12, 13, 16–18	21	10.16	9	11%

Note. MMSE = Mini-Mental State Examination.

This work has, notwithstanding, several limitations. One of the main caveats is that some relevant work in the area of cognitive interventions on AD was not included because it failed to meet our requirements for diagnostic criteria or pre/post statistics (Gaitán et al., 2013; Kurz et al., 2012; Metitieri et al., 2001; Olazarán et al., 2004; Onder et al., 2005; Simmons-Stern, Budson, & Ally, 2010; Zanetti et al., 2001). All the participants met the Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association's criteria for AD, which are biased toward memory impairments. However, there exist atypical AD presentations such as focal posterior cortical atrophy or progressive aphasia in which episodic memory could either be spared during the initial phases (Galton, Patterson, Xuereb, & Hodges, 2000) or be so pronounced that suggest a critical diagnostic feature (Lambon Ralph, Patterson, Graham, Dawson, & Hodges, 2003). In the absence of a detailed explanation of complementary exams (e.g., magentic resonance images, biomarkers), the different characteristics of persons diagnosed with AD could account for some of the variability in the effects of treatment. We recommend that future research use diagnostic criteria including both biomarkers and neuropsychological testing specifically designed for target populations, thus increasing sensitivity and specificity (Becker, Boller, Lopez, Saxton, & McGonigle, 1994; Lopez, McDade, Riverol, & Becker, 2011) instead of using more general criteria such as those included in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition and Fourth Edition Text Revision* (American Psychiatric Association, 1994, 2000).

Another limitation to this study, which could explain the failure to find any significant effect, is the reduction of power due to the small number of outcomes and studies. For example, when comparing *ES*s for cognitive outcomes between random and nonrandom studies, we could only compare 37 variables from 13 random studies with seven variables from five nonrandom studies. We cannot conclude that there are no significant differences between *ES*s, because we had not enough power to detect such differences (Hedges & Pigott, 2001). The average power of our analyses was 0.27 (median = 0.23), which limits our conclusions. This prevents us from asserting that *ES*s are not different from zero or that there are no differences in *ES*s when studies are compared according to methodological biases; rather, our analyses were underpowered to detect such an affect. However, if data from studies with lower risk of bias (randomization and independent assessors) are taken as closer to the real *ES*, our results indicate that the average *ES* is small for cognitive ($ES = 0.44$) and functional ($ES = 0.22$) outcomes.

Including underpowered studies in meta-analysis is not uncommon. Turner, Bird, and Higgins (2013) found that in 70% of a Cochrane meta-analyses review all studies were underpowered, whereas 66% of the meta-analyses themselves were underpowered. As power in random effects model is affected by both the number of studies and the variance between studies (Cohn & Becker, 2003), which changes as a function of sample size, additional studies with larger sample sizes are needed to more accurately meta-analyze the efficacy of cognitive interventions in AD according to moderator variables and powered studies.

We cannot rule out the possibility that other sources of bias affected the *ES*s. For example, while the majority of studies compared at least two groups of AD participants, Hofmann et al. (2003) compared AD participants with a healthy control group, and Cipriani et al. (2006) included a control group with MCI. Other source of bias could be the modality of the intervention delivered; for example, Hofmann et al. (2003), Cipriani et al. (2006), Tárraga et al. (2006), and Galante et al. (2007) reported delivering the cognitive intervention using some kind of software for cognitive stimulation.

Regarding the pharmacological interventions, the majority of studies in this meta-analysis included participants treated with stable doses of cholinesterase inhibitors (donepezil or memantine), whereas Farina et al. (2002) and Cahn-Weiner et al. (2003) did not report any pharmacological treatment, Jelcic et al. (2012) and Hofmann et al. (2003) included participants with no pharmacological treatment, and Heiss et al. (1993; Heiss et al. 1994) included a group with dietary supplement and other group with a drug similar to vitamin B_6. Because pharmacological treatments used to enhance cognition in AD have shown small effects on cognitive and functional outcomes (Raina et al., 2008), we cannot rule out that the *ES*s reported here are a consequence of these interventions. It is worth noting that *ES*s for cognitive ($ES = 0.42$) and functional ($ES = 0.19$) variables from studies in which patients were taking cholinesterase inhibitors are still small and negligible, respectively.

This is, to our knowledge, the first meta-analysis on cognitive interventions in AD that calculates *ES*s controlling for their intercorrelations and measuring the effects of moderators such as the comparison groups. The main finding is that *ES*s for functional and cognitive outcomes were negligible to low and that did not differ when potential sources of bias were controlled (e.g., MMSE or functional *ES*s), an issue that future research should take into account when reporting findings. As pointed out by a reviewer, some works

compared a cognitive training group with a control group receiving cognitive stimulation, whereas other studies compared an experimental group receiving cognitive stimulation with a control group receiving placebo or pharma only. Thus, cognitive stimulation is used as an experimental intervention in some studies and as a control intervention in some others. This type of comparisons will not estimate the same *ES* parameter and limits the generalization of our results. Our moderator analyses showed no differences in *ES*s according to the control group intervention, probably due to power issues, but further research is needed to clarify whether cognitive interventions are more effective than pharma only or placebo, or even differently effective between each other (e.g., cognitive training vs. cognitive stimulation). We suggest that interventions should focus on personalized goals and should include measures related to trained situations instead of or in addition to standardized outcomes.

References

References marked with an asterisk indicate studies included in the meta-analysis.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders, 4th edition*. Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders, 4th edition, text revision*. Washington, DC: American Psychiatric Association.

Bahar-Fuchs, A., Clare, L., & Woods, B. (2013). Cognitive training and cognitive rehabilitation for mild to moderate Alzheimer's disease and vascular dementia. *Cochrane Database of Systematic Reviews, 6,* CD003260.

Ballard, C., Khan, Z., Clack, H., & Corbett, A. (2011). Nonpharmacological treatment of Alzheimer disease. *Canadian Journal of Psychiatry, 56,* 589–595.

Becker, J. T., Boller, F., Lopez, O. L., Saxton, J., & McGonigle, K. L. (1994). The natural history of Alzheimer's disease. Description of study cohort and accuracy of diagnosis. *Archives of Neurology, 51,* 585–594. http://dx.doi.org/10.1001/archneur.1994.00540180063015

Benedict, R. H. B., Schretlen, D., Groninger, L., & Brandt, J. (1998). Hopkins Verbal Learning Test–Revised: Normative data and analysis of inter-form and test–retest reliability. *The Clinical Neuropsychologist, 12,* 43–55. http://dx.doi.org/10.1076/clin.12.1.43.1726

*Bergamaschi, S., Arcara, G., Calza, A., Villani, D., Orgeta, V., & Mondini, S. (2013). One-year repeated cycles of cognitive training (CT) for Alzheimer's disease. *Aging Clinical and Experimental Research, 25,* 421–426. http://dx.doi.org/10.1007/s40520-013-0065-2

Birks, J., & Harvey, R. J. (2006). Donepezil for dementia due to Alzheimer's disease. *The Cochrane Collaboration, 1,* CD001190.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, United Kingdom: Wiley. http://dx.doi.org/10.1002/9780470743386.ch4

Brodaty, H., & Arasaratnam, C. (2012). Meta-analysis of nonpharmacological interventions for neuropsychiatric symptoms of dementia. *The American Journal of Psychiatry, 169,* 946–953. http://dx.doi.org/10.1176/appi.ajp.2012.11101529

Burns, A., & Iliffe, S. (2009). Alzheimer's disease. *British Medical Journal, 338,* b158. http://dx.doi.org/10.1136/bmj.b158

*Cahn-Weiner, D. A., Malloy, P. F., Rebok, G. W., & Ott, B. R. (2003). Results of a randomized placebo-controlled study of memory training for mildly impaired Alzheimer's disease patients. *Applied Neuropsychology, 10,* 215–223. http://dx.doi.org/10.1207/s15324826an1004_3

Capitani, E., Della Sala, S., Laiacona, M., Marchetti, C., & Spinnler, H. (1994). Standardizzazione Ed.uso clinico di un test di memoria di prosa [Standarization and clinical use of a prose memory test]. *Bollettino di Psicologia Applicata, 209,* 47–56.

Casey, D. A., Antimisiaris, D., & O'Brien, J. (2010). Drugs for Alzheimer's disease: Are they effective? *Pharmacy & Therapeutics, 35,* 208–211.

*Cipriani, G., Bianchetti, A., & Trabucchi, M. (2006). Outcomes of a computer-based cognitive rehabilitation program on Alzheimer's disease patients compared with those on patients affected by mild cognitive impairment. *Archives of Gerontology and Geriatrics, 43,* 327–335. http://dx.doi.org/10.1016/j.archger.2005.12.003

Citron, M. (2010). Alzheimer's disease: Strategies for disease modification. *National Review, 9,* 387–398. http://dx.doi.org/10.1038/nrd2896

Clare, L. (2003). Cognitive training and cognitive rehabilitation for people with early-stage dementia. *Reviews in Clinical Gerontology, 13,* 75–83. http://dx.doi.org/10.1017/S0959259803013170

Clare, L., Evans, S., Parkinson, C., Woods, R., & Linden, D. (2011). Goal-setting in cognitive rehabilitation for people with early-stage Alzheimer's disease. *Clinical Gerontologist, 34,* 220–236. http://dx.doi.org/10.1080/07317115.2011.555937

*Clare, L., Linden, D. E. J., Woods, R. T., Whitaker, R., Evans, S. J., Parkinson, C. H., . . . Rugg, M. D. (2010). Goal-oriented cognitive rehabilitation for people with early-stage Alzheimer disease: A single-blind randomized controlled trial of clinical efficacy. *The American Journal of Geriatric Psychiatry, 18,* 928–939. http://dx.doi.org/10.1097/JGP.0b013e3181d5792a

Clare, L., & Woods, B. (2008). Cognitive rehabilitation and cognitive training for early-stage Alzheimer's disease and vascular dementia. *Cochrane Database of Systematic Reviews, 4,* CD003260. http://dx.doi.org/10.1002/14651858.CD003260

Clare, L., Woods, R. T., Moniz Cook, E. D., Orrell, M., & Spector, A. (2003). Cognitive rehabilitation and cognitive training for early-stage Alzheimer's disease and vascular dementia. *Cochrane Database of Systematic Reviews, 4,* CD003260.

Cohn, L. D., & Becker, B. J. (2003). How meta-analysis increases statistical power. *Psychological Methods, 8,* 243–253. http://dx.doi.org/10.1037/1082-989X.8.3.243

Cooper, H. (2010). *Applied social research methods series: Vol. 2. Research synthesis and meta-analysis. A step-by-step approach* (4th ed.). Thousand Oaks, CA: SAGE Publications Inc.

*Cotelli, M., Manenti, R., Brambilla, M., Petesi, M., Rosini, S., Ferrari, C., . . . Miniussi, C. (2014). Anodal tDCS during face-name associations memory training in Alzheimer's patients. *Frontiers in Aging Neuroscience, 6,* 38.

*Davis, R. N., Massman, P. J., & Doody, R. S. (2001). Cognitive intervention in Alzheimer disease: A randomized placebo-controlled study. *Alzheimer Disease and Associated Disorders, 15,* 1–9. http://dx.doi.org/10.1097/00002093-200101000-00001

*Farina, E., Fioravanti, R., Chiavari, L., Imbornone, E., Alberoni, M., Pomati, S., . . . Mariani, C. (2002). Comparing two programs of cognitive training in Alzheimer's disease: A pilot study. *Acta Neurologica Scandinavica, 105,* 365–371. http://dx.doi.org/10.1034/j.1600-0404.2002.01086.x

Farina, E., Fioravanti, R., Chiavari, L., Imbornone, E., Pomati, S., Pinardi, G., . . . Mariani, C. (1999). Functional Living Skills Assessment: A standardized instrument built to monitor activities of daily living in demented patients: Preliminary data. *Journal of Neurology, 246*(Suppl. 1), I/101.

*Farina, E., Mantovani, F., Fioravanti, R., Pignatti, R., Chiavari, L., Imbornone, E., . . . Nemni, R. (2006). Evaluating two group programmes of cognitive training in mild-to-moderate AD: Is there any difference between a "global" stimulation and a "cognitive-specific" one? *Aging & Mental Health, 10,* 211–218. http://dx.doi.org/10.1080/13607860500409492

*Fernández-Calvo, B., Contador, I., Ramos, F., Olazarán, J., Mograbi, D. C., & Morris, R. G. (2015). Effect of unawareness on rehabilitation outcome in a randomised controlled trial of multicomponent intervention for patients with mild Alzheimer's disease. *Neuropsychological Rehabilitation, 25,* 448–477. http://dx.doi.org/10.1080/09602011.2014.948461

Field, A. P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed- and random-effects methods. *Psychological Methods, 6,* 161–180. http://dx.doi.org/10.1037/1082-989X.6.2.161

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198. http://dx.doi.org/10.1016/0022-3956(75)90026-6

Gaitán, A., Garolera, M., Cerulla, N., Chico, G., Rodriguez-Querol, M., & Canela-Soler, J. (2013). Efficacy of an adjunctive computer-based cognitive training program in amnestic mild cognitive impairment and Alzheimer's disease: A single-blind, randomized clinical trial. *International Journal of Geriatric Psychiatry, 28,* 91–99. http://dx.doi.org/10.1002/gps.3794

*Galante, E., Venturini, G., & Fiaccadori, C. (2007). Computer-based cognitive intervention for dementia: Preliminary results of a randomized clinical trial. *Giornale Italiano di Medicina del Lavoro Ed. Ergonomia, Suppl B. Psicologia, 29,* 26–32.

Galton, C. J., Patterson, K., Xuereb, J. H., & Hodges, J. R. (2000). Atypical and typical presentations of Alzheimer's disease: A clinical, neuropsychological, neuroimaging and pathological study of 13 cases. *Brain: A Journal of Neurology, 123,* 484–498. http://dx.doi.org/10.1093/brain/123.3.484

Gardette, V., Coley, N., & Andrieu, S. (2010). Non-pharmacologic therapies: A different approach to AD. *The Canadian Review of Alzheimer's Disease and Other Dementias, 13,* 13–22.

*Giordano, M., Dominguez, L. J., Vitrano, T., Curatolo, M., Ferlisi, A., Di Prima, A., . . . Barbagallo, M. (2010). Combination of intensive cognitive rehabilitation and donepezil therapy in Alzheimer's disease (AD). *Archives of Gerontology and Geriatrics, 51,* 245–249. http://dx.doi.org/10.1016/j.archger.2009.11.008

Grandmaison, E., & Simard, M. (2003). A critical review of memory stimulation programs in Alzheimer's Disease. *The Journal of Neuropsychiatry and Clinical Neurosciences, 15*(2), 130–144.

Hedges, L. V., & Pigott, T. D. (2001). The power of statistical tests in meta-analysis. *Psychological Methods, 6,* 203–217. http://dx.doi.org/10.1037/1082-989X.6.3.203

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1,* 39–65. http://dx.doi.org/10.1002/jrsm.5

*Heiss, W. D., Kessler, J., Mielke, R., Szelies, B., Herholz, K., & Herholz, K. (1994). Long-term effects of phosphatidylserine, pyritinol, and cognitive training in Alzheimer's disease. A neuropsychological, EEG, and PET investigation. *Dementia, 5,* 88–98.

*Heiss, W. D., Kessler, J., Slansky, I., Mielke, R., Szelies, B., & Herholz, K. (1993). Activation PET as an instrument to determine therapeutic efficacy in Alzheimer's disease. *Annals of the New York Academy of Sciences, 695,* 327–331. http://dx.doi.org/10.1111/j.1749-6632.1993.tb23078.x

Higgins, J. P. T., & Green, S. (2008). *Cochrane handbook for systematic reviews of interventions: Cochrane book series*. Chichester, United Kingdom: Wiley and The Cochrane Collaboration. http://dx.doi.org/10.1002/9780470712184

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21,* 1539–1558. http://dx.doi.org/10.1002/sim.1186

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327,* 557–560. http://dx.doi.org/10.1136/bmj.327.7414.557

Hindmarch, I., Lehfeld, H., de Jongh, P., & Erzigkeit, H. (1998). The Bayer Activities of Daily Living Scale (B-ADL). *Dementia and Geriatric Cognitive Disorders, 9,* 20–26. http://dx.doi.org/10.1159/000051195

*Hofmann, M., Rösler, A., Schwarz, W., Müller-Spahn, F., Kräuchi, K., Hock, C., & Seifritz, E. (2003). Interactive computer-training as a therapeutic tool in Alzheimer's disease. *Comprehensive Psychiatry, 44,* 213–219. http://dx.doi.org/10.1016/S0010-440X(03)00006-3

*Jelcic, N., Cagnin, A., Meneghello, F., Turolla, A., Ermani, M., & Dam, M. (2012). Effects of lexical-semantic treatment on memory in early Alzheimer disease: An observer-blinded randomized controlled trial. *Neurorehabilitation and Neural Repair, 26,* 949–956. http://dx.doi.org/10.1177/1545968312440146

Jüni, P., Holenstein, F., Sterne, J., Bartlett, C., & Egger, M. (2002). Direction and impact of language bias in meta-analyses of controlled trials: Empirical study. *International Journal of Epidemiology, 31,* 115–123. http://dx.doi.org/10.1093/ije/31.1.115

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. (1963). The index of ADL: A standardized measure of biological and psychological function. *American Journal of Medical Association, 185,* 94–99. http://dx.doi.org/10.1001/jama.1963.03060120024016

Kessler, J., Denzler, P., & Moskowitsch, H. J. (1988). *Der Demenz-test*. Weinheim, Germany: Beltz.

Kurz, A., Thöne-Otto, A., Cramer, B., Egert, S., Frölich, L., Gertz, H. J., . . . Werheid, K. (2012). CORDIAL: Cognitive rehabilitation and cognitive-behavioral treatment for early dementia in Alzheimer disease: A multicenter, randomized, controlled trial. *Alzheimer Disease and Associated Disorders, 26,* 246–253. http://dx.doi.org/10.1097/WAD.0b013e318231e46e

Lambon Ralph, M. A., Patterson, K., Graham, N., Dawson, K., & Hodges, J. R. (2003). Homogeneity and heterogeneity in mild cognitive impairment and Alzheimer's disease: A cross-sectional and longitudinal study of 55 cases. *Brain: A Journal of Neurology, 126,* 2350–2362. http://dx.doi.org/10.1093/brain/awg236

Lanctôt, K. L., Rajaram, R. D., & Herrmann, N. (2009). Therapy for Alzheimer's disease: How effective are current treatments? *Therapeutic Advances in Neurological Disorders, 2,* 163–180. http://dx.doi.org/10.1177/1756285609102724

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist, 9,* 179–186. http://dx.doi.org/10.1093/geront/9.3_Part_1.179

Linn, M. W., & Linn, B. S. (1982). The rapid disability rating scale-2. *Journal of the American Geriatrics Society, 30,* 378–382. http://dx.doi.org/10.1111/j.1532-5415.1982.tb02835.x

*Loewenstein, D. A., Acevedo, A., Czaja, S. J., & Duara, R. (2004). Cognitive rehabilitation of mildly impaired Alzheimer disease patients on cholinesterase inhibitors. *The American Journal of Geriatric Psychiatry, 12,* 395–402. http://dx.doi.org/10.1097/00019442-200407000-00007

Lopez, O. L., McDade, E., Riverol, M., & Becker, J. T. (2011). Evolution of the diagnostic criteria for degenerative and cognitive disorders. *Current Opinion in Neurology, 24,* 532–541. http://dx.doi.org/10.1097/WCO.0b013e32834cd45b

Magalhães, R., Alves, J., Thomas, R. E., Chiaravalloti, N., Gonçalves, Ó. F., Petrosyan, A., & Sampaio, A. (2014). Are cognitive interventions for multiple sclerosis effective and feasible? *Restorative Neurology and Neuroscience, 32,* 623–638.

McDaniel, M. A., & Bugg, J. M. (2012). Memory training interventions: What has been forgotten? *Journal of Applied Research in Memory & Cognition, 1,* 58–60. http://dx.doi.org/10.1016/j.jarmac.2011.11.002

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology, 34,* 939–944. http://dx.doi.org/10.1212/WNL.34.7.939

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Jr., Kawas, C. H., . . . Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia, 7,* 263–269. http://dx.doi.org/10.1016/j.jalz.2011.03.005

Metitieri, T., Zanetti, O., Geroldi, C., Frisoni, G. B., De Leo, D., Dello Buono, M., . . . Trabucchi, M. (2001). Reality orientation therapy to delay outcomes of progression in patients with dementia. A retrospective study. *Clinical Rehabilitation, 15,* 471–478. http://dx.doi.org/10.1191/026921501680425199

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & the PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6,* e1000097. http://dx.doi.org/10.1371/journal.pmed.1000097

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11,* 364–386. http://dx.doi.org/10.1177/1094428106291059

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7,* 105–125. http://dx.doi.org/10.1037/1082-989X.7.1.105

Morrison, A., Polisena, J., Husereau, D., Moulton, K., Clark, M., Fiander, M., . . . Rabb, D. (2012). The effect of English-language restriction on systematic review-based meta-analyses: A systematic review of empirical studies. *International Journal of Technology Assessment in Health Care, 28,* 138–144. http://dx.doi.org/10.1017/S0266462312000086

National Collaborating Centre for Mental Health. (2007). *Dementia: A NICE-SCIE Guideline on supporting people with dementia and their carers in health and social care (NICE clinical guidelines, No. 42)*. Leicester, United Kingdom: British Psychological Society.

Novelli, G., Papagno, C., Capitani, E., Laiacona, N., Vallar, G., & Cappa, S. F. (1986). Tre tests clinici di memoria verbale a lungo termine. Taratura su soggetti normali. *Archivio di Psicologia, Neurologia, e Psichiatria, 2,* 278–296.

Olazarán, J., Muñiz, R., Reisberg, B., Peña-Casanova, J., del Ser, T., Cruz-Jentoft, A. J., . . . Sevilla, C. (2004). Benefits of cognitive-motor intervention in MCI and mild to moderate Alzheimer disease. *Neurology, 63,* 2348–2353. http://dx.doi.org/10.1212/01.WNL.0000147478.03911.28

Olazarán, J., Reisberg, B., Clare, L., Cruz, I., Peña-Casanova, J., Del Ser, T., . . . Muñiz, R. (2010). Nonpharmacological therapies in Alzheimer's disease: A systematic review of efficacy. *Dementia and Geriatric Cognitive Disorders, 30,* 161–178. http://dx.doi.org/10.1159/000316119

Onder, G., Zanetti, O., Giacobini, E., Frisoni, G. B., Bartorelli, L., Carbone, G., . . . Bernabei, R. (2005). Reality orientation therapy combined with cholinesterase inhibitors in Alzheimer's disease: Randomised controlled trial. *British Journal of Psychiatry, 187,* 450–455. http://dx.doi.org/10.1192/bjp.187.5.450

Raina, P., Santaguida, P., Ismaila, A., Patterson, C., Cowan, D., Levine, M., . . . Oremus, M. (2008). Effectiveness of cholinesterase inhibitors and memantine for treating dementia: Evidence review for a clinical practice guideline. *Annals of Internal Medicine, 148,* 379–397. http://dx.doi.org/10.7326/0003-4819-148-5-200803040-00009

*Requena, C., Maestú, F., Campo, P., Fernández, A., & Ortiz, T. (2006). Effects of cholinergic drugs and cognitive training on dementia: 2-year follow-up. *Dementia and Geriatric Cognitive Disorders, 22,* 339–345. http://dx.doi.org/10.1159/000095600

Reuben, D. B., Laliberte, L., Hiris, J., & Mor, V. (1990). A hierarchical exercise scale to measure function at the Advanced Activities of Daily Living (AADL) level. *Journal of the American Geriatrics Society, 38,* 855–861. http://dx.doi.org/10.1111/j.1532-5415.1990.tb05699.x

Rey, A. (1964). *L'examen clinique en psychologie*. Paris, France: Press Universitaire de France.

Rohling, M. L., Faust, M. E., Beverly, B., & Demakis, G. (2009). Effectiveness of cognitive rehabilitation following acquired brain injury: A meta-analytic re-examination of Cicerone et al.'s (2000, 2005) systematic reviews. *Neuropsychology, 23,* 20–39. http://dx.doi.org/10.1037/a0013659

Simmons-Stern, N. R., Budson, A. E., & Ally, B. A. (2010). Music as a memory enhancer in patients with Alzheimer's disease. *Neuropsychologia, 48,* 3164–3167. http://dx.doi.org/10.1016/j.neuropsychologia.2010.04.033

Sitzer, D. I., Twamley, E. W., & Jeste, D. V. (2006). Cognitive training in Alzheimer's disease: A meta-analysis of the literature. *Acta Psychiatrica Scandinavica, 114,* 75–90. http://dx.doi.org/10.1111/j.1600-0447.2006.00789.x

Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods, 5,* 13–30. http://dx.doi.org/10.1002/jrsm.1091

*Tárraga, L., Boada, M., Modinos, G., Espinosa, A., Diego, S., Morera, A., . . . Becker, J. T. (2006). A randomised pilot study to assess the efficacy of an interactive, multimedia tool of cognitive stimulation in Alzheimer's disease. *Journal of Neurology, Neurosurgery, and Psychiatry, 77,* 1116–1121. http://dx.doi.org/10.1136/jnnp.2005.086074

Turner, R. M., Bird, S. M., & Higgins, J. P. T. (2013). The impact of study size on meta-analyses: Examination of underpowered studies in Cochrane reviews. *PLoS ONE, 8,* e59202. http://dx.doi.org/10.1371/journal.pone.0059202

Vandenbroucke, J. P., von Elm, E., Altman, D. G., Gotzsche, P. C., Mulrow, C. D., Pocock, S. J., & Egger, M. (2007). Strengthening the reporting of observational studies in epidemiology (STROBE): Explanation and elaboration. *PLoS Medicine, 4,* e, 297, 1628–1654. http://dx.doi.org/10.1371/journal.pmed.0040297

*Viola, L. F., Nunes, P. V., Yassuda, M. S., Aprahamian, I., Santos, F. S., Santos, G. D., . . . Forlenza, O. V. (2011). Effects of a multidisciplinary cognitive rehabilitation program for patients with mild Alzheimer's disease. *Clinics, 66,* 1395–1400. http://dx.doi.org/10.1590/S1807-59322011000800015

Wechsler, D. (1987). *Wechsler memory scale–revised: Manual*. San Antonio, TX: The Psychological Corporation.

Wilson, B. A., Baddeley, A. D., & Cockburn, J. (2003). *Rivermead Behavioural Memory Test II*. Bury St. Edmunds, England: Thames Valley Test Company.

Zanetti, O., Zanieri, G., Di Giovanni, G., De Vreese, L. P., Pezzini, A., Metitieri, T., & Trabucchi, M. (2001). Effectiveness of procedural memory stimulation in mild Alzheimer's disease patients: A controlled study. *Neuropsychological Rehabilitation, 11,* 263–272. http://dx.doi.org/10.1080/09602010042000088

(*Appendices follow*)

Appendix A

Methodological Quality Indices for Research Studies

Each item of the index is scored as Y = "yes" or N = "no" for each individual research study.

Randomization: The authors clearly state that participants were randomly assigned to groups.

Double Blind: The authors clearly state that participants were blind to the study group, *and* assessors were blind to participant's allocation.

Single Blind: The authors clearly state that participants were blind to the study group *or* assessors were blind to participant's allocation.

Independence of Assessors: The authors clearly state that pre/post assessments and interventions were carried out by different professionals.

Control Intervention: The authors clearly state that there was a control intervention for comparison purposes tapping specific cognitive abilities.

Placebo Intervention: The authors clearly state that there was a control intervention *not* tapping specific cognitive abilities.

Inclusion Criteria: The authors clearly state the criteria needed to be included in the study.

Exclusion Criteria: The authors clearly state the criteria needed to be excluded from the study.

Control of Selection Bias: The authors clearly state potential sources of bias at baseline and control them some way (e.g., methodologically, statistically). This includes analyzing potential differences between groups on demographical *and* outcome variables used in the primary study (e.g., age, sex, educational level, memory outcomes, etc.)

Follow-Up Assessment: The authors include a follow-up assessment to check the generalization of any positive finding.

Standardized Assessment: Standard tests and tasks were used to measure cognitive functioning and/or memory.

Functional Assessment: Functional assessment was measured individually using functional questionnaires and scales.

Detailed Explanation of the Intervention: The authors state clearly the intervention used in the study. The intervention explanation must include timing, type of exercises, some example of the exercises, who was in charge of the intervention, the cognitive domains covered, and individually/group driven information to be considered a detailed explanation.

Attrition Causes: The authors clearly state the initial sample size, the causes of attrition and sample mortality and the final sample size.

(*Appendices continue*)

Appendix B

Formulas for Meta-Analysis

Formulas for Correlational Meta-Analysis

$$z_{r_i} = \frac{1}{2} Log_e\left(\frac{1 + r_i}{1 - r_i}\right)$$

$$ES_z = \frac{\sum_{i=1}^{k} w_i^* z_{r_i}}{\sum_{i=1}^{k} w_i^*}.$$

$$w_z^* = \frac{1}{\frac{1}{(n-3)} + \tau^2}$$

$$\tau^2 = \frac{Q - (k-1)}{c}$$

$$c = \sum_{i=1}^{k} (n-3) - \frac{\sum_{i=1}^{k} (n-3)^2}{\sum_{i=1}^{k} (n-3)}$$

$$Q = \sum_{i=1}^{k} w_z^* (z_{r_i} - ES_z)^2$$

$$SE_{ES_z} = \sqrt{\frac{1}{\sum_{i=1}^{k} w_z^*}}$$

$$ES_r = \frac{e^{(2ES_z)} - 1}{e^{(2ES_z)} + 1}$$

Formulas for Meta-Analysis on Means

ES in the experimental group: $d_E = M_{POST} - M_{PRE}/SD_{PRE}$
ES in the control group: $d_C = M_{POST} - M_{PRE}/SD_{PRE}$

$$c(df) = 1 - \frac{3}{4df - 1} [df = n - 1]$$

Adjusted $d_E = c(df) * d_E$
Adjusted $d_C = c(df) * d_C$
Effect size (*ES*): $d_E - d_C$

$$\sigma_E^2 = C_E^2 [\left(\frac{2(1 - r_E)}{n}\right)\left(\frac{n-1}{n-3}\right)\left(1 + \frac{nd_E^2}{2(1-r)}\right)] - d_E^2$$

$$\sigma_C^2 = C_C^2 [\left(\frac{2(1 - r_C)}{n}\right)\left(\frac{n-1}{n-3}\right)\left(1 + \frac{nd_C^2}{2(1-r)}\right)] - d_C^2$$

$$\sigma_{ES}^2 = \sigma_E^2 + \sigma_C^2$$

$$ES_{FE} = \frac{\sum w_i d}{\sum w_i}$$

$$w_i = \frac{1}{\sigma_d^2}$$

$$\sigma_{ES}^2 = \frac{1}{\sum w_i}$$

Confidence intervals = $ES \pm 1.96 * \sqrt{\sigma_{ES}^2}$

Hedges' $Q = \frac{k\varpi_d^2}{\varpi_e^2} [df = k - 1]$

Observed variance $(\varpi_e^2) = \frac{\sum w_i (d - ES)^2}{\sum w_i}$

Variance due to sampling error = $\frac{k}{\sum_{i=1}^{k} w_i}$

$$I^2 = (Q - df)/Q$$

$$\tau^2 = \begin{cases} \frac{Q - df}{C} & \text{if } Q > df \\ 0 & \text{if } Q \leq df \end{cases}$$

$$C = \sum w_i - \frac{\sum w_i^2}{\sum w_i}$$

$$ES_{RE} = \frac{\sum w_i^* d}{\sum w_i^*}$$

$$w_i^* = \frac{1}{\sigma_d^2 + \tau^2}$$

(*Appendices continue*)

Appendix C

Sensitivity Analysis

		95% CI		
r	*ES*	*LL*	*UL*	Var
−0.9	0.29	0.43	0.15	0.00
−0.8	0.29	0.43	0.15	0.00
−0.7	0.29	0.42	0.15	0.00
−0.6	0.29	0.42	0.15	0.00
−0.5	0.29	0.42	0.16	0.00
−0.4	0.29	0.42	0.16	0.00
−0.3	0.29	0.42	0.16	0.00
−0.2	0.29	0.41	0.16	0.00
−0.1	0.29	0.41	0.16	0.00
0	0.29	0.41	0.16	0.00
0.1	0.28	0.41	0.16	0.00
0.2	0.28	0.40	0.16	0.00
0.3	0.28	0.40	0.16	0.00
0.4	0.28	0.39	0.17	0.00
0.5	0.28	0.39	0.17	0.00
0.6	0.27	0.38	0.17	0.00
0.7	0.27	0.37	0.17	0.00
0.8	0.26	0.36	0.17	0.00
0.9	0.26	0.35	0.17	0.00
Original	0.29	0.41	0.16	0.00
With mean *r*	0.27	0.37	0.17	0.00

Note. *ES* = effect size; CI = confidence interval; *LL* = lower limit; *UL* = upper limit; Var = variance.

Appendix D

Correlations Between Moderator Variables at the Study Level

Study characteristics	3	4	5	6	7
1. Random Y	.095	−.095	−.095	.368	−.039
2. Random N	−.095	.095	.095	−.368	.039
3. I.Assess Y	1		−.321	.368	−.287
4. I.Assess N		1	.321	−.368	.287
5. C.Cogn			1	**−.587**	.039
6. C.Plac				1	**−.630**
7. C.Pharm					1

Note. *N* = 19. Y = yes; N = no; I.Asses = Independence of assessors; C.Cogn = Control: cognitive group; C.Plac = Control: placebo; C.Pharm = Control: pharma only or no intervention. Bold coefficients indicate $p < .01$. ϕ 1,2 = −1; ϕ 3,4 = −1. The conclusions drawn from the table are as follows: (a) Studies with a cognitive stimulation control group had a placebo control group less frequently; (b) studies with a placebo control group had a pharma only control group less frequently; and (c) moderators are not collinear with each other at the study level.

(*Appendices continue*)

Appendix E

Effect Size for Each Study and Outcome

Study	Outcome	Adjusted $d_{treatment}$	Adjusted $d_{control}$	d	$\sigma^2(d)$	95% CI	
						LL	*UL*
1. Heiss et al. (1993)	MMSE	−0.36	−0.14	−0.21	0.13	−0.91	0.48
	Verbal Selective reminding	0.33	−0.06	0.39	0.12	−0.30	1.09
	MMSE	−0.11	−0.14	0.03	0.12	−0.65	0.70
	Verbal Selective reminding	−0.23	−0.06	−0.16	0.12	−0.84	0.52
	MMSE	0.40	−0.14	0.54	0.13	−0.16	1.25
	Verbal Selective reminding	0.55	−0.06	0.61	0.14	−0.12	1.34
2. Heiss et al. (1994)	MMSE	−0.26	−0.22	−0.05	0.07	−0.56	0.47
	Verbal Selective reminding	0.30	0.06	0.24	0.07	−0.27	0.75
	MMSE	−0.04	−0.22	0.18	0.07	−0.33	0.69
	Verbal Selective reminding	0.11	0.06	0.05	0.07	−0.46	0.55
	MMSE	0.26	−0.22	0.48	0.07	−0.03	0.99
	Verbal Selective reminding	0.29	0.06	0.23	0.07	−0.27	0.75
3. Davis et al. (2001)	MMSE	0.04	0.05	−0.01	0.06	−0.49	0.47
	LM-Immediate	0.26	−0.04	0.30	0.06	−0.19	0.79
	LM-Delayed	0.41	0.48	−0.07	0.07	−0.60	0.46
4. Farina et al. (2002)	MMSE	0.46	0.17	0.29	0.13	−0.43	1.00
	RBMT-SS	0.25	0.43	−0.17	0.13	−0.89	0.55
	RBMT-PS	0.19	0.24	−0.05	0.12	−0.73	0.64
	ADL (Katz)	−0.23	0.01	−0.23	0.12	−0.91	0.44
	IADL (Lawton)	0.10	−0.10	0.20	0.12	−0.47	0.88
5. Cahn-Weiner et al. (2003)	HVLT-Total learning	−0.08	−0.34	0.25	0.07	−0.27	0.77
	HVLT-Delayed recall	−0.24	−0.29	0.05	0.07	−0.47	0.58
	HVLT-Recognition	−0.13	−0.33	0.20	0.07	−0.32	0.72
	ADL (Lawton)	−0.04	−0.02	−0.02	0.07	−0.52	0.48
6. Hofmann et al. (2003)	MMSE	0.00	0.67	−0.67	0.16	−1.45	0.11
7. Loewenstein et al. (2004)	Face-name assoc. 3 learning	1.29	0.38	0.91	0.09	0.31	1.51
	Face-name assoc. Delayed	1.20	0.06	1.13	0.08	0.57	1.70
	Bayer ADLS[a]	−0.34	−0.37	0.03	0.06	−0.44	0.50
8. Cipriani et al. (2006)	MMSE	0.80	0.46	0.34	0.18	−0.49	1.17
	RBMT-SS	0.19	0.55	−0.35	0.14	−1.09	0.38
	AADL (Reuben)[a]	−0.07	0.00	−0.07	0.12	−0.74	0.60
9. Farina et al. (2006)	MMSE	0.42	−0.03	0.44	0.08	−0.10	0.99
	RBMT-PS	−0.03	−0.06	0.03	0.07	−0.49	0.55
	RBMT-SS	0.26	0.07	0.19	0.07	−0.34	0.72
	FLSA	0.12	0.22	−0.10	0.07	−0.63	0.43
	ADL (Katz)	0.13	−0.35	0.48	0.08	−0.06	1.02
	IADL (Lawton)	0.27	0.02	0.25	0.07	−0.28	0.78
10. Requena et al. (2006)	MMSE	0.29	−0.43	0.72	0.06	0.22	1.21
	ADAS-Cog[a]	0.59	−0.51	1.10	0.07	0.57	1.63
	MMSE	0.29	0.29	0.00	0.06	−0.49	0.50
	ADAS-Cog[a]	0.59	0.21	0.39	0.07	−0.13	0.91
	MMSE	0.29	−1.22	1.51	0.11	0.85	2.16
	ADAS-Cog[a]	0.59	−1.00	1.59	0.10	0.97	2.22
11. Tárraga et al. (2006)	ADAS-Cog[a]	0.18	−0.19	0.36	0.08	−0.18	0.90
	MMSE	0.66	0.04	0.62	0.09	0.03	1.21
	Kurtz test	0.05	−0.34	0.39	0.08	−0.16	0.93
	RBMT - Story Recall Imm	0.37	0.08	0.29	0.08	−0.26	0.84
	ADAS-Cog[a]	0.18	−0.39	0.57	0.09	−0.03	1.17
	MMSE	0.66	−0.59	1.25	0.12	0.56	1.94
	Kurtz test	0.05	0.22	−0.17	0.09	−0.75	0.41
12. Galante et al. (2007)	MMSE	0.03	0.73	−0.70	0.60	−2.22	0.82
	Prose memory	−0.02	0.08	−0.10	0.29	−1.16	0.96
	BADL	0.09	−0.44	0.52	0.40	−0.72	1.76
	IADL (Lawton)	−0.24	0.61	−0.84	0.51	−2.24	0.56
13. Clare et al. (2010)	RBMT-PS	−0.06	−0.28	0.22	0.04	−0.16	0.60
	RBMT-PS	−0.06	−0.04	−0.03	0.03	−0.36	0.31
14. Giordano et al. (2010)	MMSE	0.53	0.09	0.43	0.03	0.12	0.74
	ADAS-Cog[a]	0.85	0.28	0.57	0.03	0.23	0.91
	ADL[a] (No reference)	0.30	0.09	0.21	0.02	−0.09	0.51
	IADL[a] (No reference)	−0.37	−0.20	−0.17	0.02	−0.48	0.13
15. Viola et al. (2011)	MMSE	−0.03	−0.22	0.19	0.04	−0.23	0.60
	SKT[a]	−0.04	−0.21	0.18	0.03	−0.17	0.52

Appendix E (*continued*)

Study	Outcome	Adjusted $d_{treatment}$	Adjusted $d_{control}$	d	$\sigma^2(d)$	95% CI *LL*	95% CI *UL*
16. Jelcic et al. (2012)	MMSE	0.69	−0.37	1.05	0.07	0.52	1.59
	Brief Stroy Recall-Immediate	0.48	0.30	0.18	0.07	−0.32	0.68
	Brief Stroy Recall-Delayed	0.50	−0.07	0.57	0.06	0.08	1.06
	RAVLT-Immediate	0.33	−0.02	0.36	0.06	−0.12	0.83
	RAVLT-Delayed	0.46	−0.25	0.72	0.06	0.22	1.21
	IADL (Lawton)	0.06	0.00	0.06	0.06	−0.40	0.52
17. Bergamaschi et al. (2013)	MMSE	0.88	−1.69	2.57	0.24	1.61	3.53
	Story Recall-immediate	0.33	−0.05	0.39	0.15	−0.37	1.14
	ADL (Katz)	−0.26	−1.15	0.89	0.10	0.26	1.51
	IADL (Lawton)	−0.14	−0.52	0.38	0.04	−0.02	0.77
18. Cotelli et al. (2014)	MMSE	0.19	0.40	−0.20	0.15	−0.97	0.56
	RBMT Story recall − Imm	0.62	0.00	0.62	0.31	−0.46	1.70
	RBMT Story recall − Del.	1.55	0.40	1.15	0.43	−0.13	2.43
	RAVLT Imm	0.15	0.27	−0.12	0.04	−0.53	0.29
	RAVLT Del	0.80	0.10	0.70	0.16	−0.08	1.48
	ADL (No reference)	−0.01	0.13	−0.14	0.07	−0.66	0.38
	IADL (No reference)	−0.01	−0.01	0.00	0.09	−0.60	0.60
	MMSE	0.19	0.08	0.11	0.13	−0.59	0.82
	RBMT Story recall − Imm	0.62	0.47	0.16	0.29	−0.90	1.21
	RAVLT Imm	0.15	0.38	−0.23	0.06	−0.71	0.25
	RAVLT Del	0.80	0.31	0.49	0.23	−0.45	1.43
	ADL (No reference)	−0.01	−0.27	0.26	0.08	−0.30	0.82
	IADL (No reference)	−0.01	−0.29	0.29	0.09	−0.29	0.87
19. Fernández-Calvo et al. (2015)	ADAS-Cog[a]	0.18	−0.87	1.05	0.06	0.59	1.51
	RDRS-2	0.09	−0.52	0.60	0.04	0.20	1.01

Note. RBMT = Rivermead Behavioral Memory Test; SS = Standard Score; PS = Profile Score; MMSE = Mini-Mental State Examination; ADL = activities of daily living; IADL = Instrumental ADL; AADL = Advance ADL; LM = Logical Memory subtest; RAVLT = Rey Auditory Verbal Learning Test; ADAS-Cog = Alzheimer's Disease Assessment Scale–Cognitive subscale; SKT = Short Cognitive Test; RDRS = Rapid Disability Rating Scale–2.
[a] The sign of the effect size is inverted for comparison purposes as lower scores indicate improvement (ADAS-Cog) or slower decline (MMSE).

Received June 22, 2015
Revision received January 18, 2016
Accepted February 15, 2016 ■