



Contents lists available at ScienceDirect

Journal of Neurolinguistics

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

Working memory treatment in aphasia: A theoretical and quantitative review

Steve Majerus*

Psychology & Neuroscience of Cognition Research Unit, University of Liège, Fund for Scientific Research FNRS, 1000 Brussels, Belgium

ABSTRACT

Working memory (WM) is a frequent and long-lasting deficit in patients with aphasia. Progress has been made in our understanding of the nature of WM impairment, by considering deficits at the level of maintenance of item and serial order information (short-term memory), deficits at the level of attentional control, and their complex interactions with language impairment. However, WM treatment studies in aphasic patients remain scarce. This theoretical and quantitative review of 15 single-case treatment studies (24 patients) reveals that WM treatment in aphasic patients is associated with robust near transfer effects and smaller far transfer effects. WM treatment can also raise verbal WM capacity to normal range performance levels. The specificity of WM treatment is more difficult to establish as control tasks/treatments are associated with treatment effects of comparable size as far transfer effects. Existing treatments cover phonological maintenance and control of attention components of WM but do not target difficulties at the level of item interference or serial order processing. In order to improve the specificity of WM treatments, future studies should define WM impairment and treatment targets in a more precise and theoretically informed manner. Also, although the vast majority of reviewed studies meet minimal methodological requirements for single case experimental designs, the use of a larger number of baseline, treatment and control measures is recommended. Finally, publication of treatment studies with negative outcomes should be encouraged as it is currently difficult to estimate the impact of publication bias on observed treatment effects.

1. Introduction

The aim of this critical review paper is to investigate the progress that has been made in the rehabilitation of working memory (WM) impairment in aphasia. WM impairment is a frequent, associated deficit of aphasia and it is also a long-lasting one, being still present in patients that appear to have recovered from their language impairment by showing no residual deficits in standard language measures (Caramazza, Basili, Koller, & Berndt, 1981; N. Martin, Saffran, & Dell, 1996). These WM deficits will continue to cause functional deficits in everyday life such as difficulties in following discussions and in taking notes, and they will prevent a full recovery of social and professional autonomy (Vallat et al., 2005). Different intervention strategies have been proposed, including both restoration and compensation techniques. In order to achieve an accurate understanding of the efficacy and specificity of these techniques, we first need to consider the complex nature of WM impairment in aphasia, in the light of current theoretical models.

* Psychology & Neuroscience of Cognition Research Unit, Université de Liège; Boulevard du Rectorat, B33, 4000 Liège, Belgium.
E-mail address: smajerus@uliege.be.

<https://doi.org/10.1016/j.jneuroling.2017.12.001>

Received 1 July 2017; Received in revised form 1 December 2017; Accepted 4 December 2017

0911-6044/ © 2017 Published by Elsevier Ltd.

2. The nature of WM impairment in aphasia

WM is a broad concept involving many different components, one distinction being the separation of storage components from processing components (e.g., Engle, Tuholski, Laughlin, & Conway, 1999). There is important theoretical confusion surrounding the terms ‘short-term memory’ (sometimes defined as the storage component) and ‘working memory’ (sometimes defined as the processing component). Cowan (2017) highlighted the theoretical conundrum that has arisen from the co-existence of (too) many different definitions and terms for the general function of WM. Given the current theoretical debates, I will use the term WM in a functional, a-theoretical manner and I will avoid opposing the terms ‘short-term memory’ and ‘working memory’. The term ‘WM’ will refer to the general function of temporary maintenance and will remain neutral as regards the components (e.g., only storage, storage and processing) involved. Critically, in order to maximize theoretical clarity, specific terms will be used to characterize the different components and processes involved in the different WM tasks and situations that will be discussed.

At the theoretical level, the concept of WM has made a long journey since the proposal of the working memory model by Baddeley and Hitch (1974). The initial working memory model contained a phonological loop system for storing verbal information, with a subdivision into in a buffer system, the phonological store, and a refreshing mechanism, the subvocal articulatory rehearsal process. This model also contained a visuo-spatial sketchpad for storing visuo-spatial information, as well as an attentional supervisory system. This model captured well the observation of patients with apparent isolated verbal WM deficits, as WM was considered to be a separate function from the language system and hence could be impaired in isolation (Vallar & Baddeley, 1984). However, this model did not capture the complex interactions that exist between verbal WM and language processing. These interactions are of particular importance for an accurate understanding of verbal WM deficits in patients with aphasia, as these deficits can be the cause or the consequence of underlying language impairment. For example, patients with reduced semantic processing abilities will show impaired storage abilities for stimuli with a semantic content such as words, but not for nonwords (Majerus, Patterson, & Norris, 2007). Even in patients with so-called isolated verbal WM impairment, a review of published cases showed that the vast majority of these patients initially presented with aphasia, and the severity of their verbal storage deficit correlated with their residual language processing abilities (Majerus, 2009). Progressively, WM models started to acknowledge more explicitly the interactions that link WM and language processing. Some models took a strong position by considering that short-term storage of verbal information is a processing property of the language system and reflects the decay rate of activated language representations (N. Martin & Saffran, 1992). Other models considered interactions between a phonological buffer system and phonological representations in the language system (Baddeley, Gathercole, & Papagno, 1998; R. Martin & Lesch, 1996) or between a specialized semantic buffer system and semantic levels of linguistic representations (R. Martin & Lesch, 1996; R. Martin, Shelton, & Yaffee, 1994). Many patients with aphasia are indeed sensitive to temporal factors in both linguistic and WM tasks, with performance decreasing when linguistic information has to be maintained over time (Majerus, Lekeu, Van der Linden, & Salmon, 2001; N. Martin et al., 1996). The importance of time-based decay in WM is also increasingly supported by studies in healthy adults (Barrouillet, Bernardin, & Camos, 2004; Ricker, Spiegel, & Cowan, 2014; Ricker, Vergauwe, & Cowan, 2016), although other authors continue to challenge the concept of time-based decay (Lewandowsky & Oberauer, 2009; Oberauer & Lewandowsky, 2013). In sum, as shown in Fig. 1, many current models of WM consider that the language processing system is at the basis of verbal WM: time-based decay of language activations can lead to verbal WM impairment particularly for the storage of verbal item information. We should note here that storage of serial order information has been considered to be supported by more specific mechanisms that I will describe below.

Control of interference within the linguistic system has been proposed to further determine verbal item storage capacities. Indeed, patients with aphasia can have intact access to their language representations, but they may have difficulties in inhibiting irrelevant

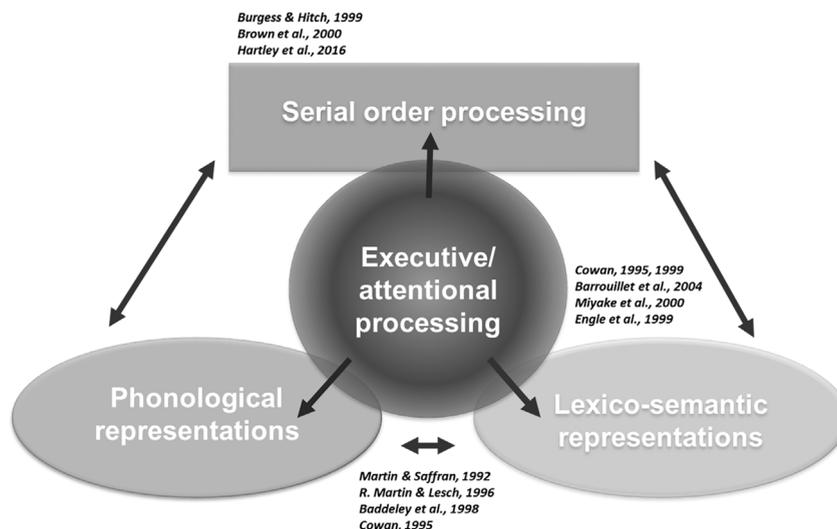


Fig. 1. Overview of WM components examined in this review.

language representations or in deactivating no longer relevant language representations (Hamilton & Martin, 2007; Hoffman, Jefferies, Ehsan, Hopper, & Lambon Ralph, 2009; Hoffman, Jones, & Lambon Ralph, 2012; Jefferies, Baker, Doran, & Lambon Ralph, 2007; Ralph, Jefferies, Patterson, & Rogers, 2017). These irrelevant or no-longer relevant activated word representations will lead to impaired performance in different language and verbal WM tasks, by interfering with the target information which will be more difficult to produce as well as by creating intrusion errors. Patients with inferior frontal or temporo-parietal lesions (Ralph et al., 2017) have been shown to be affected by phonological and/or semantic distractors in language processing and verbal WM tasks. Although the literature is still controversial, some studies further suggest that these deficits can affect selectively phonological or semantic levels of control/deactivation (Barde, Schwartz, Chrysikou, & Thompson-Schill, 2010; Gold, Balota, Kirchhoff, & Buckner, 2005). It is also important to note that this impairment at the level of the control of activated language representations can be associated with more general executive control deficits although there is important variability among patients here (e.g., Hoffman, Jefferies, Ehsan, Hopper, & Ralph, 2009).

A third aspect of WM impairment in aphasia reflects a process which has only been very recently identified and which concerns the storage of serial order information. An increasing number of models consider that the maintenance of item information and the maintenance of serial order information in WM are supported by distinct mechanisms (e.g., Burgess & Hitch, 1999, 2006; Brown, Preece, & Hulme, 2000; Hartley, Hurlstone, & Hitch, 2016; Henson, 1998; see also Fig. 1). Furthermore, there is increasing evidence for domain-general serial order coding and retention processes, given the involvement of superior parietal and intraparietal cortices and similar behavioral effects in verbal and visuo-spatial serial order WM tasks (Hurlstone, Hitch, & Baddeley, 2014; Majerus et al., 2010, 2006; Marshuetz, Smith, Jonides, DeGutis, & Chenevert, 2000). This does not exclude the existence of additional modality-specific serial order coding processes as two recent studies have shown that verbal serial order information can also be coded within the fronto-temporal language network (Kalm & Norris, 2014; Papagno et al., 2017). Critically, these studies show that even in that case, neural representations for item and serial order information are separable. The representation of item information is known to rely on language knowledge by being sensitive to different psycholinguistic variables and by leading to increased WM performance for items that can be represented more easily or more strongly within the linguistic system, as illustrated by the advantage for word list over nonword list recall (Nairne & Kelley, 2004). This is also supported by patients with semantic impairment which present reduced item recall but preserved serial order recall (Majerus, Norris, & Patterson, 2007). This does not mean that serial order information does not interact at all with linguistic variables: Poirier, Saint-Aubin, Mair, Tehan, and Tolan (2015) showed that serial order recall performance is less precise for lists composed of semantically related word lists, as opposed to unrelated word lists. Overall, these studies suggest that item information in WM reflects temporary activation of linguistic representations while arbitrary order information, when no long-term structures are available for binding a sequence of memoranda, is coded by distinct and specific mechanisms (which can however interact with the linguistic system) (Majerus, 2013; see also Fig. 1). By extension, this implies that patients with aphasia should present particularly pronounced item-based verbal WM impairment but may also present with serial order WM impairment. Importantly, the two aspects should be dissociable. This has indeed been shown recently in a study with patients with fluent aphasia. Majerus, Attout, Artielle, and Van der Kaa (2015) observed both selective item and serial order WM impairments when using WM tasks designed to maximize item or serial order retention requirements. It is further important to note that the serial order component of WM has been shown to maintain the sequential order of words in a list, but also the order of phonemes within novel word forms as for example during novel vocabulary learning (Majerus & Boukebza, 2013). Hence reduced serial order retention abilities in patients with aphasia may have a negative impact on the outcome of treatment strategies that involve the re-learning of lost word forms.

A fourth aspect of WM impairment in aphasia reflects deficient attentional and executive control (see Fig. 1). Although both terms are often used in an interchangeable manner, it is important to distinguish between attentional and executive control processes in WM (Engle & Kane, 2004). Executive control processes allow for the manipulation of information being maintained in WM; different types of executive processes have been proposed such as inhibition, shifting and updating (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). Although these different executive aspects are often considered to be domain-general and involve the dorso-lateral and ventro-lateral prefrontal cortices (Collette, Hogge, Salmon, & Van der Linden, 2006), there is also important controversy. For example, the deficit at the level of control of phonological/semantic interference discussed earlier could be considered as a language-specific executive deficit. Also, Kuzmina and Weekes (2017) recently showed that patients with fluent aphasia can have executive deficits restricted to the verbal modality. On the other hand, attentional control reflects the orientation of attention in a top-down, task related manner and has been associated with the fronto-parietal dorsal attention network (Corbetta & Shulman, 2002; Cowan, 1999). Attentional control demands not only increase with executive processing demands in WM tasks, but also with WM load in tasks with minimal executive processing demands. A number of studies have shown that in simple list-probe recognition WM tasks, the dorsal attention network is sensitive to the amount of memoranda, and this for both verbal and visuo-spatial WM tasks (Majerus et al., 2010, 2016; Majerus, Péters, Bouffier, Cowan, & Phillips, 2017; Todd & Marois, 2004). Furthermore, non-strategic attentional processes have been proposed to support WM performance, for both simple storage and more complex storage and processing WM tasks. Cowan (1995, 1999) described a basic attentional capacity which he termed the focus of attention; the focus of attention is considered to determine the amount of information we can hold in mind at one time (scope of attention) (see also, Barrouillet et al., 2004). This non-strategic capacity is considered to determine basic storage capacity and is limited to 3–4 items in young adults (Cowan, 2001). It is typically measured by tasks in which memoranda are presented at a very high speed preventing the implementation of any controlled attention processes during WM encoding (Cowan et al., 2005). The different attentional aspects, like the executive aspects, are often considered to be domain-general. A number of studies have shown cross-domain interference effects when verbal and visual WM tasks have to be performed concurrently (e.g., Cowan & Morey, 2007), although there may be some asymmetry, with visual-spatial WM tasks potentially relying more strongly on domain-general resources than verbal WM tasks

(Morey & Miron, 2016). This discussion of attentional aspects involved in WM tasks also shows how difficult it is to clearly delineate the concepts of 'working memory' and 'short-term memory' on the assumption that 'working memory' tasks (also sometimes called storage and processing tasks) would mainly recruit attentional and executive processes while these processes would not be involved in simple storage ('short-term memory') tasks. As already noted, simple storage tasks also involve attentional processes, particularly when load increases (see also Unsworth & Engle, 2007). At the same time, attentional control and executive processes are maximally recruited by tasks involving both storage and processing components as task control becomes more difficult and has to be shared by the storage and processing components.

In the area of aphasia, the aspects of executive and attentional control in WM have been mainly assessed via the administration of storage and processing tasks or via the administration of both verbal and visuo-spatial WM tasks by assuming that both tasks rely on common attentional resources. These studies have shown that attentional control and modality-general WM impairments can be observed in aphasia (Burgio & Basso, 1997; Mayer & Murray, 2012; Murray, Holland, & Beeson, 1998; Seniow, Litwin, & Lesniak, 2009). Also, the study by Kuzmina and Weekes (2017) recently showed that patients with aphasia can present deficits in both verbal and non-verbal aspects of executive processing tasks.

In sum, verbal WM impairment in aphasia is characterized by heterogeneity, with deficits affecting either the maintenance of item and/or serial order information over time or affecting the control of activated linguistic representations. Impairment at the level of executive and attentional aspects of WM has also been documented.

3. The optimization of WM

Before examining the rehabilitation of WM impairment in aphasia, a very brief overview of WM training in non-brain injured populations is provided here. WM training research has been very popular in the past 15 years, with more than 100 WM training studies published. It should be noted that the vast majority of these studies concerned healthy populations from all age groups which, by definition, show preserved WM abilities. Many of these studies used multiple-task WM training batteries, training several of the WM components described in the previous section at the same time (Klingberg, Forssberg, & Westerberg, 2002). These WM training batteries automatically and dynamically adapt the level of difficulty of the presented exercises to the participant's performance level. Initial studies appeared to show benefits of this type of WM training, with effects at both behavioral and neuroimaging levels (Constantinidis & Klingberg, 2016; Jaeggi, Buschkuhl, Jonides, & Shah, 2011; Klingberg, 2010; Klingberg et al., 2002; Kundu, Sutterer, Emrich, & Postle, 2013). However, other studies did not confirm these findings (e.g., Redick et al., 2013; Shipstead, Redick, & Engle, 2012), a result which was strengthened by subsequent meta-analyses. von Bastian and Oberauer (2014) reviewed 46 studies and observed no consistent evidence for near transfer and far transfer effects of different WM training programs; they argued that the benefits of WM training may depend on many different processes that are usually not taken into account, such as participant age, initial cognitive ability and personality factors. This conclusion was confirmed by a more recent meta-analysis on 87 training studies, controlling furthermore for publication bias (Melby-Lervag, Redick, & Hulme, 2016). This study showed reliable training-induced improvement on trained tasks, but very small or no effects for tasks close to the trained tasks (mean effect size: < 0.5) or other tasks involving WM abilities such as nonverbal reasoning abilities or verbal abilities (mean effect size approaching 0). It should be noted that the effects were however somewhat larger in older adults as compared to young adults. Also, a recent study showed that, when separating participants according to their WM capacity, WM training does not lead to greater training effects in those participants that initially had the lowest WM capacity; rather, the reverse was true, with high performers becoming even better (Foster et al., 2017).

These studies indicate that the training of different WM components at the same time, in a non-targeted and non-differentiated manner, is not highly effective. To some extent, this is not surprising given the multitude of components and processes that characterize WM as discussed earlier. In aphasic patients, we have seen that WM deficits often stem from impairment to selective components of WM, calling for a component-oriented and process-oriented treatment strategy. A patient showing verbal WM impairment due to reduced semantic-level control processes is not likely to benefit from a WM training program that simply focuses on increasing WM storage capacity by administering a mix of nonword repetition or visuo-spatial WM training tasks. Likewise, a patient presenting difficulties in maintaining phonological item information may not benefit from repeated exposure to word list immediate serial recall tasks that focus on semantic retention abilities. This logic is also supported by one of the treatment studies in healthy adults reported earlier. Foster et al. (2017) trained control of attention and focus of attention aspects of WM using separate treatments, and they observed that, while participants within each treatment group showed improvement for the respective task and component on which they were trained, there was no generalization of training effects to the non-treated component.

4. The rehabilitation of WM in aphasia

Given the heterogeneity of verbal WM impairment in aphasia, it will be very difficult to find large groups of patients that would present exactly the same type of underlying cognitive impairment. Therefore, WM rehabilitation studies in aphasia have mostly adopted single or multiple case study designs. These designs are optimally suited for developing individualized, process-oriented WM training programs informed by the patient's nature of impairment. These designs can be very powerful as they directly target the cognitive component that is impaired, and thus should be associated with the strongest rehabilitation effects, as compared to generalized WM training programs. However, in order to convincingly demonstrate the efficacy and specificity of these individualized training programs, rigorous single case study methodological designs need to be put in place (Lane & Gast, 2014; Salis, Hwang, Howard, & Lallini, 2017; Shrubsole, Worrall, Power, & O'Connor, 2017). These studies need to have stable measures of baseline, pre-

therapy performance, to demonstrate the specificity of the intervention procedure by collecting measures for both target and non-target cognitive processes or by using control treatment procedures, and to assess the ecological validity of the rehabilitation program.

Two studies have provided qualitative reviews of single-case WM treatment studies (Murray, 2012; Salis, Kelly, & Code, 2015). These reviews, based on 4 and 11 single-case studies, respectively, concluded that WM treatment in aphasic patients seems to lead to improvement in both WM and language processing domains, but also that the robustness of the evidence is, at the least, uncertain. The aim of the present review paper was to provide a quantitative evaluation of the robustness of treatment effects in 15 single case or multiple case WM treatment studies (24 patients) published up to August 2017. This was achieved by computing effect size estimates of patient treatment effects, allowing to obtain a numerical estimate of treatment efficacy and its size in terms of change of performance on treatment and non-treatment measures. This also allowed for a comparison of treatment effects between the different single-case studies. Furthermore, the clinical significance of these treatment effects was assessed by determining to what extent the patients' performance levels reached normal levels after treatment. Another, specific aim of this review was to assess the content validity of the treatment studies, by analyzing the extent to which the treatment studies are motivated at a theoretical level and target specific WM components and processes as discussed in the previous section. As we have seen, WM deficits in aphasia are marked by a large heterogeneity in terms of cognitive components involved. The theoretical precision of assessment and treatment procedures is likely to be an important factor determining the efficacy of WM treatment studies.

For the quantitative estimation of effect sizes, effect sizes of treatment studies were computed by distinguishing between near transfer measures (all measures other than the treatment task involving the component that has been treated, i.e. different types of WM tasks¹), far transfer measures (measures of aspects that are not directly trained but which are impaired due to reduced WM abilities and hence should also benefit from WM training such as sentence comprehension) and control measures (measures of aspects that should not change as a result of the target treatment, or target measures after control treatment). Effect sizes were calculated using the d_1 effect size estimator proposed by Beeson and Robey (2006) specifically for single case studies: $d_1 = (\text{mean}_{\text{post}} - \text{mean}_{\text{pre}}) / \text{standard deviation}_{\text{pre}}$. This effect size estimator is based on Cohen's d (Cohen, 1988) but mean treatment change is compared to pretreatment variance only rather than to variance over both pre- and posttreatment periods. This estimator has been considered by Beeson and Robey as being the most reliable estimator of treatment effect size based on a Monte Carlo simulation comparing seven possible estimators of change including Cohen's d . The effect size estimates were computed on patient scores for which at least two pretreatment and two posttreatment measures were reported in the reviewed studies. In order to avoid bias in the calculation of effect sizes due to measures using different scales and hence leading to inflated standard deviations, all measures were standardized by calculating accuracy scores; for span measures, this was achieved by proportionalizing reported span scores relative to the maximum span level for a given task. When reviewing the treatment studies, the effect sizes will be described in a quantitative manner by determining whether the effect size reflects a treatment effect larger than 0.5, 1 or 2 standard deviations of pretreatment performance ($SD_{\text{pretreatment}}$). I will not refer to Cohen's benchmarks for considering effects as small or large as these benchmarks were established for group studies and between-participant variance, and they are difficult to apply to the context of single case studies limited to within-participant variance (Beeson & Robey, 2006). The broader significance of the observed effect sizes relative to effect size benchmarks will be considered in the Discussion section.

For the assessment of the clinical significance of treatment change, posttreatment Z-scores for target measures were computed by comparing the patients' posttreatment scores to the means and standard deviations of control data; if sufficient information was not available for computing Z-scores, any other type of information allowing to qualify a given patient's posttreatment performance (such as reports in the text that the patient showed normal range performance) was retrieved.

Finally, for assessing the content validity of the treatments, studies focusing on the treatment of verbal item retention versus serial order retention abilities (both aspects being in some studies also referred to as 'short-term memory') were distinguished, as well as studies focusing on the treatment of executive and/or attentional aspects of WM; one final study had used a compensatory treatment technique. This subdivision of treatment strategies reflects the type of WM components that have been addressed, in an explicit or implicit manner, in the treatment studies being reviewed here.

4.1. Treatment of verbal item retention capacity

One of the first studies that aimed at improving verbal short-term storage capacity in an aphasic patient is the study by Francis, Clark, and Humphreys (2003). Their patient presented with a history of mixed aphasia and severe short-term verbal retention and sentence comprehension deficits (see Table 1 for more details). Although no detailed theoretical investigation of the WM deficit was conducted, the authors considered that their patient presented with a verbal WM impairment as indicated by severely reduced spans across several digit and word span tasks. The authors further suggested that this impairment could at least partially explain the important sentence comprehension difficulties in this patient. The treatment consisted in intensive sentence repetition exercises the patient had to carry out twice per day during five days of a week. The sentences were new for each exercise which contained 12 to 20 sentences. No cues were given and the patient had to repeat each sentence immediately after its auditory presentation. The sentences progressively increased in difficulty, from two to seven words, with the most difficult sentences further containing reversible structures, semantically confusable items or low-imageability words. After the treatment, the patient showed improved verbal spans and sentence repetition abilities, with pre-post change slightly larger than 1 $SD_{\text{pretreatment}}$ on near transfer measures (see Table 2);

¹ Note that in the review by Melby-Lervåg et al., 2016, these measures were considered as intermediate-transfer measures.

Table 1
Quantitative characteristics of reviewed studies.

Study	N/Age	Aphasia type	Lesion site ^d	Time since onset		N pre-/posttreatment measures ^a	
				(months)	(weeks)	near transfer	far transfer control
Francis et al. (2003)	1/69y	Mixed aphasia ^b	left parieto-temporal	18	85/17	8/6	7/7
Koenig-Bruhin and Studer-Eichenberger (2007)	1/47y	Conduction aphasia ^b	left MCA		31/17	8/8	1/1
Majerus et al. (2005)	1/50y	Conduction aphasia ^b	left superior temporal	36	128/64	4/4	5/5
Van der Kaa and Majerus (2007)	2/73y; 74y	Broca's aphasia; expressive aphasia	left inferior frontal; right striatum	> 36	20/20	3/3, 2/2	1/1, 1/1
Harris et al. (2014)							2/2, 3/3
Kalinyak-Fliszar et al. (2011)	1/55y	Conduction aphasia ^c	left temporal	29	137/46	9/9	21/20
McCarthy et al. (2017)	1/34y	Conduction aphasia ^c	left temporo-parietal	36	40/20	10/13	12/12
Berthier et al. (2014)	3/51; 52; 72	Conduction aphasia ^c	left fronto-temporo-parietal	17, 22, 13	40/8	11/11	2/2
Mayer and Murray (2002)	1/62y	Fluent aphasia, alexia ^c	left thalamus	> 36	11/	1/1	1/1
Vallat et al. (2005)	1/53y	Conduction aphasia ^b	left parietal	14	72/24	12/8	3/2
Vallat et al. (2014)	1/38y	Conduction aphasia ^b	left SA	48	152/76	4/2, 4/2, 6/3	4/4
Zakarias et al. (2016)	3/57; 63; 64	Transortical motor aphasia ^c (2); anomic aphasia	left fronto-temporal; left MCA (2)	12, 12, 8	13/4	4/4	1/1
Murray et al. (2006)	1/57	Conduction aphasia ^c	left temporo-parietal	> 36	> 50/11	12/12	4/4
Salis (2012)	1/73	Transcortical motor aphasia ^c	left fronto-temporal	60	26/13	5/3	3/3
Salis et al. (2017)	5/55; 60; 47; 86; 68	Chronic aphasia ^b	/	180, 48, 8, 84, 108	27-30/	5/5-6/6	1/1-2/2
Closset and Majerus (2007)	1/10	Anomic aphasia ^b	left temporo-parietal	7	16/12	7/11	1/1

MCA: middle cerebral artery territory; SA: sylvian artery territory.

^a Number of measures that were retained for calculation of effect size; measures reflecting ceiling effects at pretreatment or measures that could not be clearly categorized as near/far transfer tasks were not included.

^b Aphasia type based on standardized aphasia batteries other than WAB and/or clinical judgment.

^c Aphasia type based on WAB classification.

^d All lesions were associated with a cerebrovascular accident, except for the patient reported in Closset and Majerus (2007) who presented with open traumatic brain injury.

Table 2
Nature of treatment approaches, treatment effect sizes (d_1 ; Besson & Robey, 2006) and severity of impairment (Z -scores) after and before treatment.

Study	Type of Therapy	Treatment effect size (d_1)/Clinical significance			Ecological validity
		Near transfer	Far transfer	Control	
Studies treating verbal item retention capacities Francis et al. (2003)	Sentence repetition	1.08/normal range ^a (digit span; word ISR; sentence repetition)	0.54/normal range ^a (sentence comprehension)	0.08 (naming/synonym judgment/ nonword repetition)	Patient more confident in her language abilities
	Koenig-Bruhin and Studer-Eichenberger (2007)	0.67 (digit and word span; probe recognition)	0.64* (sentence repetition)	/	Increased sentence length in sentence production
	Majerus et al. (2005)	1.44/normal range: $Z_{post} = -0.78$ ($Z_{pre} = -2.55$) (digit, word and nonword span)	0.34/normal range ^a (word-nonword repetition/ rhyme judgment)	-0.20 (minimal pair discrimination)	Improved communication abilities; Professional activities
	Van der Kaa and Majerus (2007)	0.87 (patient DS: semantic recall and probe tasks)	no sufficient data	0.88 (patient DS: phonological tasks after control treatment)	Improved sentence comprehension in patient with semantic item WM impairment
Harris et al. (2014)	Nonword and word list repetition	1.05 (patient AK: phonological recall and probe tasks)		0.12 (patient AK: semantic tasks after control treatment)	Modest and inconsistent improvement on a battery of input and output language processing tasks
Kalinyak-Fliszar et al. (2011)	Single Cued word and nonword repetition after variable delay	2.39 (follow-up: 2.42) (single word/nonword repetition for treated items)	0.55 (follow-up: 1.08) (single word/nonword repetition for non-treated items)	na	
McCarthy et al. (2017)	word-pair repetition with enhancement of semantic/syntactic activation	5.80 (follow-up: 3.92) (isolated word repetition after semantic activation)	0.30/impaired: $Z_{post} = -4.33$ ($Z_{pre} = -5.08$) (word and nonword repetition and probe span tasks)	0 (word repetition for untreated words)	Improvement on formal ecological questionnaire
Berthier et al. (2014)	Sentence repetition Control treatment	2.77 (follow-up: 2.16) (isolated word repetition after syntactic activation)	0.19/impaired: $Z_{post} = -3.99$ ($Z_{pre} = -4.28$) (word and nonword repetition and probe span tasks)	0 (word repetition for untreated words)	No investigation of ecological validity
Studies treating attentional/executive aspects in storage and processing tasks Mayer and Murray (2002)	Sentence judgment and maintenance	0.89, 0.99, 0.93 (word/nonword list repetition)	/	0.44, 0.46, 0.49 (word/nonword list repetition)	No functional improvement reported by patient
	Verbal storage and processing tasks	1.71, 3.89, 7.07 (sentence repetition)		1.11, 2.48, 5.66 (sentence repetition)	Improvement on formal ecological questionnaire
Vallat et al. (2005)	Verbal storage and processing tasks	Reduction of errors in a complex auditory-verbal WM task/normal range: $Z_{post} = -1.00$ ($Z_{pre} = -3.39$) 1.57/normal range: $Z_{post} = -0.02$ ($Z_{pre} = -2.19$) (digit and word span; backward digit span; Brown-Peterson task)	Improved reading efficiency and comprehension	0.00 (Visual\\\\\\Verbal long-term memory\\\\\\Verbal fluency)	Improvement on formal ecological questionnaire
Vallat et al. (2014)	Verbal and visuo-spatial storage and processing tasks	10.41 (T1)/normal range: $Z_{post} = -0.03$ ($Z_{pre} = -2.06$) (digit span forward/backward)	1.87 (T1), 1.87 (T2), 2.73 (T3) (Verbal comprehension/	10.41 (T2), 11.64 (T3) (digit span forward/backward) 1.74 (T1), 4.63 (T3)	Improvement on formal ecological questionnaire (continued on next page)

Table 2 (continued)

Study	Type of Therapy	Treatment effect size (d ₁)/Clinical significance		Ecological validity	
		Near transfer	Far transfer		
		<p>4.08 (T2)/normal range: $Z_{post} = 1.21$ (visual span forward/backward) $Z_{pre} = -1.86$</p> <p>8.64 (T3)/normal range: $Z_{post} = -0.11$ ($Z_{pre} = -6.60$) (verbal and visual Brown Peterson task)</p> <p>0.67, 0.67, -0.22 (N-back tasks)</p> <p>0.40/normal range^a (digit span, visual span, listening span)</p> <p>0.88/normal range (test of everyday attention)</p>	<p>problem solving) 0 (T1), 0 (T2), 2.59 (T3) (n-back, divided attention)</p> <p>0.48 (sentence processing)</p> <p>0.60 (sentence comprehension)</p> <p>0, -0.69, 0, 0.27, 0.41 (sentence comprehension)</p> <p>/</p>	<p>(visual span forward/backward) 3.45 (T1), 6.48 (T2) (verbal and visual Brown Peterson task)</p> <p>0, (+), 0 (Naming)</p> <p>0.03 (language profile)</p> <p>0.22 (long-term memory tasks)</p> <p>0.57 (phonological discrimination\\ \\ single word comprehension)</p> <p>0.00, 0.00, 0.00, -0.70 (visual span forward and backward)</p> <p>0.85/impaird: $Z_{post} = -5.40$ ($Z_{pre} = -5.45$) (nonword/abstract word span and rhyme probe span tasks) 1.73/lower normal range: $Z_{post} = -1.19$ ($Z_{pre} = -2.43$) (word span tasks - without recoding)</p>	<p>No information</p> <p>No change on everyday attention questionnaire</p> <p>Modest improvement of verbal comprehension</p> <p>No or modest improvement in formal psycho-social assessment</p> <p>Active use of recoding strategy in everyday life</p>
Zakarias et al. (2016)	N-back training				
Murray et al. (2006)	WM and attention tasks (APT-II)				
<i>Studies treating serial order retention capacities</i>					
Salis (2012)	Serial order recognition				
Salis et al. (2017)	Computerized serial order recognition				
<i>Studies using compensatory strategies</i>					
Closset and Majerus (2007)	Visuo-semantic recoding				

* Rho trend test as reported in the study; + significant difference reported in the study; 0 no effect size could be calculated but the study reports an absence of significant change.

^a For at least one measure.

performance on far transfer sentence comprehension tasks also improved but with a smaller effect size (close to $0.5 SD_{\text{pretreatment}}$). Tasks for which no treatment effects were expected (picture naming, synonym judgment and nonword repetition) showed no change over the treatment period (see Table 1), suggesting that the treatment was both efficient and specific. Treatment also led to performance in the normal range for the digit span and some of the sentence comprehension measures. At the same time, given the nature of the treatment task (sentence repetition) there remain questions about the specific WM processes that had been treated. The authors argue that their treatment targeted more ‘dynamic’ WM processes as involved in sentence comprehension, given that nonword repetition did not improve, and hence phonological item retention processes remained impaired. More generally, it should be noted that sentence repetition involves additional linguistic components such as syntactic processing and integration of sentence-level conceptual representations with syntactic, semantic and phonological levels of representation of individual words (Potter & Lombardi, 1998; Schweppe, Rummer, Bormann, & R. Martin, 2011). In terms of the theoretical framework adapted in this review, the therapy could be considered to have involved, at the very least, the training of semantic item retention and control processes. In sum, although the study shows that sentence repetition can be an efficient method for improving performance in verbal WM and sentence processing tasks, the precise nature of the WM component(s) that were trained remains an open question.

A second study used a word list repetition procedure in a patient with conduction aphasia for treating the patient's verbal WM impairment (Koenig-Bruhin and Studer-Eichenberger 2007). The authors interpreted their patients' verbal WM impairment as reflecting a storage deficit, taking the form of a pathologically increased decay rate of phonological and semantic item representations (see Table 1). They hypothesized that these decay-based item retention difficulties had wider consequences on sentence production abilities which were also impaired in this patient. The rehabilitation procedure consisted in the repetition of subject-verb-object sentences (4–7 words) of increasing length, immediately and after a variable delay (5–12 s). Novel sentences were used for each session. When the patient made errors during the immediate repetition of the sentences, the therapist helped the patient by repeating the sentences. The treatment procedure entailed the use of multiple baselines, by collecting performance across word span and probe recognition WM tasks. The authors observed improved performance in (untrained) WM tasks amounting to less than $1 SD_{\text{pretreatment}}$ (see Table 2), and they also observed significant improvement in an (untrained) sentence repetition task as determined by a Rho trend test. The mean length of utterance further increased slightly for sentence production, from 5.26 words to 6.41 words. As regards clinical significance, no control data are reported, but digit span values at posttreatment (digit span = 4; estimated $Z = -1.92$) were at the lower end of control range as compared to norms for patients of a similar age as reported in Van der Kaa and Majerus (2007). Like the study by Francis et al. (2003), this study indicates that word list repetition training in the form of sentence repetition can improve performance in auditory-verbal WM and sentence repetition tasks. This study does however not allow to determine the specificity of the treatment procedure as the authors did not present tasks on which the patient would have been impaired and which should not have benefitted from WM training. It must be noted here that the patient showed relatively preserved performance for most aspects of language processing except for sentence production.

Another study using a similar treatment procedure has been conducted by Majerus, Van der Kaa, Renard, Van der Linden, and Poncelet (2005) (see also Van der Kaa and Majerus (2007) for a more detailed description of the patient and treatment procedure). This patient presented with severely reduced verbal WM abilities and this particularly for the retention of phonological item information, as revealed by a profound impairment for nonword span tasks. The stability of the verbal WM impairment was supported by the fact the patient's digit, word and nonword spans had shown the same level of impairment over a 4-year period preceding the start of the WM treatment phase, a period over which the patient had intermittently received speech and language therapy for her initial naming impairment. The treatment approach consisted in the delayed repetition of verbal information as in the previous studies, but focused on the repetition of two-nonword lists and thereby directly targeted phonological item retention processes. The nonword lists were minimal pairs (roba-soba) requiring the patient to maintain a detailed phonological representation of the target nonwords. At the same time, the phonological redundancy of the two nonwords also facilitated maintenance as the phonological information of the first item could be used as a cue for the second nonword. The nonword pairs were trained in sets of 10 pairs. The patient had to immediately repeat each pair after its presentation, without receiving any cues; when a set of items could be successfully repeated at 90% accuracy, training continued with a novel set of 10 nonword pairs. After all the sets ($N = 25$) had been successfully completed, the same sets were presented again but this time for delayed repetition: after presentation of a nonword-pair, the patient had to count from 1 to 3 at the rate of one digit per second, and then repeat the nonword-pair. At the end of the treatment, the patient showed normalized performance in digit, word and nonword span tasks, with the most important increase in nonword span, where her performance increased from 0 to 3. The treatment effects for (untrained) auditory-verbal WM tasks exceeded $1 SD_{\text{pretreatment}}$ (see Table 2). Treatment effects of somewhat smaller size were also observed for language processing tasks that involve temporary maintenance of phonological information such as single nonword repetition, low frequency word repetition and rhyme judgment. No improvement was observed for a speech discrimination task for which the patient had shown weak performance levels, indicating that the treatment was specifically associated with improvement of phonological short-term retention but not perceptual and phonological analysis abilities. Although the ecological validity of the treatment had not been formally assessed, the patient reported that after treatment, she could again follow group discussions, take notes on the phone and have professional activities.

A further study implemented a cognitive neuropsychological analysis in two patients with globally reduced verbal WM spans in order to determine in a detailed manner the nature of the patients' verbal WM impairment and to demonstrate the specificity of the treatment procedure using a particularly powerful design (Harris, Olson, & Humphreys, 2014). Based on a detailed linguistic and cognitive assessment, the authors considered that one of the patients (patient AK) presented with a phonological item retention deficit and that the other patient (patient DS) presented with a semantic item retention deficit. The authors submitted both patients to treatment procedures that targeted either phonological item retention (immediate repetition of nonword lists) or semantic item retention (immediate repetition of word lists). The number of nonwords/words in each list exceeded span length by 1, and were novel

for each treatment session; no cues were given but the experimenter provided corrective feedback at the end of each immediate repetition attempt. These exercises were completed by nonword/word probe recognition exercises at home. The authors hypothesized that only the patient with a semantic WM impairment should benefit from the semantic treatment, and that only the patient with a phonological WM impairment should benefit from the phonological treatment. These predictions were confirmed. As shown in [Table 2](#), phonological treatment led to slightly stronger treatment effects for phonological recall and probe recognition WM tasks in the patient with phonological retention difficulties than in the patient with semantic retention difficulties. The reverse situation was observed for semantic treatment effects on semantic recall and probe recognition WM tasks. For the patient with semantic retention difficulties impairment, although phonological and semantic treatments appeared to improve performance to a similar extent, it should be noted that the semantic treatment was administered after the phonological treatment; the increase of nearly 1 $SD_{\text{pretreatment}}$ after the semantic treatment is an additional increase relative to the phonological treatment effect. No effect sizes could be computed for far transfer measures due to an insufficient number of available data. It should however be noted that the patient with phonological retention difficulties showed increased performance accuracy from 0.60 to 0.92 on a sentence repetition measure after phonological treatment while the patient with semantic retention difficulties showed no increased performance (from 0.15 to 0.17). Conversely, the patient with semantic retention difficulties showed increased performance from 0.67 to 0.83 on a sentence comprehension measure after semantic treatment; the patient with phonological retention difficulties had shown near ceiling performance for this measure already at baseline. This study shows in a particularly compelling manner the importance of using WM treatment procedures that precisely target the WM component that is impaired. More general information about the ecological validity of the treatments proposed in this study was however not reported.

A study by [Kalinyak-Fliszar, Kohen, and Martin \(2011\)](#) also used a word and nonword repetition treatment approach but in a patient presenting not only with phonological item WM impairment (as indicated by particularly poor performance on nonword and low-imageability word span tasks), but also impairment at phonological and semantic input and output processing tasks. This patient presented a less specific short-term retention impairment than the patients in the studies reported as so far. This study used a particularly large set of baseline WM and language measures, allowing for a robust characterization of the patient's linguistic and WM impairment. As in the previous studies, the treatment procedure consisted in the repetition of word and nonword stimuli; however, unlike in previous studies, single word and nonword stimuli were presented, they had to be repeated after a 1-s or a 5-s delay, and, critically, they were cued when the repetition attempt was incorrect. A hierarchical cueing procedure was used by providing increasingly precise cues (cue 1: stimulus is re-presented; cue 2: the clinician repeats the patient's error and then presents again the target stimulus; cue 3: the clinician indicates the erroneous phoneme/syllable and presents the correct one) and by asking the patient to repeat the stimulus again after cueing. This intervention targeted both the relearning of phonological representations and their maintenance for a set of 10 words and 20 nonwords. Given that this intervention targeted specific items and their underlying representations more directly, I considered single word/nonword repetition tasks using treated stimuli as near transfer measures, and single word/nonword repetition tasks using untreated stimuli as well as word/nonword span tasks as far transfer measures. The results showed treatment effects exceeding 2 $SD_{\text{pretreatment}}$ for tasks using trained stimuli, but much smaller treatment effects for tasks using untreated stimuli or nontreated word/nonword span tasks. These results were confirmed at a 3-month follow-up, after the termination of treatment (numbers in parentheses in [Table 2](#)). Performance on word/nonword span tasks also remained strongly impaired, as indicated by mean Z-scores of posttreatment performance (see [Table 2](#)). This study did not report information about wider implications of the treatment on everyday functioning of the patient; at the same time, it should be noted that the patient showed improvement in several language processing tasks such as synonym and rhyme judgment tasks. In sum, this study contrasts with the previous studies, by showing a rather limited treatment effect on verbal WM functioning, but this is likely to be due to the stimulus-specific treatment procedure that was implemented for a patient with multiple sources of linguistic impairment.

A similar study using an item-specific treatment procedure was conducted by [McCarthy, Kalinyak-Fliszar, Kohen, and Martin \(2017\)](#) in a patient with deep-phonological dysphasia characterized by severe single word and nonword repetition impairment. The patient showed marked imageability and word frequency effects in word repetition performance. Based on the N. [Martin et al. \(1996\)](#) account of deep dysphasia, the authors considered that the patient's deficits stem from pathologically increased decay rates of the activation of phonological and semantic item representations. This deficit is considered to lead to particularly severe difficulties for items that cannot be activated easily at the semantic level and for which repetition relies mainly or exclusively on the quickly decaying activations at the phonological level. The authors aimed at improving the patient's word repetition abilities by facilitating semantic access for 10 low imageability nouns using a semantic pre-activation treatment strategy. During the treatment period, the low imageability nouns were paired with semantically related adjectives aimed at narrowing down the semantic meaning of the nouns and facilitating their access; each adjective-noun pair had to be repeated immediately after presentation. This technique was supposed to increase the strength of semantic representations for the trained nouns and facilitate their access even when the nouns were subsequently presented for repetition without the adjective. A syntactic pre-activation comparison treatment procedure was also used, where 10 different nouns were coupled with adjectives unrelated to the meaning of the target nouns. Finally, a control condition assessed repetition of 10 nouns that were not associated with any semantic or syntactic primes. Given this item-specific and indirect procedure as regards the training of short-term maintenance of information, I considered verbal WM measures (word and nonword recall and probe span tasks) as far transfer measures while isolated repetition of treated items after semantic or syntactic treatment was considered as a near transfer measure. The results showed very strong training effects for both the semantic and syntactic pre-activation treatment procedures (treatment effects larger than 2 $SD_{\text{pretreatment}}$) but, as could be expected, this treatment was item specific as no treatment effect was observed for repetition of non-treated items. This was paralleled by modest improvement in word and nonword repetition and probe span tasks; the patient remained severely impaired on these measures after treatment (see [Table 2](#)). As the previous study, this study suggests that indirect, item-specific treatment procedures do not lead to strong and

generalizable improvement of short-term retention capacity.

A final study used a verbal repetition treatment procedure for treating repetition and verbal WM impairment in conjunction with pharmacological treatment in three patients with conduction aphasia (Berthier et al., 2014). Importantly, this study is one of the few studies comparing the target treatment to a control treatment which consisted in distributed multi-level speech language therapy. The target treatment procedure consisted in the massed repetition of a closed set of sentences ranging from 2 to 7 words, with 20 sentences inducing a rich semantic context (by containing many concrete nouns and verbs) and 20 sentences inducing a poor semantic context (by containing many pronouns and functions words). Treatment was done exclusively at home, the patients repeating the audiotaped sentences; hence, no cues or corrective feedback were given to the patients during training. As shown in Table 2, the treatment effects were generally larger after the target sentence repetition treatment than after the control treatment, and this for both word and nonword list repetition tasks and sentence repetition tasks. At the same time, for the word and nonword list repetition tasks, the treatment effect sizes remained relatively small ($< 1 SD_{\text{pretreatment}}$). Although no control data were provided, the results at the digit production task (digit span was equal to 3 in all three patients posttreatment) suggest that auditory-verbal short-term retention abilities remained below control range. Furthermore, given that the sentence repetition treatment was always administered after the control treatment, the effect sizes of the sentence repetition treatment actually reflect the combined treatment effects of the control and the target treatment. Also, all patients received a pharmacological treatment (donepezil) during both treatment procedures. In sum, the sentence repetition treatment appears to have been slightly more effective than the control treatment, but the absolute size of the treatment effects remains small and the effect that is due only to the behavioral treatments is difficult to determine.

In sum, studies using repetition treatment strategies appear to lead to robust treatment effects for verbal item retention abilities when using novel items over the different treatment sessions, when providing no specific cues, and when targeting the general function of verbal maintenance rather than the relearning of specific stimuli. This treatment approach can lead to clinically significant improvement of verbal WM functioning, with performance on word and nonword span tasks back to normal range levels in most cases. Transfer effects to untreated language tasks can also be observed. On the other hand, studies using item-specific training strategies may lead to improved repetition of treated items and their use in untreated language tasks, but they do not appear to lead to general improvement of verbal retention abilities.

4.2. Treatment of attentional/executive aspects in storage and processing tasks

Other studies used a different approach and focused on executive and attentional aspects of verbal WM impairment, without however defining the exact nature of the executive or attentional control processes that were targeted. Rather these studies trained performance in storage and processing tasks, for which it is assumed that there is a strong involvement of executive and attentional control processes (see also Section 3).

The first study using this type of approach was a study by Mayer and Murray (2002). The main focus of this study was to treat an acquired reading disorder in a patient with fluent aphasia. In addition to his reading disorder, the patient also presented with inconsistently impaired performance on verbal and visual WM tasks, leading the authors to consider that the patient had additional difficulties at the level of attentional control aspects of WM (Table 1). The authors implemented an alternating reading/WM treatment. The WM treatment involved training of storage and processing abilities, by presenting written sentences to the patient which he had to judge at the grammatical level, and for which he had to determine and maintain the semantic category of the final word. The semantic category of final words had to be recalled after a predetermined number of sentences had been presented and judged. The sentences, containing 6 to 10 words, were paired, with the last word of each consecutive sentence pair sharing the same semantic category. The sentences were presented by increasing sentence length, and by increasing the number of sentences (2, 4, or 6). In the easiest condition, only one semantic category name had to be maintained and recalled, while the most difficult condition involved the maintenance and recall of three different category names that characterized the end word of each sentence. This task is very similar to a complex reading span procedure (Daneman & Carpenter, 1980) and is considered to engage to a large extent executive and attentional control processes. Given the use of a very limited number of pretreatment and posttreatment tasks, the overall efficiency of the treatment, and particularly its effect size, are difficult to estimate. The authors reported an increase in reading efficiency and in WM capacity after treatment. Posttreatment WM performance was in the normal range. The specificity of the treatment cannot be assessed, given that the treatment approach consisted in the administration of reading and WM treatments in the same session, and hence it is difficult to determine which of the two treatments caused improvement in the WM and reading measures. Furthermore, the WM treatment involved reading sentences, and hence also involved training of reading abilities. At a functional level, the patient did not report significant changes in his reading abilities.

Vallat et al. (2005) focused more specifically on executive and attentional aspects in a patient with verbal WM impairment. Their patient had recovered from conduction aphasia (Table 1). A detailed cognitive investigation of his WM profile in the light of the Baddeley and Hitch (1974) working memory model had concluded that the patient suffered from deficits at the phonological loop and central executive components of WM. The stability of the deficit was attested by multiple testing and baselines. The treatment procedure consisted in the administration of a battery of verbal WM tasks that included both storage and processing components. The patient had to perform eight different verbal WM tasks such as reconstitution of words from oral spelling, oral spelling, calculation of the odd/even number of letters in a word, word assembly from isolated syllables, recalling lists of words by reordering the words in alphabetical order, and forming new words by assembling the initial phoneme of a series of words. The tasks were presented with increasing level of difficulty, by starting with 3-item lists and by increasing list length when 90% of trials for a given task were successfully completed. Following treatment, the authors observed treatment effects exceeding $1 SD_{\text{pretreatment}}$ for both near transfer

simple and complex WM tasks (see Table 2) as well as for far transfer tasks such as arithmetic problem solving and sentence comprehension tasks. Posttreatment performance levels for the verbal WM tasks were at control levels indicating that the patient had recovered normal verbal WM capacities. No change was observed on control measures such as verbal and visual long-term memory as well as verbal fluency tasks. It could be argued that a verbal fluency task also relies on control of attention and executive processes but at the same time, this task differs from the WM tasks and training tasks given that it does not require that retrieved information is maintained throughout the task. It should also be noted that the patient did not show significant deficits on these control tasks, and so there was not much room left for improvement. Finally, unlike most other studies, this study assessed the ecological validity of the treatment procedure in a very thorough manner by administering both WM and verbal communication questionnaires. This formal ecological assessment showed a significant reduction of patient-based complaints in everyday life after treatment. Notably, the patient was able to return to work for a full-time job after treatment.

Vallat-Azouvi, Pradat-Diehl, and Azouvi (2014) extended their treatment procedure to train phonological loop, visuo-spatial sketchpad and central executive aspects in a patient with conduction aphasia presenting with a severe, global WM deficit (see Table 1). They aimed at training the three components successively, by administering tasks considered to target specifically each of these components. For training the phonological loop at T1, they used the same tasks as described in their previous study. They further developed specific tasks for training the visuo-spatial sketchpad at T2 (2-D and 3-D mental imagery tasks in which the patient had to move mentally from one position to another on an imagined chessboard or cube; visual N-back tasks with playing cards, figures or geometrical forms as stimuli) and the central executive at T3 (more demanding N-back tasks in which the patient had to detect whether a given word was from the same semantic category as the target in position N-1 or N-2; reading span tasks). All tasks varied in level of difficulty by starting with a low level of difficulty and by progressing as a function of the patient's level of performance, as in their previous study. Multiple baseline measures were collected and tasks tapping the different components targeted by this study were re-administered after each of the three training steps. The results showed very large near transfer effects (see Table 2) at the end of the entire training program, with posttreatment performance having improved by at least 4 $SD_{\text{pretreatment}}$. There was some limited evidence for training specificity: visual span tasks showed the largest improvement after visual WM training (T2) while verbal span tasks did not improve after visual WM training. However, for storage and processing WM tasks, performance progressed continuously over the three training periods, and the smallest performance increase was actually observed after central executive training at T3. The patient showed performance in the normal range for all aspects of WM after the three treatment phases (see Table 2); performance in verbal comprehension, problem solving and other executive tasks also increased, although to a lesser extent than performance on the WM tasks. The patient also reported a diminished rate of WM related complaints in everyday functioning, as measured by an ecological questionnaire. It should be noted here that the tasks used for training the phonological loop and visuo-spatial sketchpad components (such as alphabetical re-ordering word span tasks and N-back tasks) also trained executive and attentional control processes. In a similar manner, the tasks used for assessing the storage components (phonological loop, visuo-spatial sketchpad) also included processing components such as backward recall verbal and visual span tasks. Hence, it is very difficult to determine, at each treatment step, whether performance increased due to the training of modality-specific retention capacities or due to the training of executive/attentional control processes involved in both the assessment and treatment tasks at all treatment stages. In sum, this study further shows that WM training can be highly effective but the specificity of the treatment in terms of the trained WM components is difficult to establish.

Other treatment studies targeted sentence comprehension impairment by training executive and attentional control processes via storage and processing treatment tasks. Zakarias, Keresztes, Marton, and Wartenburger (2016) used N-back tasks to train storage and processing components of WM in three patients with chronic aphasia. The N-back tasks were classical N-back tasks using letter stimuli; task difficulty increased progressively, by varying the distance between targets (N-1, N-2, N-3) and by including lures (letters identical to the target but occurring at a non-target position) or not. The patients showed posttreatment improvement smaller than 1 $SD_{\text{pretreatment}}$ or even poorer performance after treatment on untrained visual and auditory N-back tasks (see Table 2). Sentence comprehension abilities improved in two patients, including in the patient showing a negative treatment effect on the N-back tasks. The conclusions that can be drawn from this study are limited given that there was no information available about the nature of WM impairment in these patients, raising the possibility that WM deficits, if present, were not trained in a targeted manner. Furthermore, as acknowledged by the authors, there were no multiple baseline measures (see Table 1). Another study conducted by Murray, Keeton, and Karcher (2006) also targeted attentional control aspects of WM as well as attentional processing more generally in a patient with mild conduction aphasia. The authors trained different dimensions of attentional control in their patient, using auditory sustained, selective, alternating and divided attention tasks as well WM-type storage and processing tasks (such as giving back words from sentence stimuli in alphabetical order) (APTII, Sohlberg, Johnson, Paule, Raskin, & Mateer, 2001); the tasks progressed in difficulty as a function of the patient's level of performance. Multiple baseline and posttreatment measures were available for WM, attention and language tasks. The authors observed that after treatment, the overall aphasia severity had not changed. Performance increased somewhat more in attentional tasks ($< 1 SD_{\text{pretreatment}}$) than for WM tasks and for far transfer sentence processing abilities ($< 0.5 SD_{\text{pretreatment}}$; see Table 2). However the patient and his partner did not report any changes in everyday life functioning. Performance in attentional and WM tasks was in the normal range for the majority of measures after treatment. For sake of completeness, note that two studies by Coelho (2005) and Sinotte and Coelho (2007) used the same treatment procedure as Murray et al. (2006) in order to improve attention and reading abilities in a patient with mild anomic aphasia; a further study by Peach, Nathan, and Beck (2017) also adopted a similar approach in four patients with fluent or non-fluent aphasia. These three studies are however not reviewed here as no detailed information was available about the patients' WM performance levels.

In sum, the studies targeting treatment of executive and attentional control processes in WM tasks show overall robust treatment effects, with patient performance reaching normal levels in many cases after treatment. At the same time, the specificity of these

treatments is difficult to determine as the nature and involvement of attentional/executive control processes in both target and control tasks is not precisely defined.

4.3. Treatment of serial order retention capacities

A further single case study used a more specific training procedure, by using a serial order recognition treatment strategy (Salis, 2012) in order to improve verbal short-term retention abilities and sentence comprehension abilities in a patient with transcortical motor aphasia. Salis observed impairment in both verbal and visual span tasks in his patient, together with sentence comprehension deficits. The treatment consisted in the presentation of pairs of word sequences with the order of the words being either identical or not, and the patient had to determine whether the order of the words in the two sequences was identical or not. The treatment task increased in level of difficulty by starting with lists of 5 monosyllabic words, and by ending with lists of 3 polysyllabic words (3–4 syllables). When the patient made an error, corrective feedback was provided by pointing out the words that had exchanged their serial position. It should be noted that this treatment approach had been selected not because the patient would have presented a specific serial order retention impairment (this aspect had not been formally assessed in this study), but because it allowed the authors to circumvent the patient's important speech output difficulties (see Table 1). Relatively large treatment effects were observed for verbal span tasks ($> 2 SD_{\text{pretreatment}}$; see Table 2). Far transfer treatment effects for sentence comprehension measures were smaller and they were of the same size as those observed for the phonological discrimination and single word comprehension control tasks ($< 1 SD_{\text{pretreatment}}$). Performance in forward verbal span tasks increased to normal levels, while performance in backward verbal span remained unchanged, in line with the treatment strategy which targeted storage rather than processing components. In sum, this study shows that WM treatment can be efficient in patients with aphasia and can be limited to the storage component. The specificity of the treatment approach as regards the storage of serial order information more specifically is however more difficult as a detailed assessment of serial order versus item retention capacities was not included in this study.

Salis et al. (2017) conducted a replication study of the Salis (2012) study in five patients with chronic aphasia and impairment in verbal or verbal and visual span tasks. This study was a methodological improvement over their initial training study as the authors used a computerized training procedure and assessed psychosocial outcome in a formal manner. The treatment procedure was otherwise very similar to the initial study, except that only monosyllabic words were used for the memory lists. Also initial treatment list length was determined by the patients' pretreatment span levels, and list length was increased by one if the patient achieved 80% correct trials for a given list length. The specificity of the treatment could also be assessed to some extent given that verbal and visual spans were measured at both pretreatment and posttreatment, while treatment only consisted in auditory-verbal serial order recognition tasks. The authors observed relatively large near transfer treatment effects in two patients ($> 3 SD_{\text{pretreatment}}$ and $> 1 SD_{\text{pretreatment}}$) but much smaller treatment effects in the other 3 patients ($< 0.5 SD_{\text{pretreatment}}$). The treatment effects were specific in the sense that performance on visual span tasks did not change, or even decreased. Transfer effects were very limited as performance on sentence comprehension measures improved very minimally ($< 0.5 SD_{\text{pretreatment}}$) or even decreased in one patient. In line with these results, formal measures of psycho-social outcome showed very limited improvement. This study contrasts with the studies reviewed so far by observing more limited near transfer WM treatment effects, and virtually no transfer effects on target language measures. At the same time, it should be noted that WM treatment was not informed by an in-depth theoretical analysis of the individual patients' WM profiles, and hence the study may not have specifically targeted the WM components that were actually impaired in the patients. Also, some of the patients showed digit spans in the normal range when assessed using a pointing procedure, raising the possibility that they showed motor speech impairment rather than WM deficits; indeed, all patients showed signs of apraxia of speech. A similar argument had also been proposed by the authors for explaining the absence of far transfer effects on sentence comprehension measures. They suggested that the nature of sentence comprehension deficits may not have been uniform across patients and could have stemmed from other deficits than WM deficits.

4.4. Compensatory treatment strategies

A final study used a compensatory strategy instead of the restoration approaches that characterized all the studies reviewed so far. Closset and Majerus (2007) reported a young aphasic patient (10 years) with mild anomic aphasia and a severe phonological item retention deficits characterized by particularly severe impairment for tasks involving the maintenance of nonwords. Performance on word span tasks was also significantly impaired, although to a lesser extent and performance on visuo-spatial span tasks was fully preserved. Multiple baselines were obtained. Given the patient's preserved semantic and visuo-spatial storage abilities, the treatment strategy aimed at favoring a visuo-semantic recoding strategy for auditory-verbal memoranda. The patient was instructed to form visual mental images for each successive word presented within an auditory word list, and to maintain the sequence of visual mental images over maintenance delays ranging from 0 to 60 s. The patient then had to recall the word lists, by reviewing and naming the mental images she had created. Given the patient's mild anomia at the moment of the study, the treatment was carried out on a set of 60 words which the patient could name easily. Before applying this strategy, the patient's ability to form mental images from auditory-verbal input had also been verified. After 16 therapy sessions, the patient was able to quickly recode auditory-verbal information in a visual format and to use the mental images to maintain and recall word lists over delays up to 60 s. This led to large treatment effects ($> 4 SD_{\text{pretreatment}}$) for those WM tasks in which the learnt strategy could be used: performance improved for tasks where information could be retained using semantic codes such as recall of concrete word lists (these tasks included different words than those used for training) (see Table 2). Performance for these tasks was at normal levels after treatment. As predicted, the treatment effect was much more limited for WM tasks where this strategy could not be used such as nonword or abstract word list

recall ($< 1 SD_{\text{pretreatment}}$). Furthermore, when the patient was instructed not to use the visuo-semantic recoding strategy, the treatment effect observed for word list immediate serial recall also dropped, indicating that the patient was able to flexibly use the strategy she had learnt. The patient also improved her sentence comprehension abilities, which went from severely impaired ($Z_{\text{pre}} = -20.00$) to lower normal range after treatment ($Z_{\text{post}} = -1.25$). The treatment effects were maintained over a 5-year follow-up period. When asked about the use of the visuo-semantic recoding strategy in everyday life, the patient reported that she used the strategy whenever possible and that this strategy had been very helpful for her studies. She complained about the fact that this strategy was not applicable in all situations (abstract information, novel words) and one of her biggest challenges remained learning foreign languages.

4.5. Aggregated analysis of single-case effect sizes

In a final section of this review, the effect sizes of the reviewed treatment studies were analyzed at an interindividual level. The individual effect sizes observed for near transfer, far transfer and control tasks/treatments were aggregated and submitted to a Bayesian analysis determining the evidence in favor of a treatment effect as compared to the null model (Lee & Wagenmakers, 2013). When the same tasks were successively considered as near transfer and control measures due to the use of multiple treatment types (as in the study by Vallat et al., 2014), differential effect sizes were computed in order to obtain effect sizes for near transfer and control treatments that are unbiased by the preceding control/target treatment. Bayesian statistics were used given their robustness in case of small-to-moderate sample sizes and non-normal distributions (Moore, Reise, Depaoli, & Haviland, 2015). A Bayesian factor (BF_{10}) greater than 3 was considered as providing moderate evidence in favor of a treatment effect, a BF_{10} greater than 10 as providing strong evidence, a BF_{10} greater than 30 as providing very strong evidence, and a BF_{10} greater than 100 as providing decisive evidence (Lee & Wagenmakers, 2013).

A Bayesian one-sample t -test revealed decisive evidence in favor of a treatment effect for the near transfer measures, with $BF_{10} = 281.00$ (error %: $2.487e - 8$). As shown in Fig. 2, a mean effect size of 1.75 was observed and the 95% credible interval (Bayesian equivalent of confidence intervals) clearly excluded zero. At the same time, evidence for far transfer treatment effects against the null was lower, $BF_{10} = 14.41$ (error %: $1.706e - 6$), but still reflected strong levels of evidence and was characterized by a 95% credible interval excluding zero. It is important to note here that the number of measures reporting far transfer effects is considerably lower than the measures reporting near transfer effects (see Table 1). Finally, control measures were associated with moderate evidence in favor of a treatment effect against the null, with $BF_{10} = 9.91$ (error %: $2.487e - 8$). This moderate evidence for control treatment effects is partially driven by studies that targeted training of executive and attentional aspects of WM which led to an aggregated mean control treatment effect size of 1.14 (range: 0.00–3.45) versus 0.73 (range: -0.7 –5.66) for the other intervention strategies. Treatment effects on control measures were also driven by the control treatment in the study by Berthier et al. (2014) which targeted many different language components and processes at the same time and which already led to relatively large treatment effects on verbal WM and sentence repetition measures before implementation of the WM treatment.

5. Discussion

The aim of this review was to provide a systematic analysis of the single case and case series studies on WM treatment in aphasia published until August 2017, with a specific focus on the quantitative assessment of treatment effect sizes and clinical significance. Fifteen studies involving 24 patients with aphasia could be identified. The vast majority of these studies fulfilled minimal methodological requirements by including multiple and/or stable baselines and by including control tasks or control treatments. Studies are also increasingly concerned by the ecological validity of the treatments, and seven studies provided a formal or informal assessment of the impact of the WM treatment on everyday life activities.

The results of this quantitative review indicate that WM treatment in patients with aphasia can lead to robust treatment effects.

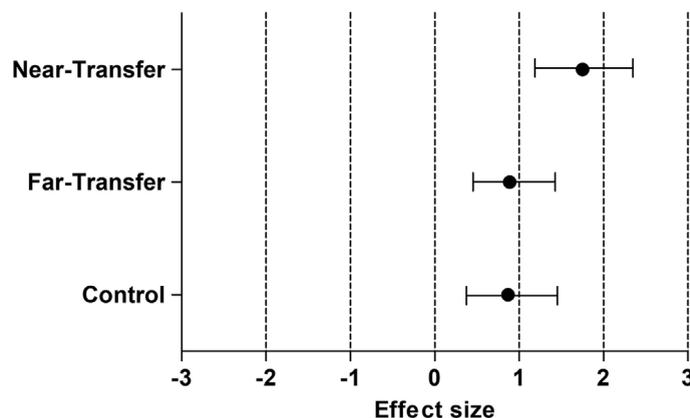


Fig. 2. Mean effect sizes and 95% credible intervals for near transfer, far transfer and control measures of reviewed studies.

The mean patient-based effect sizes observed are close to 2 (based on the d_1 estimator of effect size; Beeson & Robey, 2006). This number appears to be larger than the mean inter-individual effect sizes observed in group-based studies of healthy populations for which treatment effect sizes are often less than 0.5 (e.g., Melby-Lervag et al., 2016). At the same time, it is difficult to directly compare intra-individual effect sizes in patients and group-based effect sizes in healthy participants. The benchmarks for interpreting Cohen's d were developed for interpreting group-based effect sizes characterized by inter-individual variance. The treatment effect sizes observed in this review would be characterized as very large treatment effects according to these benchmarks (Sawilowsky, 2009). On the other hand, treatment effects in individual patients may generally be associated with larger treatment effects as pretreatment performance is often severely impaired and hence there is lots of room for improvement. Beeson and colleagues proposed benchmarks for aphasia treatment by reviewing 12 single case treatment studies for alexia and agraphia. They proposed, on the basis of the distribution of effect sizes observed in these 12 studies, that an effect size of 2 would be a small effect size (Beeson & Egnor, 2006; Beeson, Magloire, & Robey, 2005). However these more specific benchmarks only allow to determine whether a given effect size is large relative to other studies having treated the same type of linguistic impairment in the same type of patients. Different benchmarks have indeed been obtained for studies targeting syntactic production or lexical retrieval treatments (Beeson & Robey, 2006). Hence, the existing benchmarks cannot easily be generalized to other treatment targets such as verbal WM.

Moreover, treatment effects for a given patient may be considered small (relative to other aphasic patients with a similar type of deficit) and they may nevertheless reflect clinically significant improvement. Conversely, a large treatment effect size does not necessarily reflect a significant improvement in WM performance, as minimal absolute changes in scores can lead to large effect sizes when pretreatment variance is very small. For these reasons, a quantitative analysis of the clinical significance of the treatment effects was also conducted in this review by determining to what extent the treatment had allowed to bring performance back to normal range levels. For the 10 studies where sufficient information about normal range performance was available, 9 studies reported posttreatment patient performance that was back to normal range performance. Posttreatment Z-scores were within 1 standard deviation of control performance in seven out of nine cases for which these scores could be calculated. These data confirm that the treatment effect sizes observed in this review reflect in many cases clinically significant changes in WM performance.

On the other hand, this review also shows that the specificity of WM treatment strategies in aphasia is more difficult to establish. Treatment effects for far transfer and control measures were comparable, although being overall smaller than treatment effects for near transfer measures. This apparent lack of specificity is likely to be related to the often very broad and superficial definition of WM impairment, as the vast majority of studies did not carry out a detailed analysis of the nature of WM impairment in their patients. This makes it also difficult to assess the content validity of the proposed WM treatments. As discussed earlier, WM is not a unitary construct but comprises many different subcomponents and processes which can be selectively impaired. Targeting more precisely these processes may lead to treatment strategies that show stronger content validity, which in turn could result in both improved specificity and efficacy (see also Salis et al., 2015; for a similar conclusion). The study presented by Harris et al. (2014), distinguishing between phonological item and semantic item WM impairment demonstrates in a particularly convincing manner the importance of a detailed cognitive analysis of WM impairment and the implementation of treatment strategies whose content is theoretically informed by this cognitive analysis. As further illustrated by the study of Harris et al. these detailed cognitive analyses also allow for more time-efficient cross-over designs. These designs allow to directly compare two treatments which are at the same time target and control treatments depending on the patient and impairment to which they are applied.

Most studies report summary investigations of WM impairment, and mainly distinguish verbal versus visuo-spatial WM impairment. This results in the use of relatively broad treatment strategies and/or WM tasks and leads to suboptimal content validity. For example, Vallat et al. (2005, 2014) used treatment WM tasks that already included processing components (such as alphabetic word span tasks, visual N-back tasks) when treating the retention component. As a result, storage and processing WM tasks also improved during the retention treatment stages and did not improve much more when executive/attentional processing components were more specifically targeted at a later treatment stage. This also raises the more general question of whether it is possible to clearly demonstrate the specificity of WM training when targeting WM processes that have been suggested to be domain general, such as executive and control of attention aspects of WM (Cowan, 1999; Majerus et al., 2010, 2016). As already discussed earlier, it may be difficult to find WM tasks that do not require at least some minimal amount of attentional control and that would not benefit from treatment of this component.

A further difficulty for assessing the specificity of WM treatment is related to the number of domains/processes in which the patients show impairment. A number of the patients that have been reviewed here do not present severe language impairment anymore, or they present mild language impairment that is difficult to detect with standard language measures. For these patients, verbal WM impairment can be the only detectable deficit – this deficit will however lead to significant difficulties in other domains such as problem solving, mental calculation, learning of novel information, note taking or group conversations (Closset & Majerus, 2007; Majerus et al., 2005; Vallat et al., 2005). In this situation, it may be difficult to find suitable, laboratory measures for establishing the specificity of WM training. A way to circumvent these difficulties is to use control treatments instead of control tasks, as has been done by Berthier et al. (2014). This strategy is optimal for patients in which the WM deficit is the only remaining deficit as only the WM treatment should lead to improved WM performance. The use of control treatments is less optimal for more broadly impaired patients, as the control treatment, if consisting in conventional language therapy as in Berthier et al. will lead to generalized treatment effects, making it difficult to disentangle control and target treatment effects.

A further methodological aspect assessed in this review was the amount of pretreatment and posttreatment measures that were administered in the different studies. As shown in Table 1, virtually all studies respected the minimal requirement of at least 3/2 pretreatment/posttreatment measures (Beeson & Robey, 2006) for near transfer measures; in many cases, these minimal amounts were exceeded by a factor of two or higher. The numbers are however more problematic for far transfer and control measures for

which several studies only included one or two measures. In order to further demonstrate the efficacy of WM treatment in single case studies, it is important that future studies include a sufficient number of far transfer and control measures, and that they use measures with a sufficient sensitivity. Far transfer and control measures should be at least as sensitive as the target WM tasks.

A further methodological improvement for future studies could be the increasing use of experimenters/clinicians that are blind to the status of the tasks and treatment methods, in order to rule out the possibility that the effect sizes of the treatments are inflated by implicit attitudes of the experimenter/clinician (Salis et al., 2017). At a practical level, given the individualized treatment methods that are necessary for designing case study designs, it is impossible, and perhaps even not desirable, that the main experimenter/clinician supervising the study is completely naïve to the patient's nature of impairment and to the nature of the tasks/treatments being administered. However, parts of the treatment/assessments could be carried out by research assistants/clinicians that are blinded to at least some aspects of the study design. A few of the reviewed studies have adopted this 'hybrid' strategy of experimenter blinding (e.g., Salis et al., 2017). A further possibility is the fully computerized administration of assessment and treatment tasks (Salis et al., 2017). At the same time, in many cases, direct interactions between the experimenter/clinician and the patient are necessary for proper administration of assessment and treatment tasks, as well as for the interpretation of the results and subsequent adaptation of the treatment procedure. These interactions may be particularly important for WM treatment as WM tasks are often perceived as quite frustrating given the severe levels of WM impairment in many of the patients. The intervention of the experimenter/clinician remains therefore a critical factor for maintaining the patient's motivation and helping them to go through the WM treatment. It is furthermore important that WM treatment procedures are compared to control treatment procedures that are equally challenging at the motivational level for the patient (von Bastian & Oberauer, 2014).

A final issue that needs to be considered is whether the WM treatment approaches proposed so far are adapted to the nature of WM impairment that characterizes patients with aphasia. In the first part of this review paper, at least four possible types of WM impairment have been identified in patients with aphasia: short-term retention deficits resulting from excessive decay of phonological item representations, over-activation of representations leading to interference effects in WM tasks, serial order retention deficits, and impairment to executive/attentional components of WM. The treatment approaches reviewed here cover phonological and semantic decay-based impairment with both restoration and compensation (Closset & Majerus, 2007) intervention methods. The reviewed approaches also cover relatively well control of attention aspects of WM although they do not specify the type of attentional and/or executive aspects that is targeted by the treatment. For serial order retention deficits, no specific intervention method has been developed so far. Salis and colleagues (Salis, 2012; Salis et al., 2017) used a training method based on serial order recognition but the treatment approach did not explicitly target serial order WM; rather the recognition-based treatment strategy was used to circumvent the patients' linguistic output processing difficulties. Finally, there seem to be no intervention methods that would target WM impairment resulting from excessive interference at the level of semantic or phonological item representations. The development of treatment strategies for serial order retention deficits and control of semantic/phonological interference deficits is important given the growing number of patients that have been identified to show these types of WM impairment (Barde et al., 2010; Majerus et al., 2015; Ralph et al., 2017). There may also be other important distinctions in WM components such as the distinction between recall and recognition tasks and the use of verbal rehearsal strategies. Gvion and Friedmann (2012) observed that patients with conduction aphasia can have impairment in both recall and recognition WM tasks, or impairment restricted to recall tasks indicating that verbal rehearsal may still be functional in this latter type of patients (see also N. Martin, 2012; Salis et al., 2017). Taking into consideration this type of distinction is a fundamental aspect for studies that would target verbal rehearsal as a treatment strategy.

More generally, a theoretically-oriented approach of WM treatment calls for a theoretically-oriented approach of WM assessment, and this means that we also need to increase the theoretical and psychometric quality of existing WM assessment instruments. Currently, different authors have developed different verbal WM assessment instruments, inspired by different theoretical models (e.g., Harris et al., 2014; Kalinyak-Fliszar et al., 2011; Majerus et al., 2015; Vallat et al., 2014). A comprehensive WM test battery that would target the different WM components discussed here in a process-specific and integrated manner is still lacking. Although there have been some standardization efforts for a subset of the tasks and instruments presented here (e.g., Kalinyak-Fliszar et al., 2011; Majerus et al., 2015), standardized, psychometrically valid and theoretically-informed WM test batteries, particularly for use in severely impaired patients such as aphasic patients, remain scarce (Murray, Salis, Martin, & Dralle, 2016). By explicitly characterizing the storage versus processing requirements as well as the attentional and executive processes involved in all types of WM tasks, a theoretically-informed WM battery could also be a step toward the resolution of the debate around the distinction of 'working memory' versus 'short-term memory' tasks. The development of theoretically-informed and integrated WM batteries would benefit both the assessment and treatment of WM deficits in aphasia.

To conclude, despite the relatively limited amount of published studies, single-case WM treatment studies appear to show satisfactory levels of efficacy but there remain some uncertainties about their specificity and content validity. In order to improve the specificity and content validity of WM treatment, future studies should conduct theoretically motivated investigations of WM impairment in order to implement more targeted and theoretically informed WM intervention methods. Also, in order to allow for a more robust assessment of the effect sizes of WM treatment in single case studies, larger sets of pretreatment/posttreatment measures should be used, and this particularly for far transfer and controls tasks. Partial experimenter blinding to the status of tasks/treatments should also be considered for ensuring reliable and unbiased administration of WM treatment. Finally, publication of unsuccessful single case WM treatment studies should be encouraged in order to allow for an assessment of the true effect size of WM treatment in aphasia, to help establish benchmarks against which individual effect sizes can be compared, and to isolate the conditions under which WM treatment is effective or not.

Acknowledgments

This work was supported by the government of the French-speaking community of Belgium (grants 12/17-01-REST and FSR-S-SH-17/04).

References

- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105(1), 158–173.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 47–90). San Diego, CA: Academic Press.
- Barde, L. H. F., Schwartz, M. F., Chrysikou, E. G., & Thompson-Schill, S. L. (2010). Reduced short-term memory span in aphasia and susceptibility to interference: Contribution of material-specific maintenance deficits. *Neuropsychologia*, 48, 909–920.
- Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*, 133, 83–100.
- von Bastian, C., & Oberauer, K. (2014). Effects and mechanisms of working memory training: A review. *Psychological Research*, 78, 803–820.
- Beeson, P. M., & Egnor, H. (2006). Combining treatment for written and spoken naming. *Journal of the International Neuropsychological Society*, 12, 816–827.
- Beeson, P. M., Magloire, J., & Robey, R. R. (2005). Letter-by-letter reading: Natural recovery and response to treatment. *Behavioural Neurology*, 16, 191–202.
- Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychological Review*, 16, 161–169.
- Berthier, M. L., Davila, G., Green-Heredia, C., Moreno Torres, I., Juarez y Ruiz de Mier, R., De-Torres, I., et al. (2014). Massed sentence repetition training can augment and speed up recovery of speech production deficits in patients with chronic conduction aphasia receiving donepezil treatment. *Aphasiology*, 28, 188–218.
- Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, 107, 127–181.
- Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, 106, 551–581.
- Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language*, 55, 627–652.
- Burgio, F., & Basso, A. (1997). Memory and aphasia. *Neuropsychologia*, 35, 759–766.
- Caramazza, A., Basili, A. G., Koller, J. J., & Berndt, R. S. (1981). An investigation of repetition and language processing in a case of conduction aphasia. *Brain and Language*, 14, 235–271.
- Closset, A., & Majerus, S. (2007). Rééducation de la mémoire phonologique à court terme: Application chez une enfant de 10 ans. In G. Aubin, F. Coyette, P. Pradat-Diehl, & C. Vallat-Azouvi (Eds.), *Neuropsychologie de la mémoire de travail* (pp. 323–349). Marseille: Solal.
- Coelho, C. (2005). Direct attention training as a treatment for reading impairment in mild aphasia. *Aphasiology*, 19, 275–283.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.
- Collette, F., Hogge, M., Salmon, E., & Van der Linden, M. (2006). Exploration of the neural substrates of executive functioning by functional neuroimaging. *Neuroscience*, 139, 209–221.
- Constantinidis, C., & Klingberg, T. (2016). The neuroscience of working memory capacity and training. *Nature Reviews Neuroscience*, 17, 438–449.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3, 201–215.
- Cowan, N. (1995). *Attention and memory: An integrated framework*. New York: Oxford University Press.
- Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake, & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge: Cambridge University Press.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–185.
- Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, 24, 1158–1170.
- Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjattullina, A., et al. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, 51, 42–100.
- Cowan, N., & Morey, C. C. (2007). How can dual-task working memory retention limits be investigated? *Psychological Science*, 18, 686–688.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 145–199). San Diego, USA: Elsevier Academic Press.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, 128, 309–331.
- Foster, J. L., Harrison, T. L., Hicks, K. L., Draheim, C., Redick, T. S., & Engle, R. W. (2017). Do the effects of working memory training depend on baseline ability level? *Journal of Experimental Psychology: Learning, Memory and Cognition*. <http://dx.doi.org/10.1037/xlm0000426>.
- Francis, D., Clark, N., & Humphreys, G. (2003). The treatment of an auditory working memory deficit and the implications for sentence comprehension abilities in mild "receptive" aphasia. *Aphasiology*, 17, 723–750.
- Gold, B. T., Balota, D. A., Kirchoff, B. A., & Buckner, R. L. (2005). Common and dissociable activation patterns associated with controlled semantic and phonological processing: Evidence from fMRI adaptation. *Cerebral Cortex*, 15, 1438–1450.
- Gvion, A., & Friedmann, N. (2012). Phonological short-term memory in conduction aphasia. *Aphasiology*, 26, 579–614.
- Hamilton, A., & Martin, R. C. (2007). Proactive interference in a semantic short-term memory deficit: Role of semantic and phonological relatedness. *Cortex*, 43, 112–123.
- Harris, L., Olson, A., & Humphreys, G. (2014). The link between STM and sentence comprehension: A neuropsychological rehabilitation study. *Neuropsychological Rehabilitation*, 24, 678–720.
- Hartley, T., Hurlstone, M. J., & Hitch, G. J. (2016). Effects of rhythm on memory for spoken sequences: A model and tests of its stimulus-driven mechanism. *Cognitive Psychology*, 87, 135–178.
- Henson, R. N. A. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology*, 36, 73–137.
- Hoffman, P., Jefferies, E., Ehsan, S., Hopper, S., & Lambon Ralph, M. A. (2009). Selective short-term memory deficits arise from impaired domain-general semantic control mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 137–156.
- Hoffman, P., Jones, R. W., & Ralph, M. A. (2012). The degraded concept representation system in semantic dementia: Damage to pan-modal hub, then visual spoke. *Brain*, 135, 3770–3780.
- Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin*, 140, 339–373.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *PNAS*, 108, 10081–10086.
- Jefferies, E., Baker, S. S., Doran, M., & Lambon Ralph, M. A. (2007). Refractory effects in stroke aphasia: A consequence of poor semantic control. *Neuropsychologia*, 45, 1065–1079.
- Kalinyak-Fliszar, M., Kohen, F., & Martin, N. (2011). Remediation of language processing in aphasia: Improving activation and maintenance of linguistic representations in (verbal) short-term memory. *Aphasiology*, 25, 1095–1131.
- Kalm, K., & Norris, D. (2014). The representation of order information in auditory-verbal short-term memory. *The Journal of Neuroscience*, 34, 6879–6886.
- Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, 14, 317–324.
- Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical Experimental Neuropsychology*, 24, 781–791.
- Koenig-Bruhin, M., & Studer-Eichenberger, F. (2007). Therapy of short-term memory disorders in fluent aphasia: A single case study. *Aphasiology*, 21, 448–458.

- Kundu, B., Sutterer, D. W., Emrich, S. M., & Postle, B. R. (2013). Strengthened effective connectivity underlies transfer of working memory training to tests of short-term memory and attention. *Journal of Neuroscience*, *33*, 8705–8715.
- Kuzmina, K., & Weekes, B. (2017). Role of cognitive control in language deficits in different types of aphasia. *Aphasiology*, *31*, 765–792.
- Lane, J. D., & Gast, D. L. (2014). Visual analysis in single case experimental design studies: Brief review and guidelines. *Neuropsychological Rehabilitation*, *24*, 445–463.
- Lee, M. D., & Wagenmakers, E. J. (2013). *Bayesian modeling for cognitive science: A practical course*. Cambridge, UK: Cambridge University Press.
- Lewandowsky, S., & Oberauer, K. (2009). No evidence for temporal decay in working memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *35*, 1545–1551.
- Majerus, S. (2009). Verbal short-term memory and temporary activation of language representations: The importance of distinguishing item and order information. In A. S. Thorn, & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 244–276). Hove, UK: Psychology Press.
- Majerus, S. (2013). Language repetition and short-term memory: An integrative framework. *Frontiers in Human Neuroscience*, *7*, 357 doi: 310.3389/fnhum.2013.00357.
- Majerus, S., Attout, L., Artielle, M. A., & Van der Kaa, M. A. (2015). The heterogeneity of verbal short-term memory impairment in aphasia. *Neuropsychologia*, *77*, 165–176.
- Majerus, S., & Boukebzca, C. (2013). Short-term memory for serial order supports vocabulary development: New evidence from a novel word learning paradigm. *Journal of Experimental Child Psychology*, *116*, 811–828.
- Majerus, S., Cowan, N., Peters, F., Van Calster, L., Phillips, C., & Schrouff, J. (2016). Cross-Modal decoding of neural patterns associated with working Memory: Evidence for attention-based accounts of working memory. *Cerebral Cortex*, *26*, 166–179.
- Majerus, S., D'Argembeau, A., Martinez, T., Belayachi, S., Van der Linden, M., Collette, F., ... Maquet, P. (2010). The commonality of neural networks for verbal and visual short-term memory. *Journal of Cognitive Neuroscience*, *22*, 2570–2593.
- Majerus, S., Lekeu, F., Van der Linden, M., & Salmon, E. (2001). Deep dysphasia: Further evidence on the relationship between phonological short-term memory and language processing impairments. *Cognitive Neuropsychology*, *18*, 385–410.
- Majerus, S., Norris, D., & Patterson, K. (2007). What do patients with semantic dementia remember in verbal short-term memory? Sounds and order but not words. *Cognitive Neuropsychology*, *24*, 131–151.
- Majerus, S., Péters, F., Bouffier, M., Cowan, N., & Phillips, C. (2017). The dorsal attention network reflects both encoding load and top-down control during working memory. *Journal of Cognitive Neuroscience*. http://dx.doi.org/10.1162/jocn_a.01195 (in press).
- Majerus, S., Poncet, M., Van der Linden, M., Albouy, G., Salmon, E., Sterpenich, V., ... Maquet, P. (2006). The left intraparietal sulcus and verbal short-term memory: Focus of attention or serial order? *NeuroImage*, *32*, 880–891.
- Majerus, S., Van der Kaa, M. A., Renard, C., Van der Linden, M., & Poncet, M. (2005). Treating verbal short-term memory deficits by increasing the duration of temporary phonological representations: A case study. *Brain and Language*, *95*, 174–175.
- Marshuetz, C., Smith, E. E., Jonides, J., DeGutis, J., & Chenevert, T. L. (2000). Order information in working memory: fMRI evidence for parietal and prefrontal mechanisms. *Journal of Cognitive Neuroscience*, *12*, 130–144.
- Martin, N. (2012). Managing communication deficits associated with memory disorders. In R. Peach, & L. Shapiro (Eds.), *Cognition and acquired language disorders: A process-oriented approach*. Mosby.
- Martin, R. C., & Lesch, M. F. (1996). Associations and dissociations between language impairment and list recall: Implications for models of STM. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 149–178). East Sussex, UK: Hove.
- Martin, N., & Saffran, E. M. (1992). A computational account of deep dysphasia: Evidence from a single case study. *Brain and Language*, *43*, 240–274.
- Martin, N., Saffran, E. M., & Dell, G. S. (1996). Recovery in deep dysphasia: Evidence for a relation between auditory-verbal STM capacity and lexical errors in repetition. *Brain and Language*, *52*, 83–113.
- Martin, R. C., Shelton, J. R., & Yaffee, L. S. (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language*, *33*, 83–111.
- Mayer, J. F., & Murray, L. J. (2002). Approaches to the treatment of alexia un chronic aphasia. *Aphasiology*, *16*, 727–743.
- Mayer, J. F., & Murray, L. L. (2012). Measuring working memory deficits in aphasia. *Journal of Communication Disorders*, *45*, 325–339.
- McCarthy, L. M., Kalinyak-Fliszar, M., Kohen, F., & Martin, N. (2017). Effects of semantic context on access to words of low imageability in deep-phonological dysphasia: A treatment case study. *Aphasiology*, *31*(5), 542–562.
- Melby-Lervag, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far Transfer": Evidence from a meta-analytic review. *Perspectives in Psychological Sciences*, *11*, 512–534.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.
- Moore, T. M., Reise, S. P., Depaoli, S., & Haviland, M. G. (2015). Iteration of partially specified target Matrices: Applications in exploratory and bayesian confirmatory factor analysis. *Multivariate Behavioural Research*, *50*, 149–161.
- Morey, C. C., & Miron, M. D. (2016). Spatial sequences, but not verbal sequences, are vulnerable to general interference during retention in working memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *42*, 1907–1918.
- Murray, L. L. (2012). Direct and indirect treatment approaches for addressing short-term or working memory deficits in aphasia. *Aphasiology*, *26*, 3–4.
- Murray, L. L., Holland, A. L., & Beeson, P. M. (1998). Spoken language of individuals with mild fluent aphasia under focused and divided attention conditions. *Journal of Speech, Language and Hearing Research*, *41*, 213–227.
- Murray, L. L., Keeton, R. J., & Karcher, L. (2006). Treating attention in mild aphasia: Evaluation of attention process training-II. *Journal of Communication Disorders*, *39*, 37–61.
- Murray, L., Salis, C., Martin, N., & Dralle, J. (2016). The use of standardised short-term and working memory tests in aphasia research: A systematic review. *Neuropsychological Rehabilitation*, *1–43*. <http://dx.doi.org/10.1080/09602011.2016.1174718>.
- Nairne, J. S., & Kelley, M. R. (2004). Separating item and order information through process dissociation. *Journal of Memory and Language*, *50*, 113–133.
- Oberauer, K., & Lewandowsky, S. (2013). Evidence against decay in verbal working memory. *Journal of Experimental Psychology: General*, *142*, 380–411.
- Papagno, C., Comi, A., Riva, M., Bizzi, A., Vernice, M., Casarotti, A., ... Bello, L. (2017). Mapping the brain network of the phonological loop. *Human Brain Mapping*, *38*, 3011–3024.
- Peach, R. K., Nathan, M. R., & Beck, K. M. (2017). Language-specific attention treatment for Aphasia: Description and preliminary findings. *Seminars in Speech and Language*, *38*, 5–16.
- Poirier, M., Saint-Aubin, J., Mair, A., Tehan, G., & Tolan, A. (2015). Order recall in verbal short-term memory: The role of semantic networks. *Memory and Cognition*, *43*, 489–499.
- Potter, M. C., & Lombardi, L. (1998). Syntactic priming in immediate recall of sentences. *Journal of Memory and Language*, *38*, 265–282.
- Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*, 42–55.
- Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., ... Engle, R. W. (2013). No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *Journal of Experimental Psychology: General*, *142*, 359–379.
- Ricker, T. J., Spiegel, L. R., & Cowan, N. (2014). Time-based loss in visual short-term memory is from trace decay, not temporal distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1510–1523.
- Ricker, T. J., Vergauwe, E., & Cowan, N. (2016). Decay theory of immediate memory: From Brown (1958) to today (2014). *Quarterly Journal of Experimental Psychology*, *69*, 1969–1995.
- Salis, C. (2012). Short-term memory treatment: Patterns of learning and generalisation to sentence comprehension in a person with aphasia. *Neuropsychological Rehabilitation*, *22*, 428–448.
- Salis, C., Hwang, F., Howard, D., & Lallini, N. (2017). Short-term and working memory treatments for improving sentence comprehension in aphasia: A review and a replication study. *Seminars in Speech and Language*, *38*, 29–39.
- Salis, C., Kelly, H., & Code, C. (2015). Assessment and treatment of short-term and working memory impairments in stroke aphasia: A practical tutorial. *International*

- Journal of Language and Communication Disorders*, 50, 721–736.
- Sawilowsky, S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, 8, 467–474.
- Schwepe, J., Rummer, R., Bormann, T., & Martin, R. C. (2011). Semantic and phonological information in sentence recall: Converging psycholinguistic and neuropsychological evidence. *Cognitive Neuropsychology*, 28, 521–545.
- Seniow, J., Litwin, M., & Lesniak, M. (2009). The relationship between non-linguistic cognitive deficits and language recovery in patients with aphasia. *Journal of Neurological Sciences*, 283, 91–94.
- Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, 138, 628–654.
- Shrubsole, K., Worrall, L., Power, E., & O'Connor, D. A. (2017). Recommendations for post-stroke aphasia rehabilitation: An updated systematic review and evaluation of clinical practice guidelines. *Aphasiology*, 31, 1–24.
- Sinotte, M. P., & Coelho, C. A. (2007). Attention training for reading impairment in mild aphasia: A follow-up study. *NeuroRehabilitation*, 22, 303–310.
- Sohlberg, M. M., Johnson, L., Paule, L., Raskin, S. A., & Mateer, C. A. (2001). *Attention process training-II: A program to address attentional deficits for persons with mild cognitive dysfunction* (2nd ed.). Wake Forest, NC: Lash & Associates.
- Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, 428, 751–754.
- Unsworth, N., & Engle, R. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, 133, 1038–1066.
- Vallar, G., & Baddeley, A. D. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behavior*, 23, 151–161.
- Vallat-Azouvi, C., Pradat-Diehl, P., & Azouvi, P. (2014). Modularity in rehabilitation of working memory: A single-case study. *Neuropsychological Rehabilitation*, 24, 220–237.
- Vallat, C., Azouvi, P., Hardisson, H., Meffert, R., Tessier, C., & Pradat-Diehl, P. (2005). Rehabilitation of verbal working memory after left hemisphere stroke. *Brain Injury*, 19, 1157–1164.
- Van der Kaa, M. A., & Majerus, S. (2007). Evaluation et rééducation du versant « phonologique » de la mémoire à court terme. In G. Aubin, F. Coyette, P. Pradat-Diehl, & C. Vallat-Azouvi (Eds.), *Neuropsychologie de la mémoire de travail* (pp. 87–111). Marseille: Solal.
- Zakarias, L., Keresztes, A., Marton, K., & Wartenburger, I. (2016). Positive effects of a computerised working memory and executive function training on sentence comprehension in aphasia. *Neuropsychological Rehabilitation*, 1–18.