



The role of semantic assessment in the differential diagnosis between late-life depression and Alzheimer's disease or amnesic mild cognitive impairment: systematic review and meta-analysis

Sandra Invernizzi^{1,2} · Alice Bodart¹ · Laurent Lefebvre¹ · Isabelle Simoes Loureiro¹

Accepted: 25 July 2023 / Published online: 10 August 2023
© The Author(s) 2023, corrected publication 2023

Abstract

Object The cognitive complaints encountered in late-life depression (LLD) make it difficult to distinguish from amnesic mild cognitive impairment (aMCI) and Alzheimer's disease (AD) based on an analysis of neurocognitive disorders. The hypothesis of the early impairment of semantic memory in AD and aMCI is considered a potential differential cognitive clue, but the absence of this impairment has not yet been confirmed in LLD.

Method Based on the PRISMA method, we systematically seek neuropsychological assessments of individuals with LLD, the present study included 31 studies representing 3291 controls and 2820 people with LLD. Wherever possible, studies that tested simultaneously groups with LLD, AD (or aMCI) were also included. The results of the group of neuropsychological tasks relying on semantic memory were analyzed in two groups of tasks with high- or low-executive demand. The mean average effect of LLD was calculated and compared to the incremental effect of aMCI or AD on the scores. Linear regressions including education, age, and severity and type of depression were run to seek their power of prediction for the mean average effects.

Results LLD has a medium effect on scores at semantic and phonemic fluency and naming and a small average effect on the low-executive demand tasks. Differences in education is a predictor of the effect of LLD on phonemic fluency and naming but not on semantic fluency or on low-executive demand tasks. Except for semantic fluency, aMCI did not demonstrate an incremental effect on the scores compared to LLD, while AD did, for all the tasks except phonemic fluency.

Conclusion Assessment of semantic memory can be a discriminating clue for the distinction between depression and Alzheimer's disease but some methodological variables are highly influential to the scores, especially education. However, high-executive semantic tasks alone do not allow us to clearly distinguish LLD from AD or aMCI, as both pathologies seem to have a largely dialectical influential relationship, but low-executive semantic tasks appear as more sensible to this pathological distinction.

Keywords A systematic review · Semantic memory · Neuropsychology · Late-life depression · Alzheimer's differential diagnostic · Fluency task · Meta-analysis

Responsible Editor: Matthias Kliegel.

✉ Sandra Invernizzi
Sandra.invernizzi@umons.ac.be

Alice Bodart
alice.bodart@umons.ac.be

Laurent Lefebvre
laurent.lefebvre@umons.ac.be

Isabelle Simoes Loureiro
isabelle.simoesloureiro@umons.ac.be

¹ Department of Cognitive Psychology and Neuropsychology,
University of Mons, Mons, Belgium

² Fonds National de La Recherche Scientifique, Brussel,
Belgium

Introduction

It is now largely recognized that the cognitive profile of late-life depression (LLD) is affected in a manner that resembles what can be encountered in Alzheimer's disease (AD) and mild cognitive impairment (MCI).

AD is a major neurocognitive disorder encountering the criteria for a dementia syndrome listed in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition [DSM-V] and affecting the activities of daily life (Petersen 2016; Srinivasan 2014). When affected by the early stage of AD, the patient's most common results at the neuropsychological assessment demonstrate an impairment of episodic memory, language, attention, executive functions, and visuospatial abilities (Weintraub et al. 2012).

Furthermore, the prodromal phase of AD is often associated with MCI, which is determined by cognitive decline objectivated by normed tests, but for which the scores do not meet the criteria for dementia syndrome (Petersen et al. 1999). Not considered a neurodegenerative condition per se, MCI is itself a tricky condition that may or may not evolve into a neurodegenerative disorder (De Rotrou et al. 2005). Subdivided into four subtypes depending on the number of cognitive domains affected, MCI can exist as single-domain amnesic MCI (aMCI), multidomain amnesic MCI, single-domain non-amnesic MCI, and multidomain non-amnesic MCI (Díaz-Mardomingo et al. 2017). The focus of this paper is on the form of MCI that is currently identified as the most likely to progress to AD, namely aMCI, although the debate on the link between certain subtypes of MCI and their progression to one or another major neurocognitive disorder is still largely ongoing. Formerly identified as MCI of Alzheimer-type (Dubois and Albert 2004), also known as the hippocampal type, aMCI is essentially defined by an assessment of episodic memory with poor results at the free recall task, and an insufficient effect of cuing to help this recall despite adequate encoding.

In LLD, the extended damage to gray matter volume in the prefrontal, medial-temporal (Ballmaier et al. 2004; Kohler et al. 2010; Lamar et al. 2012), and subcortical cortices (Kohler et al. 2010; Lamar et al. 2012), results in cognitive impairment that also affects episodic memory (Elderkin-Thompson et al. 2006; Lamar et al. 2012) and executive functioning (Alexopoulos et al. 2005; Elderkin-Thompson et al. 2006; Henry and Crawford 2005; Morimoto et al. 2011; Rajtar-Zembaty et al. 2017; Snyder 2013), and, in a more secondary manner, attention (Lee et al. 2012; Rajtar-Zembaty et al. 2017), information processing (Lamar et al. 2012; Lee et al. 2012), learning (Lee et al. 2012), visuospatial (Lamar et al. 2012) and language (Rajtar-Zembaty et al. 2017) skills.

Consequently, these three pathological situations show important overlaps regarding discernible cognitive impairment. The diagnostic confusion is also increased by the fact that both aMCI and early stage of AD are frequently accompanied by important depressive symptoms that can influence the path of the disorder evolution, and by the fact that cases of LLD are sometimes considered as part of the prodromal phase of AD (for review see (Invernizzi et al. 2021)). The role of depression in the evolution from MCI to AD is an object of interesting research, especially about the key role of loneliness (see Lara et al. 2019; Rozzini et al. 2008).

A possible improvement to the differential diagnosis would be the assessment of semantic cognition, according to its very specific impairment in AD and aMCI (Brambati et al. 2009, 2012; Brunet et al. 2011; Callahan et al. 2015; Joubert et al. 2020; Simoes Loureiro and Lefebvre 2016a, b).

Semantic cognition enables the retrieval and use of our general knowledge of the world by combing activation in the semantic representation system (Binder et al. 2009; Hoffman 2018; Jefferies et al. 2020), an executive processes allowing the controlled retrieval of less salient semantic information (Hoffman 2018) and the selection or inhibition of task-relevant aspects of semantic knowledge (Chiou et al. 2018; Hoffman 2018; Joyal et al. 2020). The activation of the semantic representation system is an automatic process almost sufficient for ongoing tasks requiring dominant or highly accessible semantic information (Chiou et al. 2018; Jefferies et al. 2020; Teige et al. 2018). However, the executive contribution to semantic retrieval includes more controlled processes that are increasingly needed when less easily accessed aspects of knowledge are required (Chiou et al. 2018; Jefferies et al. 2020; Teige et al. 2018).

As these elements of semantic cognition rely on distinct but interacting brain regions (Copland et al. 2007; Joyal et al. 2020; Raucher-Chene et al. 2018, 2017), the full processing of semantic cognition can be almost equally impaired by a broad number of neuropathological conditions, while the separate subprocesses should be more specifically impaired by some conditions and not others.

First, the activation of the semantic representation system (Binder et al. 2009; Hoffman 2018; Jefferies et al. 2020) is expected as being affected only by AD (or aMCI). Indeed, this activation relies on the anterior temporal cortices (Binder and Desai 2011; Binder et al. 2009; Jefferies et al. 2020; Sami et al. 2018; Venneri et al. 2008) and parahippocampal regions damaged by the neuropathological injuries of AD (Venneri et al. 2008). For a review of neuroanatomical correlates of semantic impairment in AD, refers to Venneri et al. (2008).

Secondly, the executive processes allowing semantic retrieval are associated with a network comprising the inferior prefrontal cortex, posterior middle temporal gyrus, and

interparietal sulcus (Binder et al. 2009; Hoffman 2018; Venneri et al. 2008). Many psychopathological or neuropathological diseases including LLD and AD can damage one of these brain regions. Consequently to the neural differentiation, when it comes to the impairment of these subprocesses by LLD and AD, one must expect that both activation of the semantic representation system (concepts and links) (Laisney et al. 2009, 2011; Simoes Loureiro and Lefebvre 2016a, b) and executive processes of semantic retrieval will be impaired in AD, while only executive processes will be touched by LLD. Indeed, the regions involved in the activation of the semantic representation system, namely, the anterior temporal cortices (Binder and Desai 2011; Binder et al. 2009; Jefferies et al. 2020; Sami et al. 2018; Venneri et al. 2008) and parahippocampal regions, have not been reported to be affected by the neuropathological consequences of LLD. This leads to the expectation that if LLD affects semantic retrieval competencies it is most likely related to the dysexecutive consequences of the loss of volume of the superior frontal gyrus and ventromedial frontal cortex (Alexopoulos et al. 2005; Boccia et al. 2015; Morimoto et al. 2011). It is admitted that LLD affects the domain-general executive selection system involved in inhibition tasks, such as the Stroop test (Hoffman 2018), and that impaired results in recall tasks in LLD rely on impairment of the semantic executive (clustering) implied in the encoding phase of the task (Elderkin-Thompson et al. 2006; Lamar et al. 2012).

This opportunity to discriminate the semantic impairment due to AD or LLD then requires investigating in which manner these two pathological situations' cognitive impairments are reflected in the performance of semantic-based neuropsychological tasks. Following the above reasoning, it becomes relevant to consider the use of these neuropsychological tests in medical practice. To do so, let us consider two different types of tests that involve mainly, even if among other cognitive resources, semantic knowledge while attempting to assume that certain tasks can be completed with minimal contribution from executive semantics, while others are highly dependent on it.

In this systematic review of literature, we will consider that semantic cognition can therefore be assessed with two main groups of neuropsychological tests, which require more or less one of both semantic retrieval subprocesses: activation in the semantic representation system (Binder et al. 2009; Hoffman 2018; Jefferies et al. 2020), and executive processes.

The first group of tasks involves both the activation of the semantic representation system and an important contribution of executive processes. It includes mainly confrontation naming and verbal fluency tasks. Confrontation naming (e.g., Boston Naming Test (Kaplan et al. 2001), DO 80, or LEXIS denomination subtest (de Partz de Courtray et al. 2001)) is a task requiring the name of a given image and

assessing the capacity to retrieve the meaning of a concept and its lexical label. Recognition via the activation of the semantic representation system also requires an executive contribution to actively retrieve the lexical label for the verbal production of the answer (Higby et al. 2019). However, in neuroimaging research seeking the correlations between brain region disruptions in AD and impaired results in confrontation naming, Domoto-Reilly et al. (2012) demonstrated that, even if a general relationship could be demonstrated with a region of interest distributed between the frontal, parietal, and temporal cortices, thinning of the left anterior temporal lobe (directly linked to semantic representation system) was more highly correlated with impaired naming performance (Domoto-Reilly et al. 2012).

The verbal fluency task, which is almost always included in general cognitive assessments, is auto-initiated, sometimes classified as a semantic measure or a language measure, or a measure of executive functioning. In fact verbal fluency is all of these, because to perform it (to produce the maximum of words according to a given rule), either phonemic (all words have to start with the same letter) or semantic (all words have to belong to a given category), one needs to activate the lexicosemantic network (by activation of the semantic representation system), but also to maintain and refresh information in the working memory with an important contribution from the executive functions (Henry and Crawford 2005, 2004). The participant has to self-initiate the retrieval, keep track of the responses already given, and inhibit responses that are inappropriate and switch to another group of concepts when the first one is exhausted (e.g., in semantic fluency "animals," switching from house pets to farm animals). Despite this multi-determined aspect of fluency, at bases, it is first of all impossible to perform without functional access to the semantic representation system. It has been extensively studied in healthy elderly individuals and patients with AD through various meta-analyses (Henry and Crawford 2004; Laws et al. 2007, 2010). Regarding the healthy elderly, it is assumed that they have better verbal fluency performance in the semantic task than in the phonemic task, and that this advantage persists over time (Charles et al. 2019; Vaughan et al. 2016). Regarding AD, two meta-analyses (Henry and Crawford 2005, Laws et al. 2007) showed that the effect for both AD (Laws et al. 2007) and LLD (Henry and Crawford 2005), the deleterious effect on semantic fluency was significantly larger than on phonemic fluency (and confrontation naming) which is the inverse result of healthy aging participants. Notice that the samples in Henry and Crawford's research were composed of adults and not the elderly (age: 52.7 ± 13.69) (Henry and Crawford 2005) and that the authors have shown that the contradictory results are essentially the consequence of methodological artifacts (highly influenced by intra-individual variables such as education or lexical background) and may partly reflect

executive dysfunction, but also be a consequence of more generalized cognitive impairment (Henry and Crawford 2005). The effect of LLD on the results of verbal fluency and confrontation naming is however still unclear (Balsamo et al. 2018; Henry and Crawford 2005; Lee et al. 2012).

The second group includes tasks that allow to reduce of as much as possible cognitive (and especially executive) contributions, other than activation of the semantic system. It regroups the tasks that are not auto-initiated. In this category are semantic pairing and classification, designation and recognition tasks, lexical decision tasks, semantic knowledge, and vocabulary descriptions. Semantic pairing tasks require pairing images based on their semantic relationships (e.g., Pyramid and Palm Tree Test (Klein and Buchanan 2009), Camel and Cactus Test (Adlam et al. 2010)) while in classification the participant must classify images among superordinated categories (e.g., classification in Cambridge Semantic Memory Test Battery (Adlam et al. 2010) or LEXIS subtest (de Partz de Courtray et al. 2001), Size Weight Attribute Test (Warrington and Crutch 2007)). A particular type of recognition task is the lexical decision tasks (e.g., Spot-the-word (Baddeley et al. 1993)) in which participants see pairs of items comprising one word and one non-word and must identify the word. Yuspeh and Vanderploeg (2000) suggested that spot-the-word is a useful additional measure to estimate premorbid intelligence (Yuspeh and Vanderploeg 2000). When used as a semantic assessment, this usefulness can be argued against by the fact that in AD, patients are able to discriminate a word from a non-word, even if the meaning of the word is lost.

Semantic knowledge questionnaires consist of multiple choices about concepts, either about their subordinated category or their features (e.g., the Semantic Knowledge Questionnaire (Simoes Loureiro and Lefebvre 2015)) and can also contain questions about famous persons or famous public events. Last of this group are the vocabulary (e.g., Mill Hill Vocabulary Scale (Colman 2009)) and similitude tasks, often part of the verbal subtests in general intelligence scales (e.g., Wechsler Adult Intelligence Scale (Hartman 2009)). Note that in the literature, including research presenting a general cognitive assessment, some studies present the same tests exposed in this paragraph, not as part of a semantic, but as part of a language variable (Butters et al. 2004; Hoffman 2018; Yuspeh and Vanderploeg 2000). The responses for the tasks of this first group are demonstrated as being affected by the early stages of AD (Adlam et al. 2010; Croisile et al. 2011; Hernández et al. 2008; Perri et al. 2019, 2012; Zanino et al. 2015) and MCI (Belanger and Belleville 2009; Joubert et al. 2020; McLaughlin et al. 2014) while the issue of their impairment in LLD is rarely examined.

Given this information, several questions arise. Firstly, is it possible to detect an effect of LLD on the scores on the semantic evaluation tests? Secondly, is it possible to

discriminate between an effect of LLD on the tests that are included in the first and second groups? Finally, when studies compare LLD to AD (or aMCI), will the effect on semantic test scores be significantly different between the two conditions? To answer the first question, we systematically searched studies enabling us to identify outcomes at the semantic assessments of groups of patients with LLD, compared to elderly adults without cognitive impairment. To answer the second question, the results of these tests have been analyzed comparatively according to their belonging to the first or second group. To answer the last question, we also searched the literature for articles that compared the results of groups of LLD to those of groups of AD (or aMCI) within the same research, in order to see if the effect of AD (or aMCI) on these results was significantly greater than that of LLD.

Method

This systematic review was conducted following the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Liberati et al. 2009) and the guidelines for the contents of a Cochrane diagnostic test accuracy protocol (Deeks et al. 2013). The final protocol of the systematic review was submitted for registration at the PROSPERO International Prospective Register of Systematic Reviews on July 13, 2021 (register code CRD42021266476). The management of methodological steps was performed using Excel and EndNote. Intermediate tables are provided upon request.

We applied two different research strategies. The first search strategy was designed to find research enabling to identification outcomes at the semantic assessments of groups of patients with LLD, compared to elderly adults without cognitive impairment or depressive symptoms. The research identified with this strategy were included to answer the two first questions; is it possible to detect an effect of LLD on the scores on the semantic evaluation tests, and is it possible to discriminate between an effect of LLD on the tests that are included in the first and second groups? The second research strategy allowed us to identify studies that compare LLD to AD (or aMCI) to see if the effect on semantic test scores is significantly different between the two conditions.

The accepted format was non-interventional psychophysiological research, written in English or French, including case-control and cross-sectional studies. Longitudinal or retrospective studies were also included if they allowed for the extraction of a neuropsychological measurement of the groups when they were over 60 years old. To be included, the research must provide at least one measure of semantic or semantic-executive tests for at least two groups: LLD and

control, LLD and AD, or LLD and MCI. Note that the presence of data regarding LLD was the main inclusion criterion. Research with results of AD (or aMCI) were included only if they allowed a direct comparison with a group with LLD, otherwise they were not included (and will be the object of a different systematic review by the same research team). Were excluded or the review articles that (1) did not use a semantical assessment or used a semantical assessment not validated by a previous publication, (2) allowed only the extraction of data for a population under 60 years old, (3) did not include a possibility of relating semantical data with a LLD group, (4) included only an AD or aMCI population without an LLD group, (5) included only a group of LLD without a control group, (6) were meta-analyses or systematic reviews.

As reported in the flow chart (Fig. 1), at the end of the selection procedure, 55 articles were selected by two independent researchers who evaluated the quality of the selected research based on the criteria decided in the research protocol registered on PROSPERO. The quality assessment frame was created by the research team, and quality criteria were determined to identify the risks of selection and diagnostic bias, the possible presence of confounders for data, and the accuracy of data collection and statistical analysis. Studies were excluded from the review because they did not meet the criteria for individual selection or were not representative of the target population. Twenty-four studies were removed because they did not meet one of the quality criteria (Fig. 1), and a systematic review of the content was conducted on 31 articles.

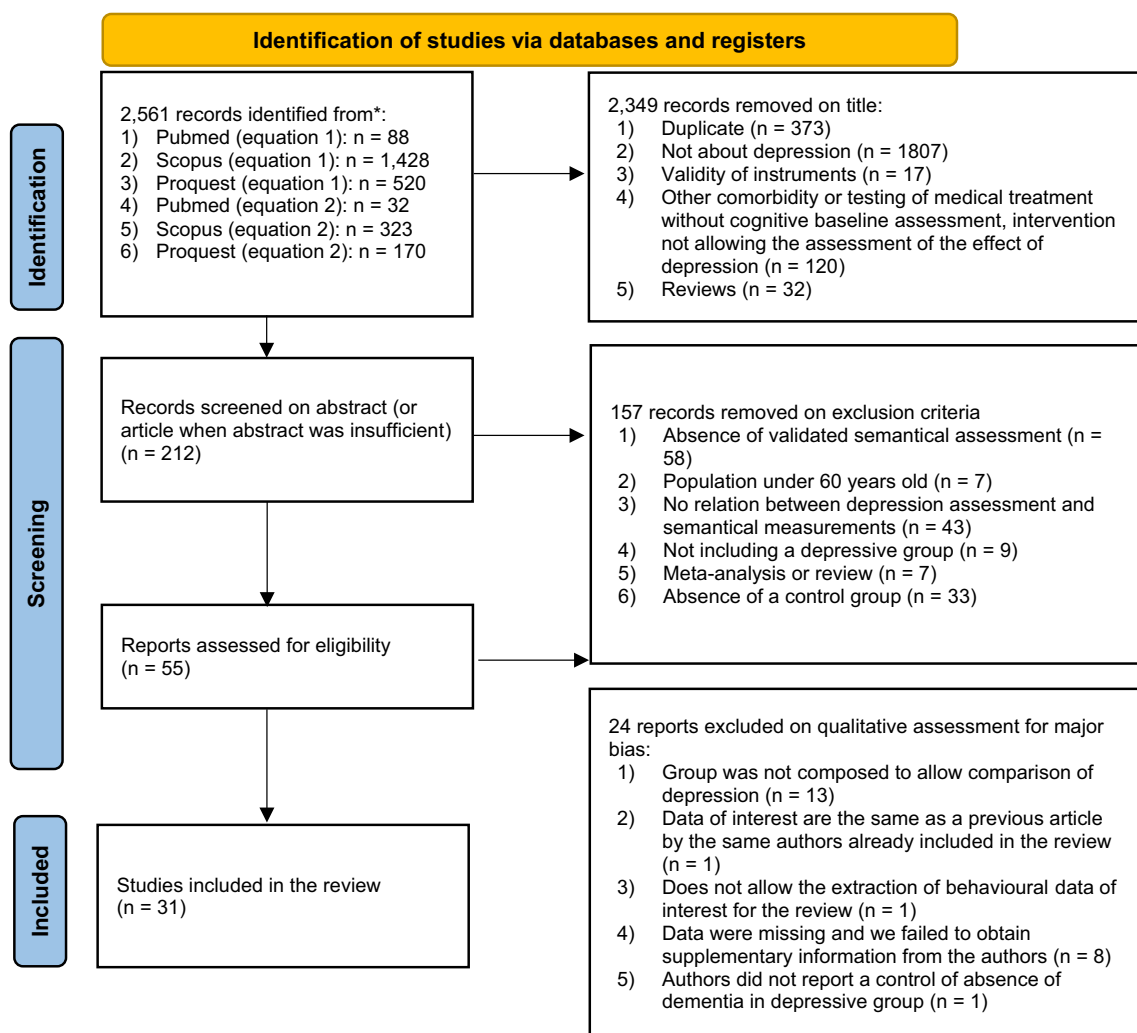


Fig. 1 PRISMA Flow diagram of study screening and selection. *Note.* The first round of selection was performed at the beginning of the systematic review process in August 2021 with no limitation in years

of publication. A second round of research was performed at the end of the systematic review process in June 2022 with a limitation on articles published in 2021 or 2022

Data extraction and analysis

R scripts and detailed databases are accessible on Open Science Framework (OSF): https://osf.io/evrwf/?view_only=935fb2bcf5aa487f90ec22a034835b0a

We systematically extracted from the 31 articles (presented in Table 1) the number of participants of each group and, in means and standard deviations, the ages, degree of depressive symptoms, quality of depression, level of education and information about the diagnosis of depression. When the studies included only groups with LLD (and no groups of AD and aMCI), we reported the information about the control of non-dementia. When the studies also included groups with aMCI or AD, we reported information about the quality of these diagnosis. The complete table of data extracted from the articles are available on OSF. From these 31 studies we extracted 52 groups comparisons; 31 comparisons between LLD and control groups, 7 comparisons between aMCI and LLD groups, and 4 comparisons between AD and LLD groups. Were compared results at the tests of semantic ($n=50$) and phonemic fluency ($n=36$), and confrontation naming ($n=30$), corresponding to the first group of tasks that involves both the activation of the semantic representation system and an important contribution of executive processes. For the second group of tasks that rely more directly to the activation of the semantic system, we extracted the subtest of the language of the ACE-R (Mioshi et al. 2006) ($n=3$), the Pyramid and Palm Tree Test (Klein and Buchanan 2009) ($n=9$), the spot-the-word (Yuspeh and Vanderploeg 2000) ($n=3$), the similarities subtest of the *Wechsler Adult Intelligence Scale* (WAIS) ^[65] ($n=2$), vocabulary subtest of the Wechsler Abbreviated Scale of Intelligence (WASI) (McCrimmon and Smith 2013) ($n=3$), and the vocabulary subtest of the WAIS (Hartman 2009) ($n=1$). All the tests of the second group were aggregated together in the same analysis of the gathered effect.

Statistical analysis

Meta-analysis of the results was conducted using the metafor package (Viechtbauer 2010) with a random effects model using the *rma* function. The preferred choice for the random effects meta-analytic model was decided due to the important differences in sample sizes between studies and between groups within studies, which allowed us to assume random variation in the effect of interests among studies. Moreover, Viechtbauer (2010) recommended the choice of a random effects model to correct a plausible selection bias, which is the consequence of the absence of unpublished studies in our selection of research (Viechtbauer 2010). Unlike the fixed-effects model, which provides an inference only suitable for the sample of selected studies, the random effects model provides an inference about the average effect in the

entire population of studies from which the included studies are assumed to be a random selection (Viechtbauer 2010).

The observed measure is the *weighted average effect*, expressed with the Hedge's g (Hedges 1981) and its derived prediction interval (95% *IC*). This result is reported with the p -value of significance calculated using a Z score. The 95% prediction interval provides the range in which the point estimate of 95% of future studies will fall, assuming that true effect sizes are normally distributed throughout the domain (Borenstein et al. 2009). The heterogeneity between studies included in one calculation of the combined effect was reported using the I^2 value of heterogeneity (Higgins and Thompson 2002). This value reports the amount of variability that cannot be justified by the sampling differences between the studies. According to the rule of thumb, the heterogeneity is considered as low ($I^2=25\%$), moderate ($I^2=50\%$) or substantial ($I^2=75\%$).

Since sampling variation is important between and within studies, the combined effect size was calculated with the SMDH function for the *standardized mean difference* with heteroscedastic population variances in the two groups (Bonett 2009). The effect sizes are reported in Tables 2, 3, 4 and 5. For the fluency and naming tasks, a single effect size on each test was calculated. For the tests of group two, given their small number, a pooled effect size was measured for all the tests. In Table 2 and 3 the effects of LLD compared to a control group are reported, and in Tables 4 and 5 the differences in effects between the depressed groups and the groups with either aMCI or AD are reported. As explained above, this research did not include direct comparisons between aMCI or AD and control groups, as the aim of this research was either to measure the effect of LLD or to identify a direct comparison between the conditions (LLD vs aMCI/AD).

Due to the high level of heterogeneity in the comparisons between LLD and control groups, we used a model of linear regression to seek the influence of the side-variables age, level of depressive symptoms and level of education. To do so, we calculated three ratio indexes (y_{Age} , y_{Edu} and y_{DS}) as the proportion between the value in the control group minus the value in the experimental group, divided by the value in control group. Analysis were run with the *lmer* (Kuznetsova et al. 2015) and the *afex* (Singmann et al. 2015) packages. We introduced in the model the medium effect size of depression on the results at the tests (y_i calculated with the SMDH function) as the dependent variable; the three ratio index (y_{Age} , y_{Edu} and y_{DS}) as fixed independent variables, and the specificity of LLD group (e.g., Major, minor, late onset, persistent, remittent, insistent, subsyndromal, dysthymia) as random effects. First, the structure of the mixed-effect modelling was determined by structuring a full model that included the fixed effects of interest and the random effects. The full model was then challenged with an ANOVA-like table for tests of random

Table 1 Thirty-one review articles with first author, year of publication, group characteristics, type of results reported, reference for the diagnostic of depression, and scale of depressive symptoms

First author, year	Groups	Diagnostic criteria LLD	Dep. scale	Extracted tests	RB1	RB2	RB3
Ravdin et al. (2003)	LLD moderate	Symptoms	GDS 30	Semantic fluency; phonemic fluency	*	*	
Butters (2004)	LLD	DSM IV	SCID	Semantic fluency; pho- nemic fluency; nam- ing; spot-the-word			
Elderkin-Thompson (2006)	LLD	DSM IV	HDRS	Semantic clustering in recall			*
Mah (2017)	LLD	DSM V	GDS 15	Semantic fluency; pho- nemic fluency; nam- ing; Wais vocabulary			
Fischer (2008)	LLD	DSM IV	GDS 30	Semantic fluency; phonemic fluency; similarities; Wais similitudes			
Avila (2009)	LLD (high—low edu- cation)	DSM IV	HDRS	Semantic fluency; phonemic fluency			
Dillon (2009)	LLD early onset LLD late onset	DSM IV	HDRS	Semantic fluency; phonemic fluency; naming			
Barabassy (2010)	LLD	DSM IV	MADRS	Semantic fluency; naming			*
Parra et al. (2010)	LLD AD	DSM IV	GDS 30	Semantic fluency; phonemic fluency	*	*	
Brunet (2011)	LLD aMCI aMCI D+	DSM IV	GDS 30	Semantic fluency; phonemic fluency; naming; PPTT			
Dillon (2011)	Dysthymia Subsynd. LLD LLD due to AD	DSM IV	HDRS	Semantic fluency; phonemic fluency; naming			
Elderkin-Thompson (2011)	LLD	DSM IV	HDRS	Semantic fluency; phonemic fluency			
Jungwirth (2011)	LLD	DSM IV	HDRS 17	Semantic fluency			
Alves (2012)	LLD	DSM IV	GDS 15	Semantic fluency; naming			
Lim et al. (2012)	LLD	DSM-IV TR	HDRS 17	Semantic fluency; naming			
Tam and Lam (2012)	LLD MCI	DSM IV	HDRS 24	Semantic fluency			*
Callahan (2015)	LLD aMCI aMCI D+	DSM IV	GDS 30	Semantic fluency; phonemic fluency; naming; PPTT			
Koenig (2015)	Euthymic	DSM IV	HDRS 17	Semantic fluency; pho- nemic fluency; nam- ing; spot-the-word			
Rotomski (2015)	LLD						

Table 1 (continued)

First author, year	Groups	Diagnostic criteria LLD	Dep. scale	Extracted tests	RB1	RB2	RB3
	AD	ICD 10	not referred	Semantic fluency phonemic fluency subtest language ACE R	*	*	
Beckert (2016)	LLD	DSM IV	GDS 15	Semantic fluency; phonemic fluency; naming; subtest language ACE R			*
Callahan (2016)	LLD aMCI aMCI D+	DSM IV-R	GDS 30	Semantic fluency; phonemic fluency; naming; PPTT			
da Costa Dias (2017)	LLD	DSM IV	GDS 15	Semantic fluency; naming			
Rajtar-Zembati (2017)	LLD	DSM V	GDS 15	Semantic fluency; phonemic fluency			*
Esteves (2019)	LLD	Symptoms	GDS 15	Semantic fluency; phonemic fluency; naming	*		*
Morin et al. (2019)	LLD	DSM IV	HDRS 24	Semantic fluency; phonemic fluency; naming; Wais similarities			
Olaya et al. (2019)	LLD lifetime LLD remittent LLD incident LLD persistent	DSM IV	SCID	Semantic fluency	*	*	*
Lin et al. (2020)	LLD	DSM V	GDS 15	Semantic fluency			
Muniswamy et al. (2020)	LLD	DSM IV R	GDS 30	Semantic fluency			*
Faoro (2021)	LLD	Symptoms	GDS 30	Semantic fluency; naming; wais vocabulary	*		
Ramos-Henderson et al. (2021)	LLD	Symptoms	GDS 15	Semantic fluency; phonemic fluency	*	*	
Liu et al. (2022)	LLD early onset LLD late onset	DSM IV	GDS 15	Phonemic fluency; naming			

Note. Groups: identification of the pathological groups for which data were extracted in the meta-analysis. Diagnostic criteria LLD: referred criteria of diagnostic for Major Depression; Diagnostic and Statistical Manual of Mental Disorders (DMS), version 4 (IV), 4 revised (IV R) and 5 (V); International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD 10), or not stated, but only referred to according to the threshold of significance of the depressive scale (Symptoms). Dep. scale represents the depression scales reported by the authors to measure the level of depressive symptoms, including the geriatric depression scale (GDS), the Hamilton depression rating scale (HDRS), the Montgomery-Åsberg Depression Rating Scale (MADRS) and the Structured Clinical Interview for DSM Disorders (SCID). Risks of bias are indications of the fact that researches do not entirely respect the non-exclusion criteria of quality of the research; (1) do not state the precise condition of LLD diagnostic of the level of depressive symptoms of the pathological group; (2) state control of non-dementia but do not provide the precise criteria of control; (3) data about the results of the control group were missing and managed by statistical procedure

Table 2 Standardized mean difference between Control and LLD groups on the results of the tasks

41 studies including LLD and control						
	Control			LLD		
N	5972			2561		
Age	69.57 (6.52)			69.48 (7.06)		
Education	11.46 (3.48)			10.53 (3.85)		
Dep. Sympt	2.67 (2.37)			13.26 (4.8)		
		Degree of freedom	Effect size Hedge's g (95% CI)	p-value	Heterogeneity (I^2)	
Phon. fluency (high-exe.)	27.44 (6.9)	24.55 (7.28)	23	0.46 (0.28 to 0.63)	0.001	81.03%
Sem. fluency (high-exe.)	22.34 (5.5)	20.03 (5.84)	34	0.46 (0.32 to 0.57)	0.001	79.90%
Naming (high-exe.)	36.74 (4.25)	32.92 (6.46)	21	0.54 (0.36 to 0.73)	0.001	79.69%
PPTT (low-exe.)	42.77 (1.91)	42.77 (2.84)	2	0.23 (-0.17 to 0.63)	ns	0%
Spot-the-word (low-exe.)	50.37 (7.66)	47.46 (9.24)	3	0.13 (-0.08 to 0.42)	0.03	66.52%
Vocabulary (low-exe.)	46.43 (11.46)	41.87 (9.51)	3	0.44 (0.14 to 0.73)	ns	0%
Similarities (low-exe.)	16.35 (2.99)	16.84 (2.54)	1	0.24 (-1.13 to 0.64)	0.02	81.43%
ACE-R Language (low-exe.)	23.56 (2.17)	23.21 (2.73)	0	0.14 (-0.12 to 0.41)	ns	0%
Semantic cluster (low-exe.)	1.70 (0.90)	1.25 (1.15)	1	0.45 (0.18 to 0.73)	ns	0%
pooled low-exe. tasks	39.91 (6.90)	37.88 (6.70)	13	0.20 (0.05 to 0.34)	0.01	44.32%

Table 3 Linear regression models for prediction of the average effect of LLD on the tasks

Average effect of LLD on:	Fitted regression model	Significance of overall regression
Semantic fluency	$= 0.48 + 0.65 * yEdu - 0.006 * yDS + 2.54 * yAge$	$R^2 = 0.02, F(3.42) = 0.35, p = 0.78$ (ns)
Phonemic fluency	$= 0.32 + 1.66 * yEdu - 0.84 * yDS + 0.24 * yAge$	$R^2 = 0.27, F(3.28) = 3.37, p = 0.03$
Naming	$= 0.48 + 1.80 * yEdu + 0.02 * yDS - 6.90 * yAge$	$R^2 = 0.33, F(3.26) = 4.20, p = 0.01$
Pooled low-executive tasks	$= 0.21 + 0.46 * yEdu - 0.03 * yDS - 5.32 * yAge$	$R^2 = 0.11, F(3.17) = 0.71, p = 0.55$ (ns)

This table report the results of the multiple linear regression used to test if the difference between LLD and control groups regarding the level of education (yEdu), of depressive symptoms (yDS) and mean age (yAge) significantly predict the average effect of LLD on the tasks of semantic fluency, phonemic fluency, naming and the pooled effect of the low-executive tasks. The fitted models were only significant for phonemic fluency and naming. (ns) indicates a p-value non significant.

effect terms in the model and simplified until all the remaining intercepts and slopes of random effects were considered as significantly useful for the model.

Results

Effect of LLD on high- and low-executive semantic tasks

The results in Table 2 show the pooled effects of depression compared to control participants of the same average age and education level. We see that scores on the fluency and naming tasks are significantly affected by depression, but in each case with substantial heterogeneity between studies, which cannot be due to differences in sample sizes within each study.

We see in the same table that the tasks in low-executive group are also affected by depression, but with a smaller effect size, however, related to studies with less heterogeneity ($I^2 = 44.32%$). The detail of this pooled result for the low-executive tasks should, however, be considered with caution, as it is dependent on very small numbers of research. When examined in detail, we see that several of these low-executive tasks are not significantly affected by depression, such as the PPTT, the vocabulary of the intelligence scales, the language sub-score of the ACE-R and the semantic cluster in a free recall task.

To investigate the possible sources of heterogeneity across studies, we performed linear regressions in which the dependent variable was the effect size of depression on the fluency, semantic and phonemic, naming and low-executive tasks, and the independent variables were the ratios calculated to represent the magnitude of the differences in

Table 4 Standardized mean difference between control and LLD groups, and between LLD and aMCI groups on the results of the tasks

7 studies including LLD and aMCI					Control	LLD	aMCI		
N					375	73			142
Age					69.84 (6.21)	71.47 (9.36)			70.55 (7.07)
Education					12.16 (3.42)	11.31 (3.68)			11.24 (3.70)
Dep. Sympt					3.11 (2.76)	12.31 (4.81)			10.49 (5.37)
High-executive tasks:						Degree of freedom	Hedge's g (95% CI)	p-value	Heterogeneity (I^2)
Phonemic Fluency	28.25 (8)	23.88 (9.03)	23.28 (9.28)	Ctrl//aMCI	6	0.52 (0.28 to 0.76)	0.001	0%	
				Ctrl//LLD	6	0.42 (0.18 to 0.66)	0.02	0%	
				LLD//aMCI	6	-0.01 (0.24 to -0.76)	ns	0%	
Semantic Fluency	14.7 (5.79)	12.67 (6.53)	10.25 (3.16)	Ctrl//aMCI	6	0.97 (0.62 to 1.33)	0.001	51.14%	
				Ctrl//LLD	6	0.39 (0.14 to 0.64)	0.001	0%	
				LLD//aMCI	6	0.29 (0.04 to 0.55)	0.02	6.42%	
Naming	16.06 (1.48)	14.29 (3.25)	15.09 (1.63)	Ctrl//aMCI	6	0.49 (0.26 to 0.73)	0.001	0%	
				Ctrl//LLD	6	0.43 (0.19 to 0.67)	0.001	0%	
				LLD//aMCI	6	0.04 (0.45 to -0.36)	ns	61.82%	
Low-executive tasks:									
PPTT	49.88 (1.83)	48.87 (2.64)	47.88 (2.35)	Ctrl//aMCI	5	0.65 (0.10 to 1.21)	0.03	78.25%	
				Ctrl//LLD	5	0.21 (-0.04 to 0.47)	ns	0%	
				LLD//aMCI	5	-0.28 (-0.73 to 0.15)	ns	61.47%	
Vocabulary	56.87 (6.01)	49.23	10.20	53.73 (5.01)	Ctrl//aMCI	0	0.55 (-0.13 to 1.26)	ns	-
					Ctrl//LLD	0	0.89 (0.15 to 1.92)	0.01	
					LLD//aMCI	0	0.54 (-0.16 to 1.26)	ns	
Pooled low-exe. tasks	50.88 (2.24)	48.92 (3.72)	48.72 (2.73)	Ctrl//aMCI	6	0.64 (0.16 to 1.11)	0.008	73.57%	
				Ctrl//LLD	6	0.28 (0.05 to 0.52)	0.02	0%	
				LLD//aMCI	6	-0.17 (-0.79 to 0.26)	ns	65.51%	

age, education and depression level between the groups. The larger these ratios are, the greater the difference in the target data (e.g., education level) between the LLD group and the control group within the same study. We also ran these regressions in mixed fixed-effect, random effect models, where the random effect was the specific type of depression reported in each study (late onset, major, minor, subsyndromal, persistent, incident, remittent, lifetime, dysthymic, euthymic). For each mixed linear regression (semantic fluency, phonemic fluency, naming, low-executive), a maximum likelihood ratio analysis showed that the model including this random effect of the specific type of depression

did not better explain the variance of the effect sizes than a model including only the fixed effects of the ratios age, education level and level of depressive symptoms. Therefore, the presented results are the outcomes of the linear regressions including only the fixed effects.

The outcome of these linear regressions, presented in Table 3, shows that the models are only significant in predicting the effects of LLD on phonemic fluency and naming tasks. When we observe the levels on these linear regression, we can see that the variable yEdu (difference in education level between the LLD group and the control group) significantly predicts the effects on the

Table 5 Standardized mean difference between control and LLD groups, and between LLD and AD groups on the results of the tasks

4 studies including LLD and AD									
		Control			LLD			AD	
N		287			264			292	
Age		68.96 (6.54)			69.83 (7.85)			71.93 (6.89)	
Education		10.57 (3.17)			9.60 (4.18)			11.15 (3.41)	
Depr. Symptoms		1.90 (1.70)			17.18 (6.36)			16.60 (5.80)	
High-executive tasks:					Degree of freedom	Hedge's g (95% C.I.)	p-value	Heterogeneity (I^2)	
Phonemic fluency	10.36 (2.52)	7.07 (3)	5.09 (2.46)	Ctrl//AD	2	1.76 (0.88 to 1.46)	0.001	0%	
				Ctrl//LLD	2	1.23 (0.97 to 1.50)	0.001	0%	
				LLD//AD	2	-0.002 (-0.27 to 0.27)	ns	10.32%	
Semantic fluency	12.15 (3.65)	9.19 (2.65)	6.14 (2.47)	Ctrl//AD	2	1.65 (1.26 to 2.04)	0.001	24.91%	
				Ctrl//LLD	2	0.55 (-0.55 to 1.65)	ns	92.05%	
				LLD//AD	2	1.37 (0.89 to 1.85)	0.001	50.22.%	
Naming	51.60 (4)	44.50 (7.20)	37.40 (10.10)	Ctrl//AD	0	1.81 (0.98 to 2.46)	0.001	/	
				Ctrl//LLD	0	1.20 (0.62 to 1.78)	0.001		
				LLD//AD	0	0.79 (0.14 to 1.44)	0.01		
Low-executive tasks:									
ACE-R language	23.56 (2.17)	23.21 (2.73)	17.21 (4.04)	Ctrl//AD	0	1.94 (1.57 to 2.33)	0.001	/	
				Ctrl//LLD	0	0.14 (-0.12 to 0.41)	ns		
				LLD//AD	0	1.73 (1.38 to 2.08)	0.001		

phonemic fluency task ($\beta = 1.66, p = 0.02$) and the naming task ($\beta = 1.80, p = 0.04$), while the variable yAge significantly predict the effects on the naming task ($\beta = -6.90, p = 0.02$). Ratios related to level of depressive symptoms did not predict these results.

Incremental effect of aMCI and AD on high- and low-executive tasks

Tables 4 and 5 show the results of the averaged effects when, in the same research, LLD groups were compared with both control and pathological groups suffering from either aMCI (Table 4) or AD (Table 5). The outcomes of Table 4 show that aMCI has no additional effect on the scores of the different tasks compared to the effect of LLD on these tasks. The only notable exception is semantic fluency, for which we see a significant medium-sized effect of the aMCI compared to the LLD. We interpret these results as showing that, apart from semantic fluency, none of the listed tasks show a different effect of the aMCI on individuals' performance or allow them to be distinguished from LLD results.

Table 5 shows that, except the phonemic fluency task, all the tasks considered show a significant effect of AD on the results when compared to LLD results. Our interpretation of this result is that phonemic fluency is not a task that shows

a significant difference in results between LLD and AD, but that the semantic fluency, naming and language subtest of the ACE-R do.

Comparison of results on high-executive tasks and methodological differences in studies

The aggregate means show that scores at phonemic, semantic fluency and naming are affected by the LLD status. However, for semantic fluency, if we compare aMCI groups to LLD groups, the aMCI has a significant incremental effect on the impairment of outcomes. This additive effect of the aMCI on LLD scores is not true for the phonemic fluency and naming tasks. This means that when the groups are tested with these three tasks, semantic fluency will show a marked difference between aMCI and LLD scores, while phonemic fluency and naming will produce scores that are equivalently impacted by both the aMCI and LLD. When comparing AD groups to LLD groups, AD has a significant incremental effect on semantic fluency and naming scores impairment, but not on phonemic fluency. In other words, in the identified studies, the semantic fluency or naming task significantly discriminated between AD and LLD scores but the phonemic fluency task did not.

The results on semantic fluency are coherent with a previous meta-analysis conducted by Henry and Crawford (2004), which showed that the adverse effect of AD on the performance of these tasks was greater than that of depression, although the meta-analysis included younger adults in the depressed groups, whereas our purpose was to assess the specificity of age in LLD.

We identified substantial heterogeneity between the studies included in the calculation of the pooled effect of LLD (when compared to the control group) for the results of the semantic, phonemic fluency and naming tasks. The linear regression model that explored the effect of age, level of depressive symptoms and level of education on these scores showed that level of education predicts the effect of depression on scores on these three tasks. To do this, we calculated ratios representing the magnitude of the difference in these variables between the LLD group and the control group for each study and introduced these ratios as variables in a linear regression to see which one best predicted the effect sizes of LLD on outcomes. Only the ratio of the difference in education level showed a significant effect in all three models.

This influence of education level on scores at semantic fluency was not true for MCI and AD groups. This may be explained by the fact that the overall effect of LLD is much lower and is more sensitive to secondary variables. It is worth noting that we also considered whether the different types of depression identified in the research could explain the substantial heterogeneity between their scores on these tasks. However, when tested by introducing this factor

as a random variable in the linear regression, no statistical significance was found. The distribution of effect sizes across studies is not better explained by the variety of types of depression reported.

However, despite the lack of proof by statistical significance, it is interesting to describe some methodological differences between the studies and to view them regarding the forest plot representing the distribution of LLD effects (compared to control groups) for the semantic fluency (Fig. 2), phonemic fluency (Fig. 3) and naming tasks (Fig. 4).

The studies that reported the most important effect of LLD on both fluency and naming are Dillon et al. (2009, 2011). The first study (Dillon et al. 2009), controlled the age at the first onset of depression between late- (> 60 years old) and early-onset (< 60 years old) depression (Dillon et al. 2009) was controlled. In the subsequent study (Dillon et al. 2011), they also controlled the variation between types of depression between groups with major depression, dysthymia, subsyndromal depression, and depression due to mild Alzheimer’s dementia (Dillon et al. 2011).

While in the studies by Dillon et al. all different groups of LLD had a major impairment on these three tasks, in Koenig et al. (2015), the distinction between two groups of depressed patients, differing in the severity of depression between euthymic and major depression, was not reflected in results of dramatically different magnitudes.

If we detail four other studies that examined the effect on semantic fluency, phonemic fluency and naming (Brunet et al. 2011; Butters et al. 2004; Callahan et al. 2015;

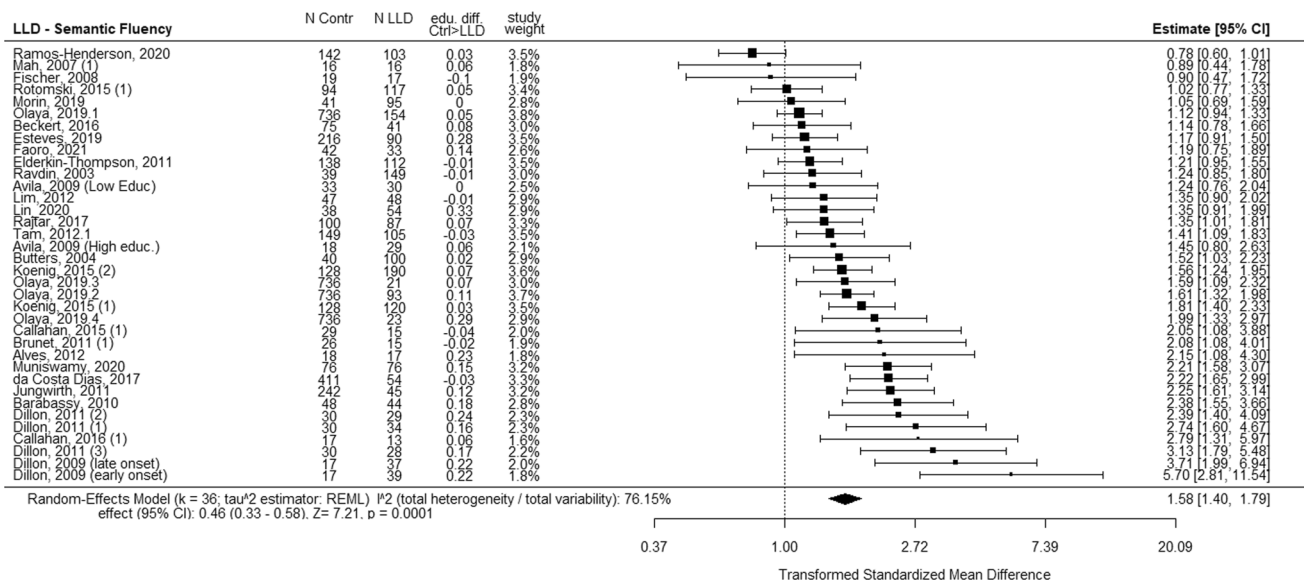


Fig. 2 Forest plot of average effects of LLD (compared to the control group) for outcomes of semantic fluency, including differences in education. Note: when a study includes several LLD groups (see details in Table 1), each comparison corresponding to a LLD group is subject to an effect size. The "edu. diff. Ctrl>LLD" represents a ratio

calculated to reflect the extent of the difference in educational level between the two groups being compared; positive values of this ratio indicate that the educational level of the control group was higher than that of the LLD group

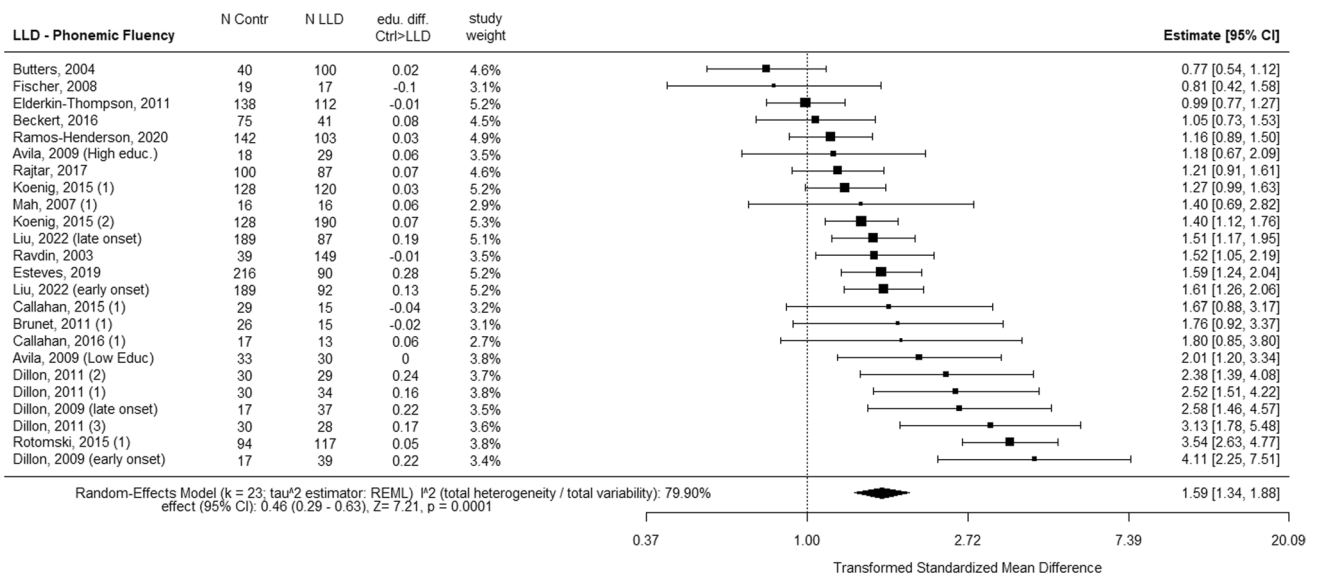


Fig. 3 Forest plot of average effects of LLD (compared to the control group) for outcomes of phonemic fluency including differences in education. *Note:* Same as for Fig. 2

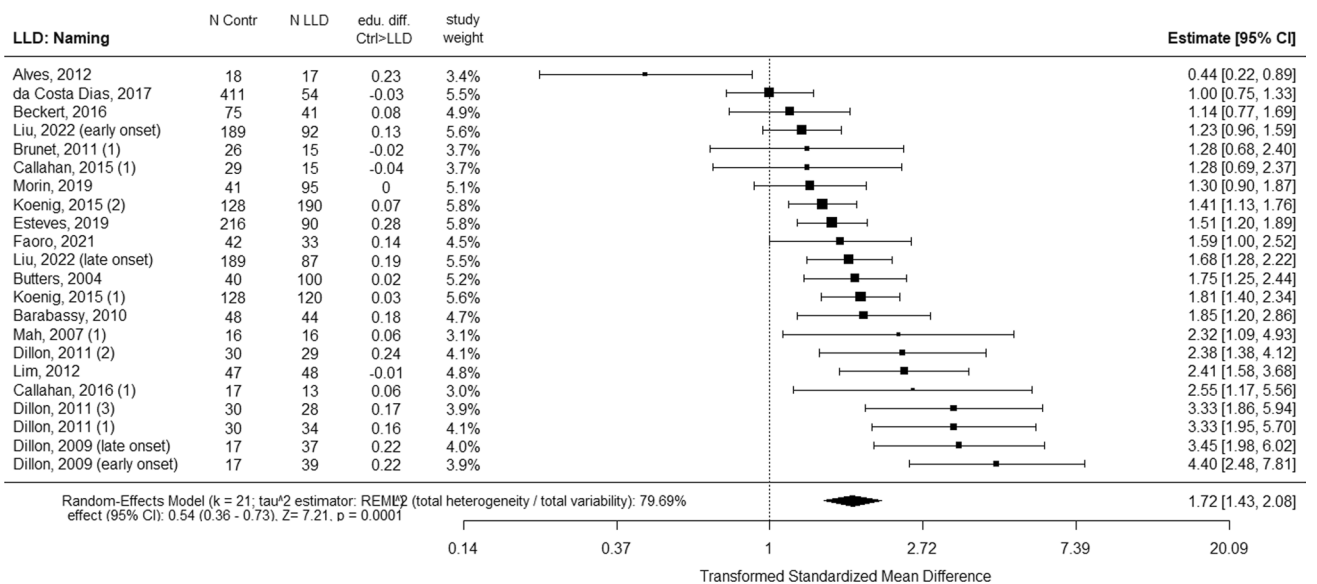


Fig. 4 Forest plot of average effects of LLD (compared to the control group) for outcomes of Naming task including differences in education. *Note:* Same as for Fig. 2

Rajtar-Zembati et al. 2017), we see that Butters et al. (2004) showed an effect of LLD only on naming but did not report a control of the severity of the depressive status of the LLD group while they did control the age of onset of depression. Brunet et al. (2011) and Callahan et al. (2015) showed no significant effect of LLD on semantic fluency, phonemic fluency or confrontation naming. On a methodological level, Brunet et al. (2011) did control for the homogeneity of education and the level of depressive symptoms, but did not

report information about the previous history of depression in the LLD groups. Callahan et al. (2015) strictly controlled the absence of neurocognitive disorder in the LLD group and the homogeneity of education between the groups. Rajtar-Zembati et al. (2017) that reported an effect on semantic fluency controlled that none of the participants had a previous history of depression, making it a late-onset LLD group, but reported that some of the individuals in LLD group were medicated with antidepressants (35%), and this variable was

not controlled in the statistical analysis of the results (Rajtar-Zembaty et al. 2017).

Four studies showed an effect of LLD on scores for phonemic fluency but not for semantic fluency (Avila et al. 2009; Esteves et al. 2019; Ravdin et al. 2003; Rotomskis et al. 2015). Avila et al.'s (2009) research was strictly balanced for the similarity of education between groups. They also measured the severity of depressive symptoms in the depressed group, showing that depressive disorders were mostly mild-to-moderate. Rotomskis et al. (2015) did not provide the level of depressive symptoms for any of the groups or participants' previous history of depression. However, the homogeneity of the education level was strictly controlled (Rotomskis et al. 2015). Esteves et al. (2019) reported an effect of LLD on phonemic fluency and on confrontation naming task, but not on semantic fluency. The article did not provide a diagnostic reference (neither DSM nor ICD) for the classification of LLD. Ravdin et al. (2003) strictly controlled the homogeneity of the education level between groups, but did not report the previous history of depression for the participants (Ravdin et al. 2003).

Four studies showed no effect of LLD on scores of either of the fluency task (Beckert et al. 2016; Elderkin-Thompson et al. 2011; Fischer et al. 2008; Mah et al. 2017). They reported strict control of non-dementia status for the participants of the LLD group. In Mah et al. (2017) and Fischer et al. (2008) all participants in the LLD group suffered from severe depression. In Elderkin-Thompson et al. (2011) the participants were separated into early- and late-onset LLD and drug free for antidepressant therapy at the time of the research (Elderkin-Thompson et al. 2011). In Beckert et al. (2016), the LLD group was of low education (half of the group had less than seven years of education and the other half had less than three years of education or were illiterate), the severity of depressive symptoms was not controlled as a variable, and participants were accepted in the research with Mini mental state evaluation (MMSE) scores of 22/30.

Five studies reported LLD scores only for semantic fluency and confrontation naming (Alves et al. 2012; Barabassy et al. 2010; da Costa Dias et al. 2017; Faoro et al. 2021; Morin et al. 2019). Barabassy et al. (2010) reported important differences in the level of education between groups, with a lower and more widely distributed level in the LLD group. In this study, the level of depressive symptoms was not reported in the control group (Barabassy et al. 2010). Da Costa Dias et al. (2017) compared groups of huge differences of sizes (411 controls for 54 LLD) (da Costa Dias et al. 2017).

Liu et al. (2022) compared the early and late onset of LLD and balanced the level of the severity of depressive symptoms between them. The level of education was significantly lower for both depressive groups than for the control group. For the early-onset group, scores were significantly lower for phonemic fluency.

When it comes to the studies including also a comparison with a MCI or AD. Brunet et al. (2011) and Callahan et al. (2015) included two groups of aMCI, with and without depressive symptoms, and strictly controlled for the homogeneity of education between groups. Their results show that aMCI with depressive symptoms had significantly lower scores for semantic fluency and confrontation naming but not for phonemic fluency; aMCI without depressive symptoms had significantly lower scores only for semantic fluency but not for phonemic fluency or confrontation naming. Mah et al. (2017) included a group of patients with aMCI without depressive symptoms in their study. The groups were homogeneous with a high level of education (> 15 years of education). Their results showed no significant differences between groups for either fluencies or confrontation naming. Finally, all the studies including groups with AD (with or without depression) showed a significant effect of AD on the three tasks. In summary, the order of magnitude of the pathological situation effects on naming and both fluency was an overwhelming dominance of AD with a much weaker and more heterogeneous influence of LLD or aMCI, likely multi-determined. This leads us to conclude that measures of fluency and naming, while demonstrating high discriminatory value for AD, do not have the same clinical value for cognitive assessment to identify LLD and aMCI. However, the effect is still present and does not allow us to claim that in the case of LLD, these tasks must not be expected to be impaired, which denies the possibility of using them as discriminating variables in the differential diagnosis.

Effects of LLD, aMCI AND AD ON low-executive tasks

Regarding tests in low-executive group, defined as not relying on a high-executive contribution, three studies (Brunet et al. 2011; Callahan et al. 2015, 2016) used the Pyramid and Palm Tree Test (PPTT) (Adlam et al. 2010). None of these studies showed an effect of LLD or aMCI. Notice that this one included groups of participants with a larger number of participants than in the other two studies (Brunet et al. 2011; Callahan et al. 2016). PPTT uses non-verbal responses to access semantic knowledge to identify that two images are semantically linked. It is a spotting task designed to reduce the reliance on other cognitive resources different from activation in the semantic representation system.

Two studies (Butters et al. 2004; Koenig et al. 2015) reported the results of a spot-the-word task (from the WAIS (Hartman 2009)) and did not show any significant effect of LLD on it. This task, also described as a lexical decision task, is conducted by presenting to the participant pairs of items comprising one word and one non-word, and requiring the subject to identify the word (Baddeley et al. 1993) and is reported to be a measure of cognitive ability

that is resistant to the effects of brain injury because it relies on crystallized measures of verbal knowledge.

Two studies used the test of similarities (Fischer et al. 2008; Morin et al. 2019) to compare the LLD and control groups and showed no effect. In the similarities subtest, the participant must identify the similarities between two concepts, which requires the retrieval of knowledge of both concepts, but also abstract thinking skills and verbal reasoning. Three studies used a vocabulary subtest (Dillon et al. 2009; Faoro et al. 2021; Mah et al. 2017) that requires to define of up to 30 words. The results of the three studies differed: Dillon et al. (2009) and Faoro et al. (2021) show no significant effect of LLD, while Mah et al. (2017) found a significant effect of LLD on the performance of this task (and no effect of aMCI with depressive symptoms). Note that the vocabulary task of the WAIS includes several words highly linked to emotions (e.g., regrets, courage, remorse, etc.) that can have a negative effect on the compliance of depressed participants to complete the task.

Rotomskis et al. (2015) and Beckert et al. (2016) compared LLD and control groups using the language subtest of the ACE-R and did not report significant differences between LLD and control group but a significant effect of AD on the outcomes.

Finally, Elderkin-Thompson et al. (2006) compared major and minor LLD to a control group with a cognitive assessment including the learning, recall, and recognition tasks of the Californian Verbal and Learning Test (Elwood 1995). The semantic clustering strategy index was measured by the authors as an indicator of executive mediation for learning task performance. The results of the recall and recognition tasks did not significantly differ from one group to another, while the learning task did ($F(8316) = 3.71$; $p < 0.001$), as did the semantic clustering index ($F(2160) = 4.79$; $p = 0.01$). The authors interpreted this result as a demonstration that the relationship between depression and verbal learning tasks is mediated by executive ability, as quantified by semantic clustering. However, they did not show a distinct effect of minor or major depression on this indicator.

General discussion

The cognitive complaints encountered in LLD make it difficult to distinguish between aMCI and AD based on an analysis of neurocognitive disorders. The hypothesis of the early impairment of semantic memory in AD and aMCI is considered a potential differential cognitive clue. By systematically seeking neuropsychological assessments of individuals with LLD, the present study included 31 studies representing 3291 controls and 2820 people with

LLD. Wherever possible, studies that tested simultaneously groups with LLD, AD (or aMCI) were also included.

The meta-analysis of group comparisons of scores at tasks that were influenced by executive resources (verbal fluency or confrontation naming) showed a moderate effect of LLD on these tasks, which was less important than the effect of AD reported by previous meta-analyses (Henry and Crawford 2005; Laws et al. 2007, 2010). Then, the aMCI showed an incremental deleterious effect on semantic fluency, but not on phonemic fluency and naming. Finally, AD had an enhanced effect on semantic fluency and naming, but not on phonemic fluency. These results show that semantic fluency is sensitive to the difference in cognitive impairment encountered between LLD and aMCI (and AD), that naming only shows an incremental effect with AD and not with aMCI, and that phonemic fluency is not sensitive to the difference between LLD, aMCI or sometimes even AD. Research comparing aMCI with and without depressive symptoms shows that aMCI with depressive symptoms has the same pattern as AD (incremental deficits in semantic fluency and naming scores compared to LLD), but that in highly educated groups this pattern does not emerge, suggesting that compensatory mechanisms are at work. These results are coherent regard to the hypothesis that the presence of depressive symptoms during aMCI is linked to a major rate of conversation to AD (Invernizzi et al. 2021), and with the hypothesis of the specific impairment of the activation of the semantic representation system as one of the earliest signs of AD.

This result can provide complementary information about aMCI to that reported by Joubert et al. (2020) in a meta-analysis of the effect of MCI on semantic tests. Their results indicated a systematic effect of MCI on semantic tests when compared with healthy controls (Joubert et al. 2020). However, in their research, the selected semantic tests were confrontation naming, free recall of semantic information, and facilitated recall of semantic information. Confrontation naming and free recall of semantic information rely on self-initiated recall or retrieval of semantic information (such as verbal fluency).

In general, the outcomes from our pooled effect sizes support the idea that during the cognitive assessment of a person with LLD, verbal fluency scores are expected to be affected, but not to the same extent as in the case of AD. However, the difference in effect between LLD and aMCI is not as clear and does not allow fluency to be identified as a discriminating tool in a differential cognitive assessment between aMCI and LLD. Moreover, in the case of LLD, the results were strongly influenced by the heterogeneity of education levels between the groups, which makes it even more difficult to interpret the results. Our results show that differences in educational level between groups predict outcomes on the phonemic and semantic fluency

and naming tasks, but do not predict outcomes of the other tasks. However, looking at each study separately, we could see that some studies had strictly controlled this balance of education between groups and sometimes showed effects and at other times not. The other possible sources of heterogeneity were the severity of the depression, the type of depression and the history of previous episodes. Future research on the cognitive profile in LLD should strictly control these three elements.

The hypothesis of late-onset LLD (i.e., without a history of depression earlier in life) reports that it is a frequent sign of a prodrome of AD (Heser et al. 2013; Ohanna et al. 2011), suggesting that when this is controlled for, the impact of this particular late-onset LLD would be more pronounced on semantic than phonemic fluency; however, this trend did not emerge. In the same way that some cognitive expressions specific to early AD do not materialize in aMCI, it is possible that this specific effect of AD on fluency is not so precocious as to make it an indication that certain types of late-onset LLD are to be observed as a possible prodrome of AD.

Since the included studies were not longitudinal, we can consider that the aMCI groups evaluated represents both those who evolve and those who remain stable; hence, we cannot draw any conclusions regarding the situation of aMCI intended to evolve into AD. An interesting line of research in this sense would be to conduct a meta-analysis on the effects of aMCI on these three tests, including only longitudinal research, allowing the classification of aMCI according to its evolution. At this point, however, our preliminary conclusion on the relationship between aMCI and verbal fluency and naming is that these tests are not sensitive enough to play a significant role in the cognitive profile of aMCI. Consistent with previous research (Henry and Crawford 2005; Laws et al. 2007, 2010), the incremental effect of AD is fully demonstrated for these tests, except for phonemic fluency.

One part of our review analyses the studies that have extracted results from experimental groups on tasks involving semantic retrieval, but assumed to require less, or at least less intense, executive input. These studies do not present fully unanimous results, but the purest semantic tasks (such as the PPTT), which are affected by AD, are spared in the case of LLD.

Conclusion

The use of neuropsychological tests relying on semantic memory functioning but involving a strong executive component such as phonemic fluency, is not of differential diagnostic interest between LLD and AD (or aMCI) because,

although of lesser magnitude, LLD in a large proportion of cases also has a negative impact on scores for these tasks. Moreover, there is considerable heterogeneity in the results of these tasks in LLD. In this sense, data such as the level of education, degree of severity of depression, its typology, and its history, must be strictly controlled in research linking this pathology to these tasks.

However, the use of semantic fluency, naming or of low-executive tests (that do not rely on the self-initiation of semantic information and therefore rely much less on executive functions) are significantly less affected by LLD than AD, which makes them relevant for differential diagnosis, in line with the hypothesis of a specific and early impairment of the activation of the semantic representation system in AD.

If a clear distinction exists between the absence of an effect of LLD and the significant effect of AD on the activation of the semantic representation system, further research must be conducted to specify the difference in impairment between these conditions on the executive semantics involved in self-initiated semantic activities. When the studies are not longitudinal, it is difficult to assert a link between the presence of depressive symptoms in aMCI as a sign that it is a prodromal aMCI to AD, confirmed by early semantic impairment. This hypothesis should also be explored through a meta-analysis of longitudinal studies on aMCI. However, our results show a clear discriminative character of these tasks in distinguishing between AD and LLD.

The exercise of pooling tests designed to assess semantic memory remains a difficult one in the framework of a systematic review of the literature. Indeed, the practice of using tests assessing semantic memory in a pure way (such as the camel and cactus test) remains rather confidential and we had to work with more disparate material, such as fluency or vocabulary subscales of intelligence tests. However, despite the heterogeneity encountered, our review shows that almost all the tests that rely on semantic memory but are also determined by other cognitive functions (such as executive functions for fluency or language for vocabulary tests) are affected by LLD. However, our results also show that for these tests, AD has a substantial incremental deleterious effect. Our results also show that tasks that minimized the use of any function other than activation and selection in semantic memory (typically semantic matching tasks, which are very sensitive to AD) show very normal results in the context of LLD. Finally, our results show that aMCI is a pathological condition whose cognitive impairment is similar to LLD, but that in the presence of aMCI and depressive symptoms simultaneously, the cognitive pattern becomes more similar to that of AD than LLD. Despite the limitations of this work, it provides an interesting basis for the differential analysis of the cognitive processes at work in these three pathologies and allows us to hypothesize, in

particular, about the continuum that may link aMCI to AD, through the presence of depressive symptoms. Future work will have to quantify the difference in effect between aMCI and AD on these tests as a whole, to complete the analytical grid constructed around the neuropsychological assessment.

Authors contributions SI: Corresponding author; management of the systematic review; selection of articles; quality evaluation and extraction of data (in double-blind with AB); statistical analysis and writing of the article. AB: Quality evaluation of articles for selection; extraction of data in double-blind with Sandra Invernizzi. LL: Revised the article critically for important intellectual content; approved the version to be submitted; agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. ISL: Revised the article critically for important intellectual content; Approved the version to be submitted; agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Availability of data and materials https://osf.io/evrwf/?view_only=935fb2bcf5aa487f90ec22a034835b0a.

Declarations

Ethical approval Not applicable.

Competing interests We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome. Each contributor has carried out the work under research funding independent of any private organization.

General statement We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us. We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property. We understand that the Corresponding Author is the sole contact for the Editorial process (including Editorial Manager and direct communications with the office). She is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs. We confirm that we have provided a current, correct email address which is accessible by the Corresponding Author.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adlam A-LR, Patterson K, Bozeat S, Hodges JR (2010) The Cambridge semantic memory test battery: detection of semantic deficits in semantic dementia and Alzheimer's disease. *Neurocase* 16(3):193–207
- Alexopoulos GS, Kiosses DN, Heo M, Murphy CF, Shanmugham B, Gunning-Dixon F (2005) Executive dysfunction and the course of geriatric depression. *Biol Psychiatry* 58(3):204–210. <https://doi.org/10.1016/j.biopsych.2005.04.024>
- Alves GS, Karakaya T, Fußer F, Kordulla M, O'Dwyer L, Christl J, Magerkurth J, Oertel-Knöchel V, Knöchel C, Prvulovic D (2012) Association of microstructural white matter abnormalities with cognitive dysfunction in geriatric patients with major depression. *Psychiatry Res Neuroimaging* 203(2–3):194–200.
- Avila R, Moscoso MA, Ribeiz S, Arrais J, Jaluol O, Bottino CM (2009) Influence of education and depressive symptoms on cognitive function in the elderly. *Int Psychogeriatr* 21(3):560–567. <https://doi.org/10.1017/S1041610209008928>
- Baddeley A, Emslie H, Nimmo-Smith I (1993) The Spot-the-Word test: A robust estimate of verbal intelligence based on lexical decision. *Br J Clin Psychol* 32(1):55–65
- Ballmaier M, Kumar A, Thompson PM, Narr KL, Lavretsky H, Estanol L, DeLuca H, Toga AW (2004) Localizing gray matter deficits in late-onset depression using computational cortical pattern matching methods. *Am J Psychiatry* 161(11):2091–2099
- Balsamo M, Cataldi F, Carlucci L, Padulo C, Fairfield B (2018) Assessment of late-life depression via self-report measures: a review. *Clin Interv Aging* 13:2021–2044. <https://doi.org/10.2147/CIA.S178943>
- Barabassy A, Beinhoff U, Riepe MW (2010) Cognitive estimation in aged patients with major depressive disorder. *Psychiatry Res* 176(1):26–29
- Beckert M, Loureiro F, Menta C, Mello EF, Nogueira EL, Gunten Av, Gomes I (2016) Performance of low-educated elders with depression on Addenbrooke's Cognitive Examination-Revised (ACE-R) test. *Dementia Neuropsychol* 10:19–25
- Belanger S, Belleville S (2009) Semantic inhibition impairment in mild cognitive impairment: a distinctive feature of upcoming cognitive decline? *Neuropsychology* 23(5):592–606. <https://doi.org/10.1037/a0016152>
- Binder JR, Desai RH (2011) The neurobiology of semantic memory. *Trends Cogn Sci* 15(11):527–536. <https://doi.org/10.1016/j.tics.2011.10.001>
- Binder JR, Desai RH, Graves WW, Conant LL (2009) Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex* 19(12):2767–2796. <https://doi.org/10.1093/cercor/bhp055>
- Boccia M, Acierio M, Piccardi L (2015) Neuroanatomy of Alzheimer's disease and late-life depression: a coordinate-based meta-analysis of MRI studies. *J Alzheimers Dis* 46(4):963–970
- Bonett DG (2009) Meta-analytic interval estimation for standardized and unstandardized mean differences. *Psychol Methods* 14(3):225
- Borenstein M, Hedges L, Higgins J, Rothstein H (2009) Power analysis for meta-analysis. *Introduction to Meta-analysis*, 257–276.
- Brambati SM, Belleville S, Kergoat M-J, Chayer C, Gauthier S, Joubert S (2009) Single- and multiple-domain amnesic mild cognitive impairment: two sides of the same coin. *Dement Geriatr Cogn Disord* 28(6):541–549
- Brambati SM, Peters F, Belleville S, Joubert S (2012) Lack of semantic priming effects in famous person recognition in Mild Cognitive Impairment. *Cortex* 48(4):414–420. <https://doi.org/10.1016/j.cortex.2011.04.001>

- Brunet J, Hudon C, Macoir J, Belleville S, Rousseau F, Bouchard RW, Verret L, Chertkow H, Chayer C, Kergoat MJ, Joubert S (2011) The relation between depressive symptoms and semantic memory in amnesic mild cognitive impairment and in late-life depression. *J Int Neuropsychol Soc* 17(5):865–874. <https://doi.org/10.1017/S1355617711000877>
- Butters MA, Whyte EM, Nebes RD, Begley AE, Dew MA, Mulsant BH, Zmuda MD, Bhalla R, Meltzer CC, Pollock BG (2004) The nature and determinants of neuropsychological functioning in late-lifedepression. *Arch Gen Psychiatry* 61(6):587–595
- Callahan BL, Joubert S, Tremblay MP, Macoir J, Belleville S, Rousseau F, Bouchard RW, Verret L, Hudon C (2015) Semantic memory impairment for biological and man-made objects in individuals with amnesic mild cognitive impairment or late-life depression. *J Geriatr Psychiatry Neurol* 28(2):108–116. <https://doi.org/10.1177/0891988714554708>
- Callahan BL, Laforce R, Dugas M, Hudon C (2016) Memory for emotional images differs according to the presence of depressive symptoms in individuals at risk for dementia. *Int Psychogeriatr* 29(4):673–685. <https://doi.org/10.1017/S1041610216002179>
- Chasles M-J, Tremblay A, Escudier F, Lajeunesse A, Benoit S, Langlois R, Joubert S, Rouleau I (2019) An examination of semantic impairment in amnesic MCI and AD: what can we learn from verbal fluency? *Arch Clin Neuropsychol* 35(1):22–30. <https://doi.org/10.1093/arclin/acz018>
- Chiou R, Humphreys GF, Jung J, Lambon Ralph MA (2018) Controlled semantic cognition relies upon dynamic and flexible interactions between the executive “semantic control” and hub-and-spoke “semantic representation” systems. *Cortex* 103:100–116. <https://doi.org/10.1016/j.cortex.2018.02.018>
- Colman AM (2009) Mill hill vocabulary scale. Oxford University Press, Oxford
- Copland DA, de Zubicaray GI, McMahon K, Eastburn M (2007) Neural correlates of semantic priming for ambiguous words: an event-related fMRI study. *Brain Res* 1131(1):163–172. <https://doi.org/10.1016/j.brainres.2006.11.016>
- Croisile B, Astier JL, Beaumont C, Mollion H (2011) The 5-word test in 37 depressed patients compared with 36 normal controls and 35 patients with mild Alzheimer’s disease. *Encephale* 37(2):127–132. <https://doi.org/10.1016/j.encep.2010.10.003>
- da Costa Dias FL, Teixeira AL, Guimarães HC, Barbosa MT, Resende E, d. P. F., Beato, R. G., Carmona, K. C., & Caramelli, P. (2017) Cognitive performance of community-dwelling oldest-old individuals with major depression: the Pietà study. *Int Psychogeriatr* 29(9):1507–1513
- de Partz de Courtray M-P, Bilocq V, De Wilde V, Seron X, Pillon A (2001) LEXIS. Tests pour le diagnostic des troubles lexicaux chez le patient aphasique.
- De Rotrou J, Wenisch E, Chausson C, Dray F, Faucounau V, Rigaud AS (2005) Accidental MCI in healthy subjects: a prospective longitudinal study. *Eur J Neurol* 12(11):879–885
- Deeks J, Wisniewski S, Davenport C (2013) Guide to the contents of a cochrane diagnostic test accuracy protocol. In: Deeks JJ, Bossuyt PM, Gatsonis C (eds), *Cochrane handbook for systematic reviews of diagnostic test accuracy version 1.0. 0*. The Cochrane Collaboration, 2013
- Díaz-Mardomingo MDC, García-Herranz S, Rodríguez-Fernández R, Venero C, Peraita H (2017) Problems in classifying mild cognitive impairment (MCI): one or multiple syndromes? *Brain Sci* 7(9):111
- Dillon C, Allegri RF, Serrano CM, Iturry M, Salgado P, Glaser FB, Taragano FE (2009) Late- versus early-onset geriatric depression in a memory research center. *Neuropsychiatr Dis Treat* 5:517–526. <https://doi.org/10.2147/ndt.s7320>
- Dillon C, Machnicki G, Serrano CM, Rojas G, Vazquez G, Allegri RF (2011) Clinical manifestations of geriatric depression in a memory clinic: toward a proposed subtyping of geriatric depression. *J Affect Disord* 134(1–3):177–187. <https://doi.org/10.1016/j.jad.2011.05.036>
- Domoto-Reilly K, Sapolsky D, Brickhouse M, Dickerson BC, Initiative A, s. D. N. (2012) Naming impairment in Alzheimer’s disease is associated with left anterior temporal lobe atrophy. *Neuroimage* 63(1):348–355
- Dubois B, Albert ML (2004) Amnesic MCI or prodromal Alzheimer’s disease? *Lancet Neurol* 3(4):246–248
- Dudas RB, Berrios GE, Hodges JR (2005) The Addenbrooke’s cognitive examination (ACE) in the differential diagnosis of early dementias versus affective disorder
- Elderkin-Thompson V, Mintz J, Haroon E, Lavretsky H, Kumar A (2006) Executive dysfunction and memory in older patients with major and minor depression. *Arch Clin Neuropsychol* 21(7):669–676. <https://doi.org/10.1016/j.acn.2006.05.011>
- Elderkin-Thompson V, Moody T, Knowlton B, Hellemann G, Kumar A (2011) Explicit and implicit memory in late-life depression. *Am J Geriatr Psychiatry* 19(4):249–255. <https://doi.org/10.1097/JGP.0b013e3181e89a5b>
- Elwood RW (1995) The California verbal learning test: psychometric characteristics and clinical application. *Neuropsychol Rev* 5(3):173–201
- Esteves CS, Lima MP, Gonzatti V, de Oliveira CR, de Lima Argimon II, Irigaray TQ (2017) Depressive symptoms and cognitive functioning of elderly from the family health strategy. *Ageing Int* 44(1):24–40. <https://doi.org/10.1007/s12126-017-9313-7>
- Faoro M, Hamdan AC (2021) Depressive symptoms may be associated with semantic memory decline in elderly adults. *Dementia Neuropsychol* 15(3):350
- First MB, Spitzer RL, Gibbon Miriam W, Janet B (1997) Structured clinical interview for DSM-IV axis I disorders: SCID-I: clinical version: scoresheet.
- Fischer C, Schweizer TA, Atkins JH, Bozanovic R, Norris M, Herrmann N, Nisenbaum R, Rourke SB (2008) Neurocognitive profiles in older adults with and without major depression. *Int J Geriatric Psychiatry* 23(8):851–856
- Hartman DE (2009) Wechsler adult intelligence scale IV (WAIS IV): return of the gold standard. *Appl Neuropsychol* 16(1):85–87
- Hedges LV (1981) Distribution theory for Glass’s estimator of effect size and related estimators. *J Educ Stat* 6(2):107–128
- Henry JD, Crawford JR (2004) A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology* 18(2):284–295. <https://doi.org/10.1037/0894-4105.18.2.284>
- Henry JD, Crawford JR (2005) A meta-analytic review of verbal fluency deficits in depression. *J Clin Exp Neuropsychol* 27(1):78–101. <https://doi.org/10.1080/138033990513654>
- Hernández M, Costa A, Juncadella M, Sebastián-Gallés N, Reñé R (2008) Category-specific semantic deficits in Alzheimer’s disease: a semantic priming study. *Neuropsychologia* 46(4):935–946. <https://doi.org/10.1016/j.neuropsychologia.2007.11.018>
- Heser K, Tebarth F, Wiese B, Eisele M, Bickel H, Kohler M, Mosch E, Weyerer S, Werle J, König HH, Leicht H, Pentzek M, Fuchs A, Riedel-Heller SG, Luppa M, Prokein J, Scherer M, Maier W, Wagner M, Age CoDe Study, G (2013) Age of major depression onset, depressive symptoms, and risk for subsequent dementia: results of the German study on Ageing, Cognition, and Dementia in Primary Care Patients (AgeCoDe). *Psychol Med*, 43(8), 1597–1610. <https://doi.org/10.1017/S0033291712002449>
- Higby E, Cahana-Amitay D, Vogel-Eyney A, Spiro A 3rd, Albert M, Obler L (2019) The role of executive functions in object- and action-naming among older adults. *Exp Aging Res* 45(4):306–330
- Higgins JP, Thompson SG (2002) Quantifying heterogeneity in a meta-analysis. *Stat Med* 21(11):1539–1558

- Hoffman P (2018) An individual differences approach to semantic cognition: Divergent effects of age on representation, retrieval and selection. *Sci Rep* 8(1):8145. <https://doi.org/10.1038/s41598-018-26569-0>
- Invernizzi S, Simoes Loureiro I, Kandana Arachchige K, Lefebvre L (2021) Late-life depression, cognitive impairment, and relationship with Alzheimer's disease. *Dementia Geriatr Cognit Disor*. <https://doi.org/10.1159/000519453>
- Jefferies E, Thompson H, Cornelissen P, Smallwood J (2020) The neurocognitive basis of knowledge about object identity and events: dissociations reflect opposing effects of semantic coherence and control. *Philos Trans R Soc Lond B Biol Sci* 375(1791):20190300. <https://doi.org/10.1098/rstb.2019.0300>
- Joubert S, Gardy L, Didic M, Rouleau I, Barbeau EJ (2020) A meta-analysis of semantic memory in mild cognitive impairment. *Neuropsychol Rev*. <https://doi.org/10.1007/s11065-020-09453-5>
- Joyal M, Groleau C, Bouchard C, Wilson MA, Fecteau S (2020) Semantic processing in healthy aging and Alzheimer's disease: a systematic review of the N400 differences. *Brain Sci*. <https://doi.org/10.3390/brainsci10110770>
- Jungwirth S, Zehetmayer S, Hinterberger M, Kudrnovsky-Moser S, Weissgram S, Tragl KH, Fischer P (2011) The influence of depression on processing speed and executive function in nondemented subjects aged 75. *J Int Neuropsychol Soc* 17(5):822–831
- Kaplan E, Goodglass H, Weintraub S (2001) Boston naming test
- Klein LA, Buchanan JA (2009) Psychometric properties of the pyramids and palm trees test. *J Clin Exp Neuropsychol* 31(7):803–808
- Koenig AM, DeLozier IJ, Zmuda MD, Marron MM, Begley AE, Anderson SJ, Reynolds III CF, Arnold SE, Becker JT, Butters MA (2015) Neuropsychological functioning in the acute and remitted States of late-life depression. *J Alzheimers Dis* 45(1):175–185. <https://doi.org/10.3233/JAD-148006>
- Kohler S, Thomas AJ, Lloyd A, Barber R, Almeida OP, O'Brien JT (2010) White matter hyperintensities, cortisol levels, brain atrophy and continuing cognitive deficits in late-life depression. *Br J Psychiatry* 196(2):143–149. <https://doi.org/10.1192/bjp.bp.109.071399>
- Kuznetsova A, Brockhoff PB, Christensen RHB (2015) Package 'lmerest.' R Package Version 2:734
- Laisney M, Eustache F, Desgranges B (2009) Évaluation de la mémoire sémantique relative aux personnes célèbres-SemPer. *Rev Neuropsychol* 1(2):175–183. <https://doi.org/10.3917/rne.012.0175>
- Laisney M, Giffard B, Belliard S, de la Sayette V, Desgranges B, Eustache F (2011) When the zebra loses its stripes: Semantic priming in early Alzheimer's disease and semantic dementia. *Cortex* 47(1):35–46. <https://doi.org/10.1016/j.cortex.2009.11.001>
- Lamar M, Charlton R, Zhang A, Kumar A (2012) Differential associations between types of verbal memory and prefrontal brain structure in healthy aging and late life depression. *Neuropsychologia* 50(8):1823–1829. <https://doi.org/10.1016/j.neuropsychologia.2012.04.007>
- Lara E, Martín-María N, De la Torre-Luque A, Koyanagi A, Vancampfort D, Izquierdo A, Miret M (2019) Does loneliness contribute to mild cognitive impairment and dementia? A systematic review and meta-analysis of longitudinal studies. *Ageing Res Rev* 52:7–16
- Laws KR, Duncan A, Gale TM (2010) 'Normal' semantic-phonemic fluency discrepancy in Alzheimer's disease? A meta-analytic study. *Cortex* 46(5):595–601
- Laws, K. R., Adlington, R. L., Gale, T. M., & Sartori, G. (2007). A meta-analytic review of category naming in Alzheimer's disease.
- Lee RS, Hermens DF, Porter MA, Redoblado-Hodge MA (2012) A meta-analysis of cognitive deficits in first-episode major depressive disorder. *J Affect Disord* 140(2):113–124. <https://doi.org/10.1016/j.jad.2011.10.023>
- Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 6(7):e1000100. <https://doi.org/10.1371/journal.pmed.1000100>
- Lim HK, Jung WS, Ahn KJ, Won WY, Hahn C, Lee SY, Kim I, Lee CU (2012) Regional cortical thickness and subcortical volume changes are associated with cognitive impairments in the drug-naïve patients with late-onset depression. *Neuropsychopharmacology* 37(3):838–849. <https://doi.org/10.1038/npp.2011.264>
- Lin C, Huang CM, Fan YT, Liu HL, Chen YL, Aizenstein HJ, Lee TM, Lee SH (2020) Cognitive reserve moderates effects of white matter hyperintensity on depressive symptoms and cognitive function in late-life depression. *Front Psychiatry* 11:249
- Liu M, Chen B, Zhong X, Zhang M, Wang Q, Zhou H, Wu Z, Hou L, Peng Q, Zhang S, Yang M (2022) Differences in odor identification in early-onset and late-onset depression. *Brain Sci* 12(2):276.
- Mah L, Anderson ND, Verhoeff NPL, Pollock BG (2017) Negative emotional verbal memory biases in mild cognitive impairment and late-onset depression. *Am J Geriatr Psychiatry* 25(10):1160–1170
- McCrimmon AW, Smith AD (2013) Review of the Wechsler abbreviated scale of intelligence, (WASI-II). Sage, Los Angeles, CA
- McLaughlin PM, Wright MJ, Larocca M, Nguyen PT, Teng E, Apostolova LG, Ringman JM, Zhou Y, Cummings JL, Woo E (2014) The "Alzheimer's type" profile of semantic clustering in amnesic mild cognitive impairment. *J Int Neuropsychol Soc* 20(4):402–412. <https://doi.org/10.1017/S135561771400006X>
- Mioshi E, Dawson K, Mitchell J, Arnold R, Hodges JR (2006) The Addenbrooke's cognitive examination revised (ACE-R): a brief cognitive test battery for dementia screening. *Int J Geriatr Psychiatry* 21(11):1078–1085
- Morimoto SS, Gunning FM, Murphy CF, Kanellopoulos D, Kelly RE, Alexopoulos GS (2011) Executive function and short-term remission of geriatric depression: the role of semantic strategy. *Am J Geriatr Psychiatry* 19(2):115–122. <https://doi.org/10.1097/JGP.0b013e3181e751c4>
- Morin RT, Gonzales MM, Bickford D, Catalinotto D, Nelson C, Mackin RS (2019) Impaired financial capacity in late-life depression: revisiting associations with cognitive functioning. *J Int Neuropsychol Soc* 25(10):1088–1093. <https://doi.org/10.1017/S1355617719000705>
- Muniswamy NR, Mukku SSR, Sivakumar PT, Rajur S, Rajeswaran J, Mehta UM, Thirthalli J, Kumar KJ, Varghese M (2020) Study of neuropsychological deficits in late onset depression. *Asian J Psychiatr* 54:102435. <https://doi.org/10.1016/j.ajp.2020.102435>
- Ohanna I, Golander H, Barak Y (2011) Does late onset depression predispose to dementia? A retrospective, case-controlled study. *Compr Psychiatry* 52(6):659–661. <https://doi.org/10.1016/j.comppsy.2010.10.016>
- Olaya B, Moneta MV, Miret M, Ayuso-Mateos JL, Haro JM (2019) Course of depression and cognitive decline at 3-year follow-up: the role of age of onset. *Psychol Aging* 34(4):475
- Parra MA, Abrahams S, Logie RH, Della Sala S (2010) Visual short-term memory binding in Alzheimer's disease and depression. *J Neurol* 257(7):1160–1169. <https://doi.org/10.1007/s00415-010-5484-9>
- Perri R, Zannino G, Caltagirone C, Carlesimo GA (2012) Alzheimer's disease and semantic deficits: a feature-listing study. *Neuropsychology* 26(5):652
- Perri R, Carlesimo GA, Monaco M, Caltagirone C, Zannino GD (2019) The attribute priming effect in patients with Alzheimer's disease. *J Neuropsychol* 13(3):485–502. <https://doi.org/10.1111/jnp.12168>

- Petersen RC (2016) Mild cognitive impairment. *Continuum Lifelong Learn Neurol* 22(2):404
- Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E (1999) Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol* 56(3):303–308
- Rajtar-Zembaty A, Salakowski A, Rajtar-Zembaty J, Starowicz-Filip A (2017) Executive dysfunction in late-life depression. *Psychiatr Pol* 51(4):705–718
- Ramos-Henderson M, Ledezma-Dámas A, López N, Machado Goyano Mac Kay AP (2021) Executive functions and functional impairment in Latin seniors suffering from depression. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn* 28(4):543–558. <https://doi.org/10.1080/13825585.2020.1796915>
- Raucher-Chene D, Terrien S, Gobin P, Gierski F, Kaladjian A, Besche-Richard C (2017) Modulation of the N400 component in relation to hypomanic personality traits in a word meaning ambiguity resolution task. *Psychiatry Clin Neurosci* 71(9):637–646. <https://doi.org/10.1111/pcn.12534>
- Raucher-Chene D, Terrien S, Gierski F, Obert A, Caillies S, Besche-Richard C, Kaladjian A (2018) Neural correlates of semantic inhibition in relation to hypomanic traits: an fMRI study. *Front Psychiatry* 9:108. <https://doi.org/10.3389/fpsy.2018.00108>
- Ravdin LD, Katzen HL, Agrawal P, Relkin NR (2003) Letter and semantic fluency in older adults: effects of mild depressive symptoms and age-stratified normative data. *Clin Neuropsychol* 17(2):195–202. <https://doi.org/10.1076/clin.17.2.195.16500>
- Rotomskis A, Margeviciute R, Germanavicius A, Kaubrys G, Budrys V, Bagdonas A (2015) Differential diagnosis of depression and Alzheimer's disease with the Addenbrooke's cognitive examination-revised (ACE-R). *BMC Neurol* 15:57. <https://doi.org/10.1186/s12883-015-0315-3>
- Rozzini L, Chilovi BV, Trabucchi M, Padovani A (2008) Predictors of progression from mild cognitive impairment to Alzheimer disease. *Neurology* 70(9):735–736
- Sami S, Williams N, Hughes LE, Cope TE, Rittman T, Coyle-Gilchrist ITS, Henson RN, Rowe JB (2018) Neurophysiological signatures of Alzheimer's disease and frontotemporal lobar degeneration: pathology versus phenotype. *Brain* 141(8):2500–2510. <https://doi.org/10.1093/brain/awy180>
- Simoes Loureiro I, Lefebvre L (2015) The SQK: a semantic knowledge questionnaire to specify the severity of semantic deterioration in Alzheimer's disease patients. *Geriatr Psychol Neuropsychiatr Vieil* 13(2):225–233. <https://doi.org/10.1684/pnv.2015.0535>
- Simoes Loureiro I, Lefebvre L (2016a) Distinct progression of the deterioration of thematic and taxonomic links in natural and manufactured objects in Alzheimer's disease. *Neuropsychologia* 91:426–434. <https://doi.org/10.1016/j.neuropsychologia.2016.09.002>
- Simoes Loureiro I, Lefebvre L (2016b) Retrogenesis of semantic knowledge: comparative approach of acquisition and deterioration of concepts in semantic memory. *Neuropsychology* 30(7):853–859. <https://doi.org/10.1037/neu0000272>
- Simoes Loureiro I, Taverne M, Lefebvre L (2018) Le mini-QCS: un outil de dépistage rapide des troubles de la mémoire sémantique de la maladie d'Alzheimer. *Gériatrie Et Psychologie Neuropsychiatrie Du Vieillessement* 16(4):429–438
- Singmann H, Bolker B, Westfall J, Aust F (2015) Package 'afex'. In: Vienna
- Snyder HR (2013) Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. *Psychol Bull* 139(1):81–132. <https://doi.org/10.1037/a0028727>
- Srinivasan S (2014) Mild cognitive impairment: profile of a cohort from a private sector memory clinic. *Ann Indian Acad Neurol* 17(3):340–344
- Tam C, Lam L (2012) Cognitive and functional impairment in Chinese elderly with late-onset depression. *East Asian Arch Psychiatry* 22(1):25–30
- Teige C, Mollo G, Millman R, Savill N, Smallwood J, Cornelissen PL, Jefferies E (2018) Dynamic semantic cognition: characterising coherent and controlled conceptual retrieval through time using magnetoencephalography and chronometric transcranial magnetic stimulation. *Cortex* 103:329–349
- Vaughan RM, Coen RF, Kenny R, Lawlor BA (2016) Preservation of the semantic verbal fluency advantage in a large population-based sample: Normative data from the TILDA study. *J Int Neuropsychol Soc* 22(5):570–576
- Venneri A, McGeown WJ, Hietanen HM, Guerrini C, Ellis AW, Shanks MF (2008) The anatomical bases of semantic retrieval deficits in early Alzheimer's disease. *Neuropsychologia* 46(2):497–510. <https://doi.org/10.1016/j.neuropsychologia.2007.08.026>
- Viechtbauer W (2010) Conducting meta-analyses in R with the metafor package. *J Stat Softw* 36(3):1–48
- Warrington EK, Crutch SJ (2007) A within-modality test of semantic knowledge: the size/weight attribute test. *Neuropsychology* 21(6):803
- Weintraub S, Wicklund AH, Salmon DP (2012) The neuropsychological profile of Alzheimer disease. *Cold Spring Harb Perspect Med* 2(4):a006171
- Yuspeh RL, Vanderploeg RD (2000) Spot-the-word: a measure for estimating premorbid intellectual functioning. *Arch Clin Neuropsychol* 15(4):319–326
- Zannino GD, Caltagirone C, Carlesimo GA (2015) The contribution of neurodegenerative diseases to the modelling of semantic memory: A new proposal and a review of the literature. *Neuropsychologia* 75:274–290. <https://doi.org/10.1016/j.neuropsychologia.2015.06.023>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.