

# UC Berkeley

## UC Berkeley Previously Published Works

### Title

Neural mechanisms of two different verbal working memory tasks: A VLSM study.

### Permalink

<https://escholarship.org/uc/item/7f76c3gm>

### Authors

Ivanova, MV  
Dragoy, O  
Kuptsova, SV  
[et al.](#)

### Publication Date

2018-07-01

### DOI

10.1016/j.neuropsychologia.2018.03.003

Peer reviewed



Published in final edited form as:

*Neuropsychologia*. 2018 July 01; 115: 25–41. doi:10.1016/j.neuropsychologia.2018.03.003.

## Neural Mechanisms of Two Different Verbal Working Memory Tasks: a VLSM Study

M.V. Ivanova<sup>a,b,\*</sup>, O.V. Dragoy<sup>a,c</sup>, S.V. Kuptsova<sup>a,d</sup>, Yu.S. Akinina<sup>a,e</sup>, A.G. Petrushevskii<sup>d</sup>, O.N. Fedina<sup>d</sup>, A. Turken<sup>b</sup>, V.M. Shklovsky<sup>c,d</sup>, N.F. Dronkers<sup>a,b,f</sup>

<sup>a</sup>National Research University Higher School of Economics, Neurolinguistics Laboratory, 21/4 Staraya Basmannaya street, office 510, 105066, Moscow, Russia <sup>b</sup>Center for Aphasia and Related Disorders, VA Northern California Health Care System, 150 Muir Road 126R, 94553, Martinez, California, USA <sup>c</sup>Moscow Research Institute of Psychiatry, Department of Speech Pathology and Neurorehabilitation, 3 Potesnaya street 3, 107076, Moscow, Russia <sup>d</sup>Center for Speech Pathology and Neurorehabilitation, 20 Nikoloyamskaya street, 109240, Moscow, Russia <sup>e</sup>University of Groningen, Graduate School for the Humanities, P.O. Box 716 NL-9700 AS Groningen, Groningen, The Netherlands <sup>f</sup>University of California, Davis, 1 Shields Ave, 95616, Davis, California, USA

### Abstract

Currently, a distributed bilateral network of frontal-parietal areas is regarded as the neural substrate of working memory (WM), with the verbal WM network being more left-lateralized. This conclusion is based primarily on functional magnetic resonance imaging (fMRI) data that provides correlational evidence for brain regions involved in a task. However, fMRI cannot differentiate the areas that are fundamentally required for performing a task. These data can only come from brain-injured individuals who fail the task after the loss of specific brain areas. In addition to the lack of complimentary data, is the issue of the variety in the WM tasks used to assess verbal WM. When different tasks are assumed to measure the same behavior, this may mask the contributions of different brain regions. Here, we investigated the neural substrate of WM by using voxel-based lesion symptom mapping (VLSM) in 49 individuals with stroke-induced left hemisphere brain injuries. These participants completed two different verbal WM tasks: complex listening span and a word 2-back task. Behavioral results indicated that the two tasks were only slightly related, while the VLSM analysis revealed different critical regions associated with each task. Specifically, significant detriments in performance on the complex span task were found with lesions in the inferior frontal gyrus, while for the 2-back task, significant deficits were seen after injury to the superior and middle temporal gyri. Thus, the two tasks depend on the structural integrity of different, non-overlapping frontal and temporal brain regions, suggesting distinct

\*Corresponding author: Mailing address: Maria V. Ivanova, Center for Aphasia and Related Disorders, VA Northern California Health Care System, 150 Muir Road 126R, 94553, Martinez, California, USA, mvimaria@gmail.com.

**Conflicts of interest:** none.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

neural and cognitive mechanisms triggered by the two tasks: rehearsal and cue-dependent selection in the complex span task, versus updating/auditory recognition in the 2-back task. These findings call into question the common practice of using these two tasks interchangeably in verbal WM research and undermine the legitimacy of aggregating data from studies with different WM tasks. Thus, the present study points out the importance of lesion studies in complementing functional neuroimaging findings and highlights the need to consider task demands in neuroimaging and neuropsychological investigations of WM.

### Keywords

working memory; complex span; N-back; VLSM; prefrontal cortex; temporal gyri

---

## INTRODUCTION

### 1.1 Neural foundations of verbal WM

Working memory (WM) is broadly defined as the ability and/or resources to concurrently process and store task relevant information or coordinate processing of multiple competing inputs. Within the last 50 years it has become the cornerstone of cognition, demonstrating a key role in many higher-level cognitive processes (Baddeley, 2003; Caplan & Waters, 1999; Daneman & Merikle, 1996; Gathercole & Baddeley, 1993; Just & Carpenter, 1992; Kane et al., 2004). With the advancement of neuroimaging techniques, it has become increasingly important to understand the neural correlates of WM to further expand and validate existing theories of cognition (D'Esposito & Postle, 2015). The literature concerning the neural substrate of verbal WM reports similar functional magnetic resonance imaging (fMRI) activations as with general WM, namely lateral premotor cortex, dorsal cingulate, medial premotor cortex, dorsolateral and ventrolateral prefrontal cortex (PFC), frontal pole, and bilateral and medial posterior parietal cortex (D'Esposito & Postle, 2015; Ma, Husain, & Bays, 2014; Owen, McMillan, Laird, & Bullmore, 2005; Rottschy et al., 2012; Wager & Smith, 2003). Some meta-analyses report more extensive recruitment of the left ventrolateral PFC, particularly the inferior frontal gyrus (Rottschy et al., 2012), while others do not (Owen et al., 2005; Wager & Smith, 2003). Unfortunately, when specific studies are examined, there is a discrepancy in the location of activations within and across hemispheres. There are two factors that likely affect the existing evidence on the neural basis of verbal WM: certain limitations of the functional neuroimaging methodology, and task variability.

### 1.2 Methodological limitations of existing neuroimaging studies

Most of what we know today about the neural substrate of verbal WM is based on functional neuroimaging studies. Functional neuroimaging provides correlational evidence of the role of brain regions in the studied processes. Consequently, fMRI detects additional, associated activation, thus showing recruitment of areas that might not be central to the psychological function under investigation. Not surprisingly, studies comparing mapping of cognitive functions using fMRI and intraoperative electrical stimulation have repeatedly shown that fMRI is not able to detect critical functional areas reliably, with excision of tissue activated

in fMRI paradigms often not resulting in post-surgical deficits (Babajani-Feremi et al., 2016; Ottenhausen, Krieg, Meyer, & Ringel, 2015). Thus, functional neuroimaging data need to be further complemented and validated by lesion studies, that detect critical neural areas supporting the cognitive function under examination and demonstrate a causal role for a specific region (Müller & Knight, 2006; Rorden & Karnath, 2004). In other words, exploration of lesion data, especially employing modern methods of analyses, such as voxel-based lesion-symptom mapping (VLSM), establishes the necessity of certain brain areas for cognitive functions. Thus, human lesion and functional imaging studies can provide complementary information on brain-behavior relationships (D'Esposito, Cooney, Gazzaley, Gibbs, & Postle, 2006; Karnath, Sperber, & Rorden, 2018).

Despite the need, there are few investigations of verbal WM using lesion methods (Müller & Knight, 2006) and most of them are based on relatively small sample size of patients with varied etiologies. Those that exist have identified a much more circumscribed network of areas in the left hemisphere compared to fMRI studies, encompassing posterior temporal, inferior parietal and lateral prefrontal cortex with the posterior regions being sufficient for supporting performance on undemanding recall tasks (Baldo & Dronkers, 2006; D'Esposito & Postle, 1999; for review see also Müller & Knight, 2006), suggesting that they may serve the role of the phonological buffer (Baldo, Katseff, & Dronkers, 2012). Another insight from lesion studies is that the executive component of WM is not localized to a specific circumscribed area of the PFC, but rather is distributed along the ventro- and dorsolateral PFC (Müller, Machado, & Knight, 2002; Szczepanski & Knight, 2014) with greater amount of damage to the PFC leading to more impairment on WM tasks.

The dissociation between functional neuroimaging and lesion methods is particularly evident when one compares the two techniques within the same investigation. This was done in one study, where involvement of brain regions in performance on immediate and delayed span tasks was compared using both fMRI in non-brain damaged controls and patients with unilateral frontal lesions (D'Esposito et al., 2006). It was clearly demonstrated that, the ability to temporarily store and rehearse information in delayed response tasks (even with a distraction) was not significantly affected by unilateral lesions of the lateral PFC, although activations in the frontal cortex for these tasks were observed in healthy individuals. Similar discrepancies between the two methods have been observed in other studies (e.g., Volle et al., 2008).

### 1.3 Task limitations: task selection & variability

Another limitation and the source of variability in observed activation patterns may be due to the employment of various WM tasks, such as N-back (Cohen et al., 1997), variants of Sternberg's item recognition (Awh et al., 1996), continuous performance (Carter et al., 1998), and self-ordered selection tasks (Curtis, Zald, & Pardo, 2000), most of which involve multiple executive processes. Wager and Smith (2003), in one of the first reviews of neuroimaging studies of WM, provided evidence that various cognitive mechanisms, such as manipulation, order storage, updating and selective attention, implicated differentially in these different WM tasks recruit distinctive brain areas. However, many following meta-analyses either tended to generalize from a single type of WM task (Owen et al., 2005) or

grouped all data coming from characteristically different WM tasks (Rottschy et al., 2012), thus possibly confounding the neural underpinnings of the various executive processes involved.

Moreover, there is a striking discrepancy between WM tasks used in functional neuroimaging investigations, lesion studies and behavioral research. The most common verbal WM tasks used in functional neuroimaging research are item recognition tasks, particularly letter N-back tasks, due to their easy adaptability to the scanner testing environment (Chein, Moore, & Conway, 2011; Owen, McMillan, Laird, & Bullmore, 2005). In N-back tasks, participants are required to recognize identical items in a continuous stream of items by indicating whether the current item matches the stimulus presented N items back. However, most of what we know today about the psychological structure of WM is based on studies using complex span tasks, the gold standard for assessing WM capacity in cognitive psychology, where participants are required to perform a processing task and simultaneously remember a set of items for later recall (Chein et al., 2011; Conway et al., 2005). From a conceptual standpoint, these two tasks implicate different cognitive mechanisms: N-back tasks require updating and recognition of items in the focus of attention, while successful performance of the complex span task depends on efficient shifting between competing representations and recall (Chein et al., 2011; Akira Miyake & Friedman, 2012). Not surprisingly, behavioral studies of WM provide convincing evidence that the N-back task and the classical complex span task do not index the same underlying processes and performance on these tasks is at best weakly correlated (Ivanova, Kuptsova, & Dronkers, 2017; Jaeggi, Buschkuhl, Perrig, & Meier, 2010; Kane, Conway, Miura, & Colflesh, 2007; Redick & Lindsey, 2013). The tasks used in the lesion literature remain even more discrepant, with most studies employing classical forward and backward word/digit span. The traditional complex listening/reading span task (Conway et al., 2005), surprisingly, has never been used in lesion studies, possibly because of its difficulty for the patient population. Consequently, this leads to a dramatic gap between what we know about the psychological structure and functioning of WM and about its neural bases (Chein et al., 2011; Smith et al., 2001).

A further concern is the overuse of the letter N-back task to investigate verbal WM in fMRI research. The existing evidence indicates that letter identification is supported by a specific perceptual processing system distinct from the language system and that the right hemisphere can at least partially support this letter identification process (up to access of the symbolic information) in healthy controls (Nestor, Behrmann, & Plaut, 2013; Simos et al., 2002) and in individuals with neurogenic language disorders (Cohen et al., 2003; McCloskey & Schubert, 2014). Consequently, most neuroimaging studies of letter N-back report extensive right hemisphere activations and recruitment of visual areas (e.g., Awh et al., 1996; Cohen et al., 1997; Volle et al., 2008; for reviews see Owen et al., 2005, Rottschy et al., 2012; Wager & Smith, 2003). Therefore, it is somewhat inaccurate that letter N-back tasks have been classified as “verbal”.

Out of a total of 213 studies using various types of N-back tasks included in recent meta-analyses of WM (Owen et al., 2005; Rottschy et al., 2012), only two studies actually employed word N-back tasks. Reynolds and colleagues (2009) used a word N-back task and

demonstrated that with increased maintenance load (3-back compared to 1-back), activation in the right dorsolateral PFC, bilateral parietal cortex, and the cerebellum was detected. Only in one study did word N-back elicit left-lateralized activation of the ventrolateral PFC, along with activation in more posterior regions (Braver et al., 2001). In stark contrast to letter and word N-back tasks, those fMRI studies that used adaptations of verbal complex span tasks showed much more extensive activations in the left dominant hemisphere, including superior and middle temporal gyri, anterior temporal lobe, left middle and inferior PFC, and anterior cingulate cortex (Bunge, Klingberg, Jacobsen, & Gabrieli, 2000; Chein et al., 2011; Just, Carpenter, & Keller, 1996; Kondo et al., 2004; M. Osaka et al., 2003; N. Osaka et al., 2004). Thus, despite the abundance of functional neuroimaging literature on verbal WM, due to the variation in task demands these data yield very little regarding the neural correlates of typical complex span tasks and word N-back tasks used in cognitive psychology or neurolinguistics research.

Lesion studies of the neural basis of N-back tasks also provide inconsistent findings. Tsuchida and Fellows (2009) found the dorsolateral anterior cingulate cortex and adjacent dorsal fronto-medial cortex to be critical for letter 2-back performance, but not dorsolateral and ventrolateral PFC areas reported in fMRI studies. Volle and colleagues (2008) demonstrated that the posterior portion of the inferior frontal gyrus contributed to the deficits in performance on the letter 3-back task, but not on 1-back and 2-back tasks. Thus, the left PFC is recruited only when task demands increase and the specific role of the prefrontal lateral cortex in N-back tasks remains to be ascertained. The largest published lesion study of WM to date (n=158) employing a letter N-back task revealed an almost exclusively right hemisphere network of cortical regions, including right ventrolateral PFC, right inferior parietal cortex, and right middle temporal gyrus (Barbey, Colom, Paul, & Grafman, 2014). These findings again demonstrate that the letter N-back task can be successfully completed based on visual perceptual processes: matching of visual stimulus to previously presented one. Consequently, there is strong evidence coming from both activation and lesion studies that the letter N-back task has been previously erroneously classified as being a “verbal” WM task in neuroimaging research.

To conclude, despite active involvement in fMRI tasks, PFC does not seem to be required for WM tasks that involve simple and even delayed span memory. Instead, span seems to be more dependent on posterior cortical networks, particularly inferior parietal and superior-middle temporal areas for verbal stimuli. The PFC plays a more crucial role when the information being stored requires further manipulation to become useful for goal-directed behavior or when items have to be maintained in active memory in the face on-going interference. Thus, the question – which brain areas are necessary for accurate performance on particular verbal WM tasks or for proper functioning of verbal WM – remains unanswered. Exploration of the neural substrate of verbal WM remains limited due to the type of method and tasks used to investigate them. The lesion literature on verbal WM remains sparse. As we can see from the reviewed studies, the classical WM paradigm from cognitive psychology, the complex listening/reading span task, surprisingly has never been used in lesion studies. Additionally, small sample sizes are a perpetual challenge for lesion studies. Also, different etiologies, preselected lesion sites and different criteria used to define subgroups of patients make it difficult to synthesize findings across studies, with very few

studies attempting to use VLSM to define neural correlates of WM. In sum, much remains to be ascertained about the brain substrate of verbal WM. Lesion evidence is necessary to complement, validate and refine findings from activation studies and there is a great need for a larger-scale lesion study of WM using contemporary methods of analysis, such as VLSM, and employing various WM tasks.

#### 1.4 Aims of the current study

In the present study, we argue that the neural substrate of verbal WM might have been misrepresented due to over reliance on functional neuroimaging data and lack of consideration for the cognitive processes underlying different WM tasks. Given the existing literature and the lack of consensus on cortical areas essential for the verbal WM network, the aim of the current study was two-fold. First, we wanted to demonstrate the critical neural substrate of WM using a contemporary lesion method for establishing brain-behaviors relations: VLSM. Second, we sought to explore the impact of the type of task on critical areas and thus compare performance on two tasks commonly used to investigate WM in behavioral studies in psychology – complex span task, and in functional neuroimaging investigations of WM – N-back task. We anticipated, given the cognitive differences between these two types of tasks and previously reported discrepancies in performance on the two tasks, that they would rely on distinct neural circuitries.

## 1. METHODS AND MATERIALS

### 2.1 Participants

Persons with left hemisphere stroke at least 3-months post onset were recruited from the Center for Speech Pathology and Neurorehabilitation in Moscow, Russia. Presence of stroke was verified based on medical history and radiological reports. If radiological report or inspection of scans by a certified radiologist (the fifth and sixth authors of the study) indicated significant cortical atrophy, lacunes in the right hemisphere or diffuse white matter damage, then those individuals were excluded from further participation. Additional exclusion criteria were: diagnose of neurodegenerative disorders, epilepsy, or history of alcohol or substance abuse. To be included, participants had to be right-handed, native speakers of Russian and pass vision and hearing screenings. All participants gave informed consent prior to testing. The study was approved by the Ethics Committee of the National Research University HSE, Moscow, Russia.

49 individuals with chronic stroke participated in the study (23 male, 26 female;  $M_{age} = 51.98$  years,  $SD = 9.85$ , age range: 33 – 74 years). Participants had education levels ranging from completing secondary school (10 years) to a university degree (15 years) ( $M_{years\ of\ education} = 12.98$  years,  $SD = 1.91$ ). Thirty-three individuals had a clinical diagnosis of a single ischemic stroke in the middle cerebral artery distribution, one person (participant # 21) had a diagnosis of recurrent ischemic strokes in the left hemisphere occurring within 2 months, and 15 had a clinical diagnosis of a hemorrhagic stroke. Testing and MRI scanning took place at least 3-months post stroke ( $M_{post.onset} = 25.51$  months,  $SD = 30.03$ , post-onset range: 3 – 146 months).

Each participant was examined by a speech-language pathologist and a neuropsychologist of the Center. Based on language assessment and neuropsychological examination, forty-five individuals were diagnosed with aphasia following stroke and four individuals with dysarthria. Aphasia language deficits were classified according to Luria's system (Akhutina, 2015; Luria, 1980). Consequently, 21 persons with aphasia were classified as having non-fluent aphasia (either motor aphasia or dynamic aphasia), and 17 as having fluent aphasia (either sensory or acoustic-mnemonic aphasia), and seven individuals as having mixed types of aphasia. Also, all individuals with aphasia were assessed with the Assessment of Speech in Aphasia test (ASA; Tsvetkova, Akhutina, & Pylaeva, 1981), a comprehensive quantitative standard aphasia battery in Russian that includes both comprehension and production subtests along with rating of conversational speech (for a more detailed description of this test see Ivanova et al., 2016). Based on the overall performance on this test, aphasia severity ranged from severe to mild ( $M_{\text{aphasia severity}} = 79.43\%$ ,  $SD = 13.88$ , severity range: 42.17 – 99 %). Individual participant data are presented in Appendix A.

## 2.2 Working memory assessment

Participants were assessed with two working memory tasks: an adapted version of a complex span task, the type of a task that has been used predominantly in behavioral studies, and an N-back task, analogous to tasks found in neuroimaging studies. Similar, though not identical verbal stimuli, were used in both tasks to maximize correspondence of processing requirements between tasks. Various measures of performance on these tasks were used as dependent variables in the VLSM analysis.

**2.2.1 Complex span task**—This was a simplified version of the traditional complex listening span task (Conway et al., 2005) adapted specifically for individuals with aphasia. In this modified listening span task, for the processing component, participants had to listen to spoken sentences and match them to the target image in an array of four images on the computer screen and simultaneously, for the storage component, remember a word presented aurally for subsequent recognition. The theoretical and methodological rationale for this task and its features are described in more detail in the publication on the original English version of this task (see Ivanova & Hallowell, 2014). The adapted Russian version of the task (described in detail in Ivanova et al., 2017) included only the condition with short and simple syntactically-reversible sentences (e.g., *“the girl is pushing the boy”*) and high frequency concrete disyllabic words as to-be-remembered items.

For the processing component of the task, visual displays that accompanied sentences included the target picture (i.e., ‘the girl is pushing the boy’), a syntactic distractor (where the agent and patient roles were reversed – ‘the boy is pushing the girl’), a lexical distractor (where the target agent and patient were performing a different action – ‘the girl is drawing the boy’) and a reversed semantic distractor (where the agent and patient roles were reversed while performing a different action – ‘the boy is drawing the girl’). This choice of distractors precluded participants from using any strategies for selection of the matching picture, thus requiring participants to fully attend to the presented sentences. Sentences with corresponding visual arrays were presented interchangeably with words to-be-recalled (storage component). That is, participants had to alternate between two parallel tasks:



comprehension of sentences and remembering a set of distinct items. All the verbal stimuli (sentence and words to-be-remembered) were previously recorded and, along with the visual stimuli (picture arrays for sentences and recall arrays), presented on the computer. The processing component, visual array with four images, was presented to the participants until they provided a response. The interval between these visual arrays was two seconds and, during that time, the word to-be-remembered was aurally presented. At the end of each set, participants had to recall the to-be-remembered words by pointing to the target pictures amongst distractors in a visual array. The number of distractors was always equal to the number of target pictures. This format of non-verbal registration of responses to the processing and recall components minimized the influence of expressive speech and language deficits on performance. See Figure 1 for an example of a set from the complex span task. Set sizes ranged from two to six items per set and were presented to participants in ascending order; three sets at each set size were presented. The examiner recorded each participant's response on a score sheet. Prior to testing, four training sets were provided. Performance on the task was scored based on the maximum set size correctly recalled. Also, the proportion of correctly processed sentences (when the aurally presented sentence was correctly matched to the presented picture) was recorded.

**2.2.2 Verbal 2-back task**—This was a typical 2-back word task with auditory verbal stimuli. Participants were aurally presented with a continuous string of words and were instructed to judge whether a word matched the one they had heard 2 items before. The 2-back task contained 20 different stimuli consisting of high-frequency disyllabic concrete words similar but not identical to the words used for to-be-remembered items in the complex span task. The auditory presentation of stimuli and the use of words with comparable psycholinguistic characteristics allowed us to make the perceptual processing requirements between the two WM tasks maximally similar. Altogether, 150 items containing 36 targets (24%) were presented. Items were presented in 3 blocks of 50 items each with an inter-stimulus interval of two seconds. Participants were allowed to take a break between blocks. The participants responded with their non-dominant left hand by pressing the spacebar on a keyboard to indicate that the current probe was the same as the one presented two probes before. The percentage of targets and the length of the task was selected to be similar to existing tasks in the literature (Christensen & Wright, 2010; Mayer & Murray, 2012; Wright, Downey, Gravier, Love, & Shapiro, 2007). As with the complex span task, all the auditory stimuli were previously recorded; the task was presented on a computer using E-Prime 2.0 software. Performance on the task was indexed via an accuracy measure that was computed by subtracting the false positive rate from the hit rate.

Participants were also presented with 0-back word task as a control condition. The task was analogous to the 2-back task in terms of stimuli and format of presentation, but participants simply had to monitor for presentation of a single target word and press the space bar with their non-dominant left hand whenever they heard it. Altogether, 100 items containing 24 targets (24%) were presented with an inter-stimulus interval of two seconds. Performance on the task was again indexed via an accuracy measure computed by subtracting the false positive rate from the hit rate. Due to a technical error, data for participant # 26 was not

available for this task. The order of 0-back and 2-back tasks was counterbalanced across participants.

### 2.3 MRI data acquisition, preprocessing and lesion reconstruction

MRI data were acquired using a 1.5T Siemens Magnetom Avanto scanner. For each participant we acquired a high-resolution structural T1-weighted volume of the whole brain ( $1 \times 1 \times 1 \text{ mm}^3$  voxels, TR = 1900 ms, TE = 3.37 ms, FoV =  $256 \times 256 \text{ mm}$ , 176 axial slices), T2-weighted ( $0,6 \times 0,4 \times 4 \text{ mm}^3$  voxels, TR = 5000 ms, TE = 93 ms, FoV =  $208 \times 230 \text{ mm}$ , slice thickness = 4 mm, 28 axial slices) and FLAIR ( $0,9 \times 0,9 \times 4 \text{ mm}^3$  voxels, TR = 9000 ms, TE = 89 ms, FoV =  $201 \times 230 \text{ mm}$ , slice thickness = 4 mm, 28 axial slices).

Structural MRI data was preprocessed using SPM8 software (<http://www.fil.ion.ucl.ac.uk/spm>). Prior to lesion reconstruction the three images (T1, T2 and FLAIR) for each participant were first manually reoriented along the AC-PC plane. Then the T1 was co-registered and resliced to MNI152 (Montreal Neurological Institute) template with  $1 \text{ mm}^3$  resolution using 4th degree B-spline transformation, and T2 and FLAIR were yoked and resliced to the new T1 using trilinear transformation.

Lesions were then traced manually in native space based on T1 images in MRIcron (Rorden & Brett, 2000; MRIcron, <http://people.cas.sc.edu/rorden/mricron/index.html>) and ITK-SNAP (Yushkevich et al., 2006; ITK-SNAP, [www.itksnap.org](http://www.itksnap.org)). T2 and FLAIR images were then additionally examined to confirm extent of lesion and to identify regions of hyper- and hypo-intensities surrounding the lesion that would suggest additional white matter degeneration, gliosis and hemosiderin deposition. These regions were also included as a part the lesion. Prior to further processing lesion segmentation was verified by at least three experts in the field.

Finally, the T1 image and subsequently the binary lesion mask were normalized to an MNI template using a modified version of the unified segmentation/normalization algorithm implemented in SPM8 (“Seg” toolbox in the SPM8 distribution; Crinion et al., 2007). This algorithm was customized to optimize normalization of deep white matter and ventricles by using an age relevant template and by additionally incorporating a head model (Turken, D’Esposito, & Dronkers, 2010). Further, a two-step normalization procedure was implemented: in the first step, the major anatomical landmarks were aligned while cost function masking the lesion, in the second step, detailed anatomy was further fine-tuned. The transformations from the two passes were then combined to produce a single deformation field for normalization of the T1 and the binary lesion mask. This approach allows for a tighter fit to the reference space without distorting overall anatomy (Crinion et al., 2007).

At least two experts in the field then inspected reliability of normalization individually for each participant. This was done by yoking the T1-weighted images in native and standard space, overlaying on them the native and normalized lesions respectively, and visually comparing the lesion’s extent in standard space to the original reconstruction in native space. Cases of misalignment, such as lesion masks inside the ventricles or outside the meninges, as well as inconsistent extension to the cortical rim or the side of the ventricles,

were manually corrected using segmentation tools in ITK-SNAP software. We corrected the lesion masks in template space for 39 participants in our cohort. For the majority of participants alterations in template space were limited to elimination of the lesion mask outside of the brain and its correction around the ventricles. Anatomical corrections were only required in a couple of cases with large lesions that distorted the brain anatomy and thus lead to minor misalignment during the normalization. This thorough procedure of lesion segmentation allowed for maximizing the accuracy of the initial segmentation and subsequent normalization.

## 2.4 VLSM analysis

To determine the anatomical correlates of WM, a series of VLSM analyses (Bates et al., 2003) were performed using custom VLSM software (VLSM 2.55; freely available at <https://langneurosci.mc.vanderbilt.edu/resources.html> ). In the VLSM analysis for each voxel linear regression functions are estimated, comparing behavioral performance (scores on the WM tasks: maximum set size correctly recalled for the complex span task and accuracy measure for the 2-back task) in participants with and without a lesion in that voxel. We applied 4mm FWHM Gaussian smoothing to the lesion mask images in order to account for residual anatomical variability post-normalization (as is done in fMRI). Our regression equations computed at each voxel can automatically handle the resulting values between 0 and 1 generated by such smoothing at lesion edges and should result in reducing isolated t-value peaks that are generally considered unimportant while improving our ability to detect true behavior-to-brain correlations via permutation testing. Analyses were confined to those voxels in which at least 10% (n=4) of participants had lesions. This was done in order to minimize potentially spurious results by comparing unbalanced groups. Additionally, participants' age, months post-onset and lesion size were used as covariates in all the VLSM analyses in order to account for possible influence of those variables.

To demonstrate robustness of primary findings VLSM analyses were rerun with additional behavioral covariates in order to take into account varying processing demands of the two tasks. For the complex span task, we used the proportion of correctly-processed sentences as an additional covariate, to account for the possible influence of general language comprehension processing abilities on performance. For the 2-back task, we used accuracy on the 0-back task to account for possible difficulties with single-word auditory comprehension. To further explore the relative importance of different brain regions for the two WM tasks we conducted two supplementary analyses. First, we performed a VLSM analyses of each WM score while using the score from the other WM task as an additional covariate. Second, we used lesion load (i.e., percent damaged) to brain areas associated with impaired performance on the other WM task (as identified in the primary VLSM analyses) as a covariate in the analyses of the two WM tasks. These tests would indicate whether the two WM tasks are supported by distinct brain regions even when common behavioral mechanisms are accounted for (first analysis) and when brain areas relevant for the other task are partialled out (second analysis).

For all of the VLSM analyses outlined above voxel-wise significance thresholding was set at  $p < .001$ , significant clusters were determined by permutation testing (1000 iterations), a

conservative correction method for multiple comparisons (Kimberg, Coslett, & Schwartz, 2007), with cluster threshold set to  $p < .05$ . Permutation-based thresholding is a non-parametric approach to FWER correction and is based on iterative permutations of the actual data – a method adapted from functional neuroimaging (Nichols & Holmes, 2001) and previously shown to provide conservative and reliable correction for multiple-comparisons in mass-univariate lesion mapping approaches (Kimberg et al., 2007; Mirman et al., in press). While voxel-wise permutation thresholding controls familywise error rate robustly, it has been shown to be overly conservative with small to medium sample sizes. Additional limitation of this approach is that it is centered on only a single attribute of the obtained results – extreme statistical scores. However, our results should differ from random noise not only based on statistical scores, but also on spatial coverage, that is, we expect that a cluster of voxels and not a single sparse voxel supports the psychological function under investigation. Since no real conclusion can be made based on a discrete voxel, then, in turn, a single false positive voxel cannot undermine an inference about brain-behavior relations. The cluster-based permutation method takes this into account by controlling for the rate of false positive voxel clusters instead of individual voxels (see Karnath et al., 2018 and Mirman et al., in press, for more on this). Permutations (in our case,  $n = 1000$ ) were performed to determine the null distribution of cluster sizes that pass the voxel-wise cluster-forming threshold ( $p < .001$ ) and use that distribution to set a minimum cluster size threshold that would occur by chance (i.e., if there was no relationship between behavioral scores and lesion patterns) in less than 5% of cases ( $p < .05$ ). In other words, this FWER approach controls the rate of a single false positive cluster of voxels. Voxel-wise cluster-forming thresholds in VLSM studies using permutation-based clustersize thresholding typically range from .05 to .001 (e.g., Banerjee et al., 2015; Binder et al., 2016; Mirman et al., 2015; Pillay, Stengel, Humphries, Book, & Binder, 2014); as stated above we have opted to use the more conservative thresholding level ( $p < .001$ ). Figures 3 and 4 shown below in the Results section include only significant voxels surpassing this permutation-based threshold. Identification of the brain regions associated with the significant voxels was made using the Automated Anatomical Labeling (AAL) atlas and Natbrainlab atlas of white matter pathways (Thiebaut de Schotten et al., 2011) in MRICron.

Also, we conducted preliminary analysis of white matter disconnections contributing to observed WM deficits via track-wise lesion-deficit analysis. This was done by mapping the normalized lesion mask from each patient onto tractography reconstructions of white matter pathways obtained from a group of healthy controls (Rojkova et al., 2016) and quantifying the severity of the disconnection by measuring the probability of the tract to be disconnected (Thiebaut de Schotten et al., 2014) using Tractotron software as part of the BCBtoolkit (Foulon et al., 2018; <http://www.toolkit.bcblab.com>). Following standard recommendations, we considered a given white matter tract to be disconnected (binary measure) if the individual patient lesion overlapped on a voxel within the white matter pathway map with probability of at least 50% (above the chance level). We also measured proportion of tract damaged (i.e., lesion load) by calculating the number of lesioned voxels in the tract divided by the total volume of the tract (thresholded at 50%). We used both the binary measure of disconnection and the continuous measure of proportion of tract damaged in the statistical track-wise lesion-deficit analyses based on linear regression (Chechlacz, Rotshtein, &

Humphreys, 2014). In the linear regression we entered lesion volume, age, months post-onset as covariates (similar to the VLSM analyses above) and each individual pathway disconnection measure/proportion damaged as independent variables to test whether the disconnection/damage within specific pathways (controlling for lesion volume, age, months post-onset) predicted performance on the two WM tasks. We analyzed association (3 segments of the arcuate, cingulum, frontal aslant, frontal inferior longitudinal, frontal orbito-polar, frontal superior longitudinal, inferior frontal occipital, inferior longitudinal, superior longitudinal I, II, III, uncinata), commissural (corpus callosum, anterior commissure) and projection (anterior thalamic, cortico-spinal, fornix, fronto-striatal) tracts in the left hemisphere – a total of 20 tracts. Binary disconnection measure and continuous measure of tract damaged were included in separate regression models for each of the tracts. Each tract-wise lesion deficits analysis was subjected to Bonferroni correction for multiple comparisons ( $\alpha$  level;  $p = .0025$  based on 20 tracts analyzed).

## 2. RESULTS

### 3.1 Behavioral findings

Mean maximum set size recalled for the complex span task was 4 ( $SD=1.15$ , range: 2 to 6), indicating a range of performance. Participants generally performed well on the processing component of the task, demonstrating that they were paying sufficient attention to this component of the task ( $M = 90.64 \pm 10.69\%$ , range: 55 to 100%). Performance on the 2-back task demonstrated similar variability. Mean accuracy of performance (proportion of correct hits minus proportion of false alarms) was 85.65 % ( $SD = 14.31\%$ , range: 36.11 to 100%). Mean accuracy of performance on 0-back task was 99.35% ( $SD = 1.39$ , range: 95.83 to 100%). Pearson correlation between performance on the complex span and 2-back tasks was  $r = .366$ ,  $p = .01$ . Individual participant scores for all WM tasks are included in Appendix A. Histograms of scores on both WM tasks along with the scatterplot of the correlation between the two WM scores are presented in Appendix B.

### 3.2 VLSM findings

The distribution of lesion locations for all 49 individuals is presented in Figure 2. It should be noted that the lesion overlay map only shows those brain regions affected in a minimum of four participants in each voxel consistent with the constraints of the VLSM (described above), where analyses were confined to those voxels in which at least 10% (4 individuals) have a lesion. Thus, the lesion overlay clearly outlines those brain regions about which specific conclusions may be drawn based on the analyses performed below. Here the greatest degree of overlap is seen in the periventricular white matter underlying the frontal, temporal, and parietal lobes in the left hemisphere. In the cortex, the lesion overlap was greatest in the left perisylvian cortex and insula.

First, to identify brain regions required for performance on the WM tasks, we ran two separate VLSM analyses using the maximum set size from the complex span task and accuracy measure from the 2-back task as dependent variables. In both analyses, the standard set of covariates – lesion size, age and time post-onset – was used. The VLSM analysis of performance on the complex span task revealed one significant cluster (volume =

6457 voxels,  $p = .011$ ) with t-values ranging from 3.29 to 4.12, located predominantly in the left inferior frontal area. The peak MNI coordinates were  $(-28\ 35\ 2)$ , and the center coordinates were  $(-34\ 27\ -5)$ . The VLSM analysis of performance on the 2-back task exposed one significant cluster (volume = 4359 voxels,  $p = .035$ ) with t-values ranging from 3.29 to 4.56, located primarily in the left superior and middle temporal gyri. The peak MNI coordinates were  $(-58\ -47\ -6)$ , and the center coordinates were  $(-62\ -20\ 0)$ . Results of these analyses are overlaid on an MNI template in Figure 3. As can be clearly seen in Figure 3, complex span and 2-back maps (shown in orange and blue, respectively) are completely distinct, with no overlap between the two maps. Then using AAL and Natbrainlab atlases we determined which specific anatomical areas were covered by the two VLSM maps. These regions are presented in Table 1. Only areas that were at least 3% covered by the VLSM map for either task are listed. Note that the VLSM map for the 2-back task did not contain any white matter pathways as identified by the standardized atlas.

Second, to further isolate regions specifically involved in WM processes, we ran another set of VLSM analyses with extra behavioral covariates (in addition to the standard set) in order to take into account language processing requirements of the two tasks. For the complex span task, we used the proportion of correctly-processed sentences as an additional covariate, for the 2-back task – accuracy on the 0-back task. The VLSM analysis of performance on the complex span task showed one significant cluster (volume = 12522 voxels,  $p = .001$ ) with t-values ranging from 3.29 to 4.83, predominantly covering the left inferior frontal gyrus. The peak MNI coordinates were  $(-41\ 31\ -4)$ , and the center coordinates were  $(-36\ 26\ -5)$ . The VLSM analysis of performance on the 2-back task revealed one significant cluster (volume = 3723 voxels,  $p = .036$ ) with t-values ranging from 3.30 to 5.04, primarily encompassing the left superior temporal gyrus. The peak MNI coordinates were  $(-58\ -47\ -6)$ , and the center coordinates were  $(-62\ -19\ -3)$ . Obtained VLSM maps were largely identical to the areas identified in the primary analyses and are shown in Figure 4 (again orange for the complex span task and blue for the 2-back task). Once more, the complex span and 2-back maps remain distinct with no common anatomical areas or voxels. Anatomical regions according to AAL and Natbrainlab atlases covered by the VLSM maps (at least 3%) are presented in Table 2.

Finally, we run a set of supplementary analyses to further corroborate our main findings and highlight the relative importance of the identified regions. VLSM analyses of each WM score while using the score from the other WM task as a covariate again revealed distinct cortical regions, generally similar to areas identified in the primary analyses. When for the complex span task analysis accuracy measure from the 2-back task was used as a behavioral covariate, one significant cluster was identified (volume = 10053 voxels,  $p = .004$ ), predominantly covering the left inferior frontal area and basal ganglia structures (inferior frontal triangular, inferior frontal orbital, insula, caudate, putamen, pallidum). When similar analysis was done for the 2-back task with maximum set size recalled from the complex span task as a covariate, no significant clusters were identified at the initial threshold level (volume = 1427,  $p = .109$ ). When we rerun the same analysis with a more lenient voxel-wise threshold of  $p < .005$ , then a significant cluster was identified (volume = 9847, corrected  $p = .044$ ), encompassing again superior and middle temporal gyri. Next to evaluate the relative contribution of the identified brain areas to the two tasks we calculated lesion load (i.e.,

percent of area damaged) for each participant for each significant cluster identified in the primary analyses (maps presented in Figure 2, VLSM analyses of the two main variables with the standard set of covariates). We then used this lesion load as an additional covariate in our VLSM analyses. When the lesion load for the 2-back VLSM map was used as a covariate in the analysis of the complex span task, one significant cluster was identified (volume = 15234 voxels,  $p = .001$ ), predominantly covering the left inferior frontal area and basal ganglia structures (inferior frontal triangular, inferior frontal orbital, insula, caudate, putamen). Similarly, when the lesion load for the complex span VLSM map was used as a covariate in the analysis of the 2-back task, one significant cluster was identified (volume = 6909,  $p = .018$ ), encompassing once again superior and middle temporal gyri. As expected, when lesion load for the same region was used as a covariate for the analysis (i.e., lesion load to the region critical for the complex span task was used as a covariate for that particular task) no significant clusters were identified for either of the WM tasks. Detailed results and anatomical regions according to AAL and Natbrainlab atlases covered by the VLSM maps from these supplementary analyses are provided in Appendix B.

We also conducted an exploratory analysis to investigate the role of white matter pathways supporting performance on our two WM tasks. Linear regression analyses (controlling for lesion volume, age, months post-onset) with disconnection or proportion damaged measure for each individual tract were conducted separately for each WM task. We only report negative standardized beta coefficients here, as those imply that damage to a tract (greater probability of disconnection or greater proportion damaged) leads to compromised WM performance, thus reflecting possible underlying neural mechanism of impairment. Positive beta coefficients in turn imply the opposite (greater damage leads to better performance) and simply reflect the fact that lesion elsewhere in the brain has a positive effect on the process of interest. For instance, higher probability of disconnection of the anterior thalamic projection led to better performance on the 2-back task ( $\beta = .515$ ,  $p = .001$ ). This finding implies that individuals with frontal lesions were more likely to be unimpaired on the 2-back task. This is logical given that typically frontal lesions do not extend into the temporal lobe – areas found to be critical for performance on the 2-back task. Only proportion damaged of the inferior frontal occipital fasciculus (IFOF;  $\beta = -.717$ ,  $p < .001$ ) and uncinate fasciculus (UF;  $\beta = -.555$ ,  $p = .001$ ) significantly contributed to impaired performance on the complex span task. When we further added lesion load of the VLSM map identified in our primary analysis to the model to see whether damage to these tracts predicted WM impairment beyond the contribution of the regions identified in the VLSM analysis, both tracts became non-significant. We failed to identify any tracts within the left hemisphere as being crucial for performance on the 2-back task.

### 3. DISCUSSION

The goal of the present study was two-fold. First, we wanted to demonstrate the critical neural substrate of verbal WM using VLSM in a large group of 49 individuals with left hemisphere stroke. In the era of functional neuroimaging, it is also important to provide complementary findings using contemporary lesion methods to delineate brain regions that are critical for different psychological functions. The second aim was to discern whether different brain regions are involved in different WM tasks, namely the complex span and

word N-back. Although these two tasks have been used interchangeably in neuroimaging research, behavioral studies suggest that they actually share little in common (Ivanova et al., 2017; Jaeggi et al., 2010; Kane et al., 2007; Redick & Lindsey, 2013).

Foremost, we discovered differentiated brain regions critical for successful performance on the complex span and word N-back tasks in the frontal and temporal lobes, respectively. Compared to previous functional neuroimaging studies, a more refined neural substrate was revealed in the current study. Specifically, results of the VLSM analysis demonstrated that critical regions for successful performance on the complex span task were in the inferior frontal gyrus (particularly in the triangular and orbital regions), insula and subcortical white matter (particularly the IFOF and UF). For the 2-back task, critical areas were found in the superior and middle temporal gyri. As predicted, the regions necessary for performance on each WM task were non-intersecting (see Figure 3). Also, when further VLSM analyses were conducted with supplementary behavioral covariates to account for the different language processing requirements of the two tasks, the critical regions associated with each task remained largely unchanged and non-overlapping (see Figure 4). The identified areas remained significant when the score of the other WM task was used as a covariate. Further, when the area critical for the other WM task was taken into account the spatial distribution of the results largely remained unchanged. This showed that the two WM tasks were supported by distinct brain regions even when common behavioral mechanisms and brain areas relevant for the other task were accounted for. In sum, these additional analyses demonstrate the robustness of the primary findings and clearly imply that the identified neural substrates are critical for core cognitive processes involved in the execution of each of these tasks. The difference between the two tasks was also evident in the behavioral data, as the low correlation between the two tasks accounted for only 13% of the variance. Taken together, this supports the initial hypothesis that these two tasks involve distinct cognitive mechanisms – rehearsal and cue-dependent selection in the complex span task, versus updating/auditory recognition in the 2-back task (Ivanova et al., 2017; Redick & Lindsey, 2013) – and are supported by separate neural substrates.

As postulated in previous research on WM, lateral frontal areas become involved when active manipulation or rehearsal of information held in WM (Baldo & Dronkers, 2006; Müller et al., 2002), or, in other words, strategic control of WM contents is required (D'Esposito et al., 2006; Müller & Knight, 2006). Our results demonstrated that regions previously reported to be involved in these exact processes, the inferior frontal gyrus and insula, in both functional neuroimaging (Chein et al., 2011; M. Osaka et al., 2003; N. Osaka et al., 2004) and lesion studies (Baldo & Dronkers, 2006; Szczepanski & Knight, 2014) were essential for successful performance on the complex span task. In this task, participants are required to shift swiftly between several concurrent tasks: processing of sentences and encoding a separate set of words for later recall into short-term memory while constantly rehearsing previously-presented items (Conway et al., 2005). These processing requirements clearly explain the role that the inferior frontal gyrus plays in performance on that task, as that region has been known to be important for sub-vocal rehearsal and coordinating multiple representations in WM (Flinker et al., 2015; Novick, Trueswell, & Thompson-Schill, 2005; Rogalsky & Hickok, 2011). Not surprisingly, the vital role of the inferior frontal gyrus only increased once performance on the purely language component of the task



was taken into account (see Table 2). Contribution of the IFOF and UF to WM has been reported previously (Charlton, Barrick, Lawes, Markus, & Morris, 2010; Golestani et al., 2014). Possibly white matter integrity along these pathways reflects information processing capacity. Demonstrated involvement of these association fibers offers preliminary evidence that a larger network requiring coordination between multiple brain regions is needed for successful performance on the complex span task.

Implicated in performance on the complex span task subcortical structures, the caudate nucleus, putamen and globus pallidus, also play an important role in cognitive processes related to WM. The caudate nucleus, that has been shown to have a strong functional relationship with executive frontal areas (Grahn, Parkinson, & Owen, 2009), might contribute to performance by selection of appropriate sub-goals based on an evaluation of action outcomes (Grahn, Parkinson, & Owen, 2008). Activation foci in the caudate have been noted in functional neuroimaging studies of WM (Levitt et al., 2002). The putamen, previously only considered important for motor functioning, is now also being related to learning and higher cognitive abilities (Ell, Helie, & Hutchinson, 2012), with circumscribed damage to the left putamen also leading to WM impairments (Sefcsik et al., 2009). The globus pallidus has been proposed as the gate-keeper for representations to enter WM (McNab & Klingberg, 2008). While determination of the specific and differential role of these subcortical structures in WM is beyond the scope of the current study, it is clear that the dorsal-prefrontal subcortical loop involving the striatum is important for performance on WM tasks requiring shifting between competing processes and inhibition of irrelevant stimuli (O'Reilly & Frank, 2006; Voytek & Knight, 2010). Clearly, these basal ganglia structures contribute to some core domain-general WM processes, as their involvement in execution of the complex span task becomes pronounced once processing properties of the 2-back task and damage to cortical areas associated with it are partialled out (see Table 1 & 2 in Appendix B).

In contrast to the complex span task, processing requirements of the N-back task are drastically different, as it simply requires updating the contents of WM. Correspondingly, successful performance on this task only needs preserved temporal regions, specifically superior and middle temporal gyri. It has recently been postulated that WM tasks involve primary and secondary perceptual areas of the stimulus modality implicated (Chein et al., 2011; D'Esposito & Postle, 2015). In other words, these domain-specific areas seem to have a dual function: precise sensory encoding, and short-term storage of the incoming information. Involvement of such areas in the temporal lobes seems to be sufficient for the word 2-back task, as no rehearsal is required to hold just 3 items in the focus of attention and, consequently, no cue-dependent recall mechanism is implicated either. Performance on the N-back task did not require any basal ganglia structures that form the dorsal-prefrontal subcortical loops (again in stark contrast to the complex span task). Also, no contribution of white matter pathways to task performance was detected in our preliminary track-wise lesion-deficit analysis. Together these findings further reinforce our claim that the neural network supporting performance on the 2-back task is largely circumscribed to the temporal lobe.

It should be noted that the version of the N-back task used in the current study did not include any lures, that is items matching the target in  $n - 1$  or  $n + 1$  positions. Here we tried to use a version of the N-back task that was comparable to previous investigations using N-back tasks in studies with patients with brain injury and at the same time similar in design to most N-back tasks used in functional neuroimaging studies, that typically do not include lures. It is well known that lack of lures greatly simplifies the task, as the task can be successfully performed based on familiarity alone as the basis for a correct response (Kane et al., 2007). Conversely, when lures are included it becomes necessary to remember the order of items alongside their content and be able to unambiguously recollect what item appeared  $n$ -items back (Harbison, Atkins, & Dougherty, 2011). Possibly, if lures were included, then more frontal areas would have been involved in performance; though, this might have made the task too difficult for patients with brain injury. Also, previous lesion studies show that prefrontal areas become involved in the N-back task only when complexity of the task is elevated to the level of 3-back, possibly necessitating recruitment of control and rehearsal mechanisms (Volle et al., 2008). In contrast, the current 2-back version of the task could be completed via simple updating and matching of items in the focus of attention, leading to exclusive involvement of temporal regions.

While the critical contribution of temporal regions to performance on the N-back task is expected, one might wonder why in the current study there was greater contribution of anterior versus posterior temporal regions. Anterior temporal regions have not been typically implicated in WM tasks in either the fMRI or lesion literature. It is possible that activation is not typically seen in the anterior temporal lobes in functional neuroimaging investigations because these areas are particularly vulnerable to field inhomogeneities and magnetic susceptibility artefacts (Devlin et al., 2000). While the anterior temporal lobes were not considered classical language areas, an accumulation of recent evidence coming from studies with various clinical populations repeatedly point to their role in lexical-semantic and syntactic processing (Gorno-Tempini et al., 2004; Mesulam et al., 2013; Tsapkini, Frangakis, & Hillis, 2011; Wilson et al., 2014), necessitating their inclusion in future models of language processing (Friederici & Singer, 2015). Thus, if one considers anterior temporal regions as language areas, it is not surprising to see their involvement alongside more typical areas for primary phonological and lexical language processing, such as middle and posterior portions of the superior and middle temporal gyri. Also, evidence from studies of anterior temporal lobe resection in epilepsy patients highlights the potential role of this region in processes somewhat related to WM (Stretton et al., 2014). Another explanation for more anterior versus posterior involvement might be due to a limitation of the current study, as better coverage of the anterior compared to posterior temporal regions was present in our cohort. Larger scale lesion studies with more expansive and uniform coverage of the left hemisphere are needed to clarify this and to discern the specific role of the anterior temporal regions in language processing, in general, and verbal WM, in particular.

As has been stated above, complex span tasks involve rehearsal, shifting flexibly between tasks or mental sets, while N-back tasks primarily rely on updating, that is rapid addition/deletion of WM contents (Conway et al., 2005; Miyake & Friedman, 2012). Previous research investigating these aspects of executive functions indicates that the processes of shifting and updating, while moderately related, are clearly separable (Miyake et al., 2000).

Another perspective on the distinction in neural substrates required for performance on the two tasks might be that prefrontal cortex becomes involved only when active maintenance and controlled cue-dependent search of memory is required (Unsworth & Engle, 2007), while simple recognition mechanisms can be accommodated by modality specific perceptual processing areas. In addition, the two tasks target distinctive cognitive processes according to Cowan's embedded processes WM model (Cowan, 1999), with N-back involving only retrieval from the central focus of attention and complex span depending on processing and retrieval of items in both the focus of attention and activated long-term memory. The findings of the current study lend support to the theoretical claims made within these approaches to WM.

One surprising finding from the current study is the lack of necessity of parietal areas for either of our two WM tasks, while according to neural models of WM it is the inferior parietal areas that provide the substrate for the focus of attention/temporary storage within WM (Chein, Ravizza, & Fiez, 2003). Potentially, temporal regions act as alternative short-term storage buffers for verbal WM. This might be in line with the functional neuroimaging literature, where typically activation foci for word recall tasks do not encompass the parietal areas (Chein et al., 2011). While the inferior parietal cortex has been previously implicated in span tasks in lesion studies (Baldo & Dronkers, 2006), they have generally also included the posterior temporal cortex (Baldo et al., 2012), which we believe is actually more critical for temporary storage of verbal stimuli in simple recall tasks according to contemporary views of WM as encompassing related domain-specific areas (D'Esposito & Postle, 2015). Also, it is possible that we did not find significant contribution of the inferior parietal cortex to WM tasks, as only a limited number of patients had lesions in those regions.

### **Limitations of the current study**

As with all VLSM studies, the conclusions that can be drawn from the current findings are limited by the lesion coverage of the patient sample (see Figure 2). Consequently, no inferences can be made regarding the role of the right hemisphere or any other area that has not sustained damage in sufficient numbers of patients. For example, it would be incorrect to conclude that inferior parietal cortex is not important for WM performance, as there simply were not enough patients to detect its role. Thus, it remains to be determined what role regions outside of our coverage area might play in sustaining verbal WM.

Also, inferences about the role of specific white matter pathways in verbal WM remain limited. Analyses using Tractotron provides some insight into what tracts might be implicated in performance on each WM task, however, they should be regarded with caution and only suggestive of possible contribution or lack thereof for a particular tract. Moreover, in the current analysis we could not detect specific contributions of tracts beyond regions identified in the primary VLSM analyses. Future studies using advanced diffusion-weighted sequences and tractography techniques will help to determine the specific role of various white matter bundles in sustaining WM.

## 4. CONCLUSIONS

It should be made clear that the findings of the current study do not imply that we have established all the necessary and sufficient areas for verbal WM, or that word N-back tasks in general do not index verbal WM memory, or that all versions of the N-back task (such as 3- or 4-back versions, or variations with lures) will not require the frontal areas for successful execution. Rather, the behavioral and neural results clearly suggest that the two tasks in their current form involve separable cognitive mechanisms and underlying brain substrate. Therefore, they should not be used interchangeably to investigate verbal WM memory, as has often been the case previously. Instead, future research in this area should attempt to map onto the brain different cognitive mechanisms contributing to efficient WM performance. Second, the study clearly highlights that the lesion method – even in the era of functional neuroimaging – continues to make important contributions to delineating the neural substrate underlying psychological functions, allowing researchers to pinpoint areas critical for the process under investigation. Future VLSM studies supplemented by specialized imaging modalities, more advanced multivariate approaches and connectome-based analyses (Karnath et al., 2018) recruiting larger and more varied patient groups will enable researchers to further discern the specific neural substrate of WM.

Overall, the present study demonstrates that the neural substrates of verbal WM can best be revealed by taking into account the cognitive components that comprise each task and avoiding the assumption that the two tasks can be used interchangeably. Our results highlight the multidimensional nature of WM and the role of executive functioning mechanisms in supporting WM performance. Both functional neuroimaging studies and those using advanced lesion methods will surely benefit by probing the sub-mechanisms of verbal WM to refine existing neural models of WM and establish the indispensable regions for WM processing.

## Acknowledgements

We are extremely grateful to the staff of the Center for Speech Pathology and Neurorehabilitation in Moscow, Russia for assistance with participant recruitment. We thank Ekaterina Iskra and Anastasia Kobzeva for their help with data collection and transcription. We are especially grateful to all participants who took part in the study. Also, we would like to thank the members of the Aphasia Center – Juliana Baldo, Krista Schendel, Carl Ludy, Brian Curran, XJ Kang, and Rita Barakat – for their insightful comments on an earlier version of this manuscript, and Timothy Herron for his guidance and help with neuroimaging analyses.

### Funding

The article was prepared within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) and supported within the framework of a subsidy by the Russian Academic Excellence Project ‘5–100’. This project was also supported by a Research Career Scientist Award to the senior author by the Department of Veterans Affairs Clinical Science Research and Development Program and the NIH/NIDCD grant R01 DC016345. The contents reported within do not represent the views of the Department of Veterans Affairs or the United States Government.

## APPENDIX A

Table A.1

Individual characteristics of participants.

No	Age	Sex	Years of education	Months post-onset	Etiology	Localization of lesion based on MRI report	Type of aphasia	ASA total score	Complex span task-maximum set size recalled	Complex span task- proportion of correctly processed sentences	2-back task- accuracy rate	2-back task- hit rate	2-back task- false positive rate	0-back task- accuracy rate
1	55	f	12	100	hemorrhagic	LH	non-fluent	89.83%	2	0.92	0.82	0.83	0.02	1.00
2	47	f	13	19	ischemic	distribution of LMCA central and cortical branches	fluent	83.67%	5	0.95	1.00	1.00	0.00	1.00
3	40	m	12	26	hemorrhagic	LH	dysarthria	-	5	1.00	0.95	0.97	0.02	1.00
4	70	f	15	16	ischemic	LMCA distribution	fluent	60.50%	2	0.77	0.58	0.61	0.03	0.96
5	40	f	15	12	hemorrhagic	left temporal lobe	fluent	96.17%	4	0.98	0.76	0.78	0.02	1.00
6	50	m	12	59	ischemic	LMCA distribution	non-fluent	90.00%	6	0.98	0.97	0.97	0.00	1.00
7	57	f	12	23	ischemic	distribution of LMCA cortical branches	non-fluent	97.17%	5	1.00	0.94	0.94	0.01	1.00
8	43	m	15	10	hemorrhagic	deep sections of left fronto-temporal region	dysarthria	-	4	0.95	0.99	1.00	0.01	1.00
9	50	m	10	5	ischemic	LMCA distribution	fluent	89.67%	6	0.95	0.96	0.97	0.01	1.00
10	39	f	12	48	ischemic	LMCA distribution	non-fluent	95.00%	6	0.98	0.98	1.00	0.02	1.00
11	48	m	15	49	ischemic	LMCA distribution	non-fluent	75.17%	6	0.95	1.00	1.00	0.00	1.00
12	59	f	15	36	hemorrhagic	left temporal lobe	fluent	94.33%	5	0.88	0.55	0.61	0.06	1.00
13	66	f	10	36	hemorrhagic	left fronto-parietal region	non-fluent	92.67%	4	0.92	0.91	1.00	0.09	0.99
14	50	m	15	51	ischemic	LMCA distribution, lacune in right frontal lobe	mixed	78.67%	5	1.00	0.94	0.97	0.04	1.00
15	63	f	10	17	ischemic	LMCA distribution	non-fluent	90.50%	4	1.00	0.83	0.83	0.00	1.00
16	68	f	15	3	ischemic	distribution of the LMCA, cortical branches, lacune in left frontal lobe	fluent	80.83%	3	0.78	0.85	0.89	0.04	1.00
17	54	f	12	12	hemorrhagic	left basal ganglia, changes in right parietal lobe	mixed	88.00%	6	0.97	0.84	0.86	0.02	1.00
18	50	m	12	29	ischemic	LMCA distribution	non-fluent	77.33%	3	0.97	0.72	0.89	0.17	0.99
19	48	m	12	4	ischemic	left temporoparietal region	non-fluent	42.17%	4	0.65	0.88	0.92	0.04	0.96
20	41	f	15	28	ischemic	LMCA distribution	mixed	88.17%	3	0.98	0.97	1.00	0.03	0.96
21	67	m	15	10	recurrent ischemic	LMCA distribution	non-fluent	69.67%	3	0.95	0.91	0.92	0.01	0.99
22	42	f	15	14	ischemic	LH	non-fluent	99.00%	4	0.95	0.95	0.97	0.03	1.00
23	51	f	15	52	ischemic	LMCA distribution	non-fluent	61.00%	3	0.73	0.96	1.00	0.04	1.00
24	56	m	15	10	hemorrhagic	left thalamus	mixed	92.67%	4	0.98	0.93	0.94	0.02	1.00
25	53	m	15	37	hemorrhagic	left temporoparieto-occipital region	fluent	72.17%	4	0.92	0.98	1.00	0.02	0.99
26	54	m	10	6	ischemic	distribution of LMCA central branches	non-fluent	72.33%	3	0.93	1.00	1.00	0.00	n/a
27	73	f	15	4	ischemic	LMCA distribution	fluent	47.17%	4	0.93	0.91	0.92	0.01	1.00
28	50	m	12	26	ischemic	LMCA distribution	fluent	83.17%	4	0.97	0.94	0.94	0.01	1.00
29	52	m	10	13	ischemic	LMCA distribution	fluent	44.67%	2	0.62	0.56	0.56	0.00	1.00
30	44	f	15	9	ischemic	left temporoparietal region	non-fluent	77.00%	5	0.93	0.89	0.89	0.00	1.00
31	33	f	15	26	ischemic	lentiform nucleus, left frontal lobe	non-fluent	92.83%	4	1.00	0.88	0.92	0.04	1.00
32	44	m	12	19	ischemic	left temporoparietal region	fluent	78.17%	4	0.73	0.58	0.58	0.00	1.00

No	Age	Sex	Years of education	Months post-onset	Etiology	Localization of lesion based on MRI report	Type of aphasia	ASA total score	Complex span task - maximum set size recalled	Complex span task - proportion of correct responses	2-back task - accuracy rate	2-back task - hit rate	2-back task - false positive rate	0-back task - accuracy rate
33	48	m	10	9	ischemic	LACA distribution	non-fluent	63.00%	2	0.93	0.72	0.72	0.00	1.00
34	68	f	10	15	ischemic	left fronto-temporal region	non-fluent	85.00%	5	0.95	0.99	1.00	0.01	1.00
35	55	m	12	5	hemorrhagic	left fronto-temporal region	fluent	62.50%	2	0.82	0.79	0.83	0.04	0.99
36	44	m	12	6	ischemic	left temporoparietal region	fluent	65.33%	5	0.83	0.75	0.75	0.00	0.96
37	40	f	12	4	hemorrhagic	left fronto-temporal region	non-fluent	83.33%	4	1.00	0.52	0.56	0.04	1.00
38	52	f	12	3	ischemic	left fronto-parietotemporal region	fluent	81.33%	4	0.55	0.87	0.92	0.04	1.00
39	53	m	10	3	hemorrhagic	left putamen; lacune in left frontal lobe	fluent	78.83%	4	0.97	0.85	0.86	0.01	1.00
40	37	m	15	146	ischemic	left fronto-temporal region	mixed	85.17%	2	0.98	0.82	0.83	0.01	1.00
41	49	f	12	7	ischemic	subcortical, vascular encephalopathy	dysarthria	-	5	0.93	0.88	0.89	0.01	1.00
42	34	f	12	39	ischemic	left basal ganglia, cortical atrophy in insular region	non-fluent	91.67%	5	0.98	1.00	1.00	0.00	1.00
43	57	f	12	121	hemorrhagic	left fronto-parietotemporal region	non-fluent	70.33%	4	0.83	0.67	0.81	0.14	0.96
44	55	f	15	4	ischemic	left fronto-parietotemporal region	mixed	71.67%	4	0.77	0.82	0.83	0.01	1.00
45	54	f	12	55	ischemic	left fronto-parietotemporal region	non-fluent	79.00%	3	0.87	0.94	0.94	0.01	1.00
46	53	m	15	3	hemorrhagic	left basal ganglia	mixed	91.50%	3	1.00	1.00	1.00	0.00	1.00
47	60	m	15	8	ischemic	left fronto-insular region	fluent	83.00%	3	0.95	0.85	0.97	0.12	1.00
48	74	m	15	4	hemorrhagic	left temporal lobe	dysarthria	-	5	0.92	0.88	0.89	0.01	1.00
49	57	f	12	9	ischemic	left fronto-temporal region	fluent	83.00%	4	0.90	0.87	0.92	0.04	0.96

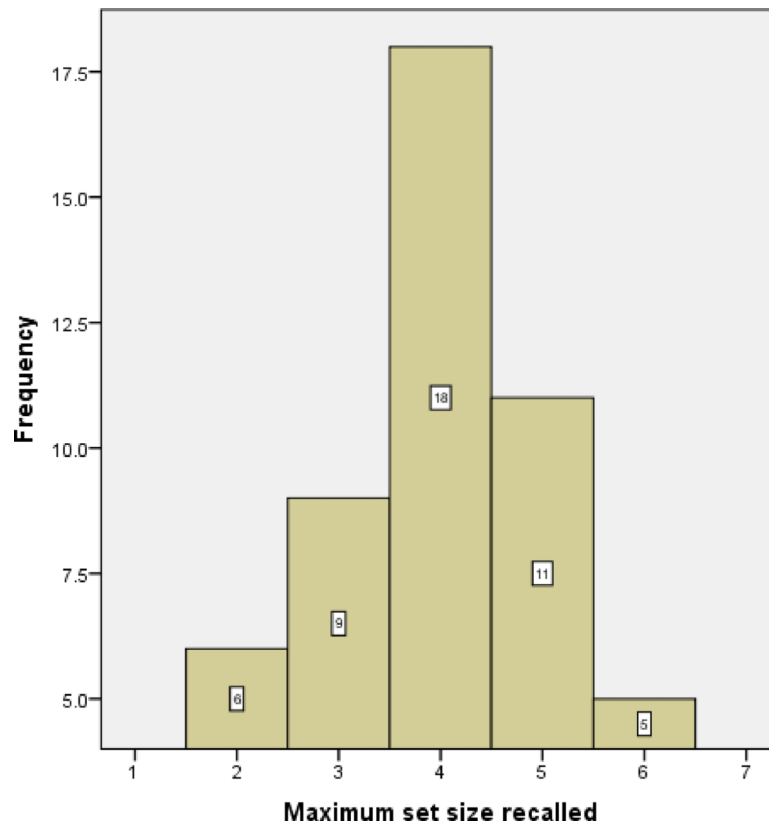
## APPENDIX B

Author Manuscript

Author Manuscript

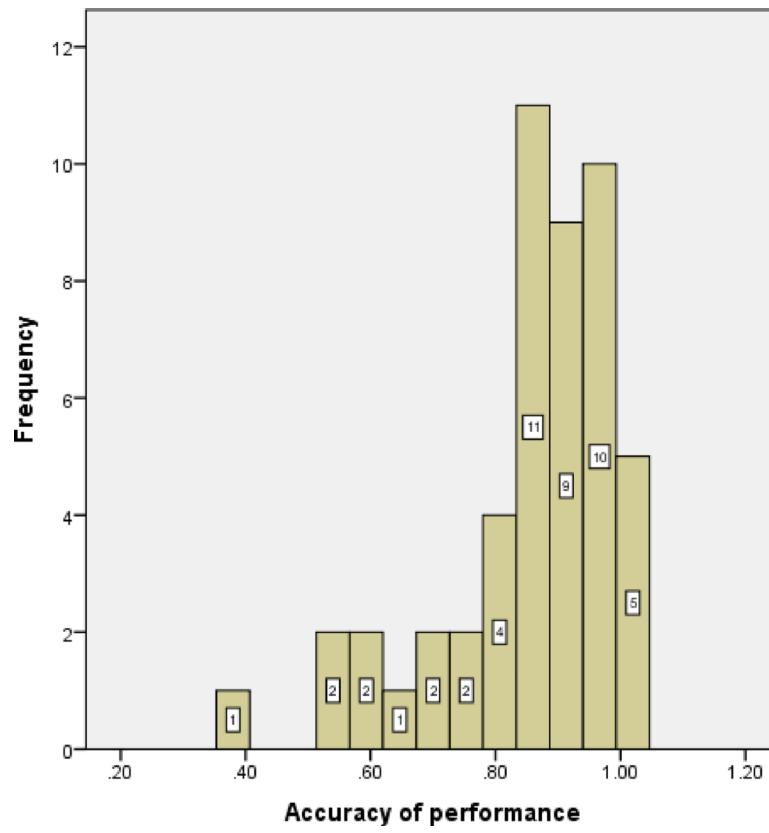
Author Manuscript

Author Manuscript

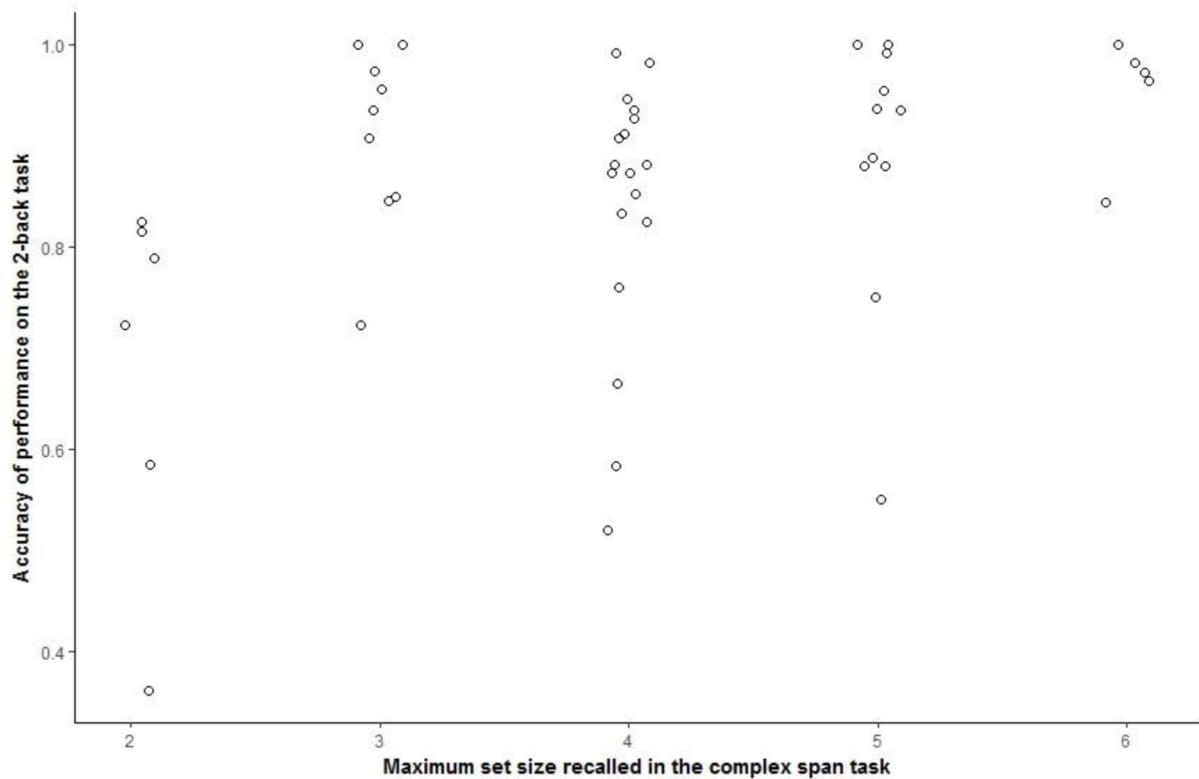


**Figure B.1.**  
Histogram of scores on the complex span task





**Figure B.2.**  
Histogram of scores on the 2-back task



**Figure B.3.**

Jittered scatterplot between WM scores from the complex span task and from the 2-back task.

### Supplementary analyses

When for the complex span task analysis accuracy measure from the 2-back task was used as a behavioral covariate, one significant cluster was identified (volume = 10053 voxels,  $p = .004$ ) with t-values ranging from 3.29 to 4.29, predominantly covering the left inferior frontal area and basal ganglia structures. The peak MNI coordinates were  $(-23\ 13\ -9)$ , and the center coordinates were  $(-29\ 22\ -2)$ . When similar analysis was done for the 2-back task with maximum set size recalled from the complex span task as a covariate with a voxel-wise threshold of  $p < .005$ , then a significant cluster was identified (volume = 9847,  $p = .044$ ) with t-values ranging from 2.7 to 4.17, encompassing again superior and middle temporal gyri. The peak MNI coordinates were  $(-65\ -5\ -1)$ , and the center coordinates were  $(-62\ -25\ 5)$ . See Table B.1 for a list of anatomical regions covered by these VLSM maps.

**Table B.1.**

List of anatomical regions covered by the VLSM maps of the complex span and 2-back tasks while using the score from the other WM task as a covariate (all regions are located in the left hemisphere).

	Complex span task		2-back task	
	Number of voxels covered	Percent of area covered	Number of voxels covered	Percent of area covered
<b>AAL atlas</b>				
Inferior frontal triangular	859	4.3%	0	0%
Inferior frontal orbital	3021	22.2%	0	0%
Insula	1944	12.9%	0	0%
Caudate	1171	15.2%	0	0%
Putamen	1385	17.4%	0	0%
Pallidum	124	5.4%	0	0%
Superior temporal	0	0%	4502	24.6%
Middle temporal	0	0%	4440	11.3%
<b>Natbrainlab atlas</b>				
Inferior occipital frontal fasciculus	1345	12.2%	0	0%
Uncinate fasciculus	53	7.3%	0	0%

*Note.* Results for the 2-back task are presented with a voxel-wise threshold of  $p < .005$ .

When the lesion load for the 2-back task VLSM map was used as a covariate in the analysis of the complex span task, one significant cluster was identified (volume = 15234 voxels,  $p = .001$ ) with t-values ranging from 3.29 to 4.83, predominantly covering the left inferior frontal area and basal ganglia structures. The peak MNI coordinates were  $(-33\ 12\ -12)$ , and the center coordinates were  $(-31\ 22\ -3)$ . Similarly, when the lesion load for the complex span task VLSM map was used as a covariate in the analysis of the 2-back task, one significant cluster was identified (volume = 6909,  $p = .018$ ) with t-values ranging from 3.29 to 4.84, encompassing once again superior and middle temporal gyri. The peak MNI coordinates were  $(-58\ -47\ -6)$ , and the center coordinates were  $(-62\ -24\ 2)$ . See Table B.2 for a list of anatomical regions covered by these VLSM maps.

**Table B.2.**

List of anatomical regions covered by the VLSM maps of the complex span and 2-back tasks while covarying for damage to brain areas associated with impaired performance on the other WM task (all regions are located in the left hemisphere).

	Complex span task		2-back task	
	Number of voxels covered	Percent of area covered	Number of voxels covered	Percent of area covered
<b>AAL atlas</b>				
Inferior frontal triangular	1961	9.8%	0	0%

	Complex span task		2-back task	
	Number of voxels covered	Percent of area covered	Number of voxels covered	Percent of area covered
Inferior frontal orbital	4171	30.7%	0	0%
Insula	3742	24.9%	0	0%
Caudate	1372	17.9%	0	0%
Putamen	911	11.5%	0	0%
Superior temporal	0	0%	2864	15.6%
Middle temporal	0	0%	3580	9.1%
<b>Natbrainlab atlas</b>				
Inferior occipital frontal fasciculus	1455	13.2%	0	0%
Uncinate fasciculus	761	10.4%	0	0%

## REFERENCES

- Akhutina T (2015). Luria's classification of aphasias and its theoretical basis. *Aphasiology*, (August), 1–20. 10.1080/02687038.2015.1070950
- Awh E, Jonides J, Smith EE, Schumacher EH, Koeppel RA, & Katz S (1996). Dissociation of Storage and Rehearsal in Verbal Working Memory: Evidence from Positron Emission Tomography. *Psychological Science* 10.1111/j.1467-9280.1996.tb00662.x
- Babajani-Feremi A, Narayana S, Rezaie R, Choudhri AF, Fulton SP, Boop FA, ... Papanicolaou AC (2016). Language mapping using high gamma electrocorticography, fMRI, and TMS versus electrocortical stimulation. *Clinical Neurophysiology*, 127(3), 1822–1836. 10.1016/j.clinph.2015.11.017 [PubMed: 26679420]
- Baddeley A (2003). Working memory and language: an overview, 36, 189–208. 10.1016/S0021-9924(03)00019-4
- Baldo JV, Katsseff S, & Dronkers NF (2012). Brain regions underlying repetition and auditory-verbal short-term memory deficits in aphasia: Evidence from voxel-based lesion symptom mapping. *Aphasiology*, 26(3–4), 338–354. 10.1080/02687038.2011.602391 [PubMed: 24976669]
- Baldo JV, & Dronkers NF (2006). The Role of Inferior Parietal and Inferior Frontal Cortex in Working Memory, 20(5), 529–538. 10.1037/0894-4105.20.5.529
- Banerjee P, Leu K, Harris RJ, Cloughesy TF, Lai A, Nghiemphu PL, ... Ellingson BM (2015). Association between lesion location and language function in adult glioma using voxel-based lesion-symptom mapping. *NeuroImage: Clinical*, 9, 617–624. 10.1016/j.nicl.2015.10.010 [PubMed: 26740915]
- Barbey AK, Colom R, Paul EJ, & Grafman J (2014). Architecture of fluid intelligence and working memory revealed by lesion mapping. *Brain Structure and Function*, 219(2), 485–494. 10.1007/s00429-013-0512-z [PubMed: 23392844]
- Bates E, Wilson SM, Saygin AP, Dick F, Sereno MI, Knight RT, & Dronkers NF (2003). Voxel-based lesion-symptom mapping. *Nature Neuroscience*, 6(5), 448–50. 10.1038/nn1050 [PubMed: 12704393]
- Binder JR, Pillay SB, Humphries CJ, Gross WL, Graves WW, & Book DS (2016). Surface errors without semantic impairment in acquired dyslexia: A voxel-based lesion-symptom mapping study. *Brain*, 139(5), 1517–1526. 10.1093/brain/aww029 [PubMed: 26966139]
- Braver TS, Barch DM, Kelley WM, Buckner RL, Cohen NJ, Miezin FM, ... Petersen SE (2001). Direct comparison of prefrontal cortex regions engaged by working and long-term memory tasks. *NeuroImage*, 14(1 Pt 1), 48–59. 10.1006/nimg.2001.0791 [PubMed: 11525336]

- Bunge SA, Klingberg T, Jacobsen RB, & Gabrieli JD (2000). A resource model of the neural basis of executive working memory. *Proceedings of the National Academy of Sciences of the United States of America*, 97(7), 3573–3578. 10.1073/pnas.97.7.3573 [PubMed: 10725372]
- Caplan D, & Waters GS (1999). Verbal working memory and sentence comprehension. *The Behavioral and Brain Sciences*, 22(1), 77–94-126. [PubMed: 11301522]
- Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, & Cohen JD (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, 280(5364), 747–750. 10.1126/science.280.5364.747 [PubMed: 9563953]
- Charlton RA, Barrick TR, Lawes INC, Markus HS, & Morris RG (2010). White matter pathways associated with working memory in normal aging. *Cortex*, 46(4), 474–489. 10.1016/j.cortex.2009.07.005 [PubMed: 19666169]
- Chechla M, Rotshtein P, & Humphreys GW (2014). Neuronal substrates of Corsi Block span: Lesion symptom mapping analyses in relation to attentional competition and spatial bias. *Neuropsychologia*, 64, 240–251. [PubMed: 25281309]
- Chein JM, Moore AB, & Conway ARA (2011). Domain-general mechanisms of complex working memory span. *NeuroImage*, 54(1), 550–9. 10.1016/j.neuroimage.2010.07.067 [PubMed: 20691275]
- Chein JM, Ravizza SM, & Fiez JA (2003). Using neuroimaging to evaluate models of working memory and their implications for language processing. *Journal of Experimental Psychology: Applied*, 9(4), 315–339. 10.1037/1076-898X.9.4.315 [PubMed: 12911604]
- Christensen SC, & Wright HH (2010). Verbal and non-verbal working memory in aphasia: What three n-back tasks reveal. *Aphasiology*, 24(6–8), 752–762. 10.1080/02687030903437690
- Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, & Smith EE (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, 386(6625), 604–611. 10.1038/386604a0 [PubMed: 9121583]
- Cohen L, Martinaud O, Lemer C, Lehericy S, Samson Y, Obadia M, ... Dehaene S (2003). Visual Word Recognition in the Left and Right Hemispheres: Anatomical and Functional Correlates of Peripheral Alexias. *Cerebral Cortex*, 13(12), 1313–1333. 10.1093/cercor/bhg079 [PubMed: 14615297]
- Conway ARA, Kane MJ, Bunting MF, Hambrick DZ, Wilhelm O, & Engle RW (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769–786. [PubMed: 16523997]
- Crinion J, Ashburner J, Leff A, Brett M, Price C, & Friston K (2007). Spatial normalization of lesioned brains: Performance evaluation and impact on fMRI analyses. *NeuroImage*, 37, 866–875. 10.1016/j.neuroimage.2007.04.065
- Curtis CE, Zald DH, & Pardo JV (2000). Organization of working memory within the human prefrontal cortex: A PET study of self-ordered object working memory. *Neuropsychologia*, 38(11), 1503–1510. 10.1016/S0028-3932(00)00062-2 [PubMed: 10906375]
- D'Esposito M, Cooney JW, Gazzaley A, Gibbs SEB, & Postle BR (2006). Is the prefrontal cortex necessary for delay task performance? Evidence from lesion and FMRI data. *Journal of the International Neuropsychological Society: JINS*, 12(2), 248–260. 10.1017/S1355617706060322 [PubMed: 16573858]
- D'Esposito M, & Postle BR (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, 66, 115–142. 10.1146/annurev-psych-010814015031.THE
- Daneman M, & Merikle PM (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3(4), 422–33. 10.3758/BF03214546 [PubMed: 24213976]
- Devlin JT, Russell RP, Davis MH, Price CJ, Wilson J, Moss HE, ... Tyler LK (2000). Susceptibility-induced loss of signal: Comparing PET and fMRI on a semantic task. *NeuroImage*, 11(6 I), 589–600. 10.1006/nimg.2000.0595 [PubMed: 10860788]
- Ell SW, Helie S, & Hutchinson S (2012). Contributions of the putamen to cognitive function. In Costa A & Villalba E (Eds.), *Horizons in Neuroscience Research* (pp. 29–52).
- Flinker A, Korzeniewska A, Shestuyk AY, Franaszczuk PJ, Dronkers NF, Knight RT, & Crone NE (2015). Redefining the role of Broca's area in speech. *Proceedings of the National Academy of Sciences*, 112(12), 3697–3702. 10.1073/pnas.1418811112 [PubMed: 25811111]

- Sciences of the United States of America, 112(9), 2871–5. 10.1073/pnas.1414491112 [PubMed: 25730850]
- Foulon C, Cerliani L, Kinkingnéhun S, Levy R, Rosso C, Urbanski M, ... Thiebaut de Schotten M (2018). Advanced lesion symptom mapping analyses and implementation as BCBtoolkit. *GigaScience*, (2 2018). 10.1093/gigascience/giy004
- Friederici AD, & Singer W (2015). Grounding language processing on basic neurophysiological principles. *Trends in Cognitive Sciences*, 1–10. 10.1016/j.tics.2015.03.012
- Gathercole SE, & Baddeley AD (1993). *Working memory and language*. East Sussex, UK: Psychology Press.
- Golestani AM, Miles L, Babb J, Castellanos FX, Malaspina D, & Lazar M (2014). Constrained by Our Connections: White Matter's Key Role in Interindividual Variability in Visual Working Memory Capacity. *Journal of Neuroscience*, 34(45), 14913–14918. 10.1523/JNEUROSCI.2317-14.2014 [PubMed: 25378158]
- Gorno-Tempini ML, Dronkers NF, Rankin KP, Ogar JM, Phengrasamy L, Rosen HJ, ... Miller BL (2004). Cognition and anatomy in three variants of primary progressive aphasia. *Annals of Neurology*, 55(3), 335–46. 10.1002/ana.10825 [PubMed: 14991811]
- Grahn JA, Parkinson JA, & Owen AM (2008). The cognitive functions of the caudate nucleus. *Progress in Neurobiology*, 86(3), 141–155. 10.1016/j.pneurobio.2008.09.004 [PubMed: 18824075]
- Grahn JA, Parkinson JA, & Owen AM (2009). The role of the basal ganglia in learning and memory: Neuropsychological studies. *Behavioural Brain Research*, 199(1), 53–60. 10.1016/j.bbr.2008.11.020 [PubMed: 19059285]
- Harbison JI, Atkins SM, & Dougherty MR (2011). N-back Training Task Performance: Analysis and Model Experiment. *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, 120–125. Retrieved from <https://mindmodeling.org/cogsci2011/papers/0026/paper0026.pdf>
- Ivanova MV, Isaev DY, Dragoy OV, Akinina Yu S, Petryshevskii AG, Fedina ON, ... Dronkers NF (2016). Diffusion-tensor imaging of major white matter tracts and their role in language processing in aphasia. *Cortex*, 1–17. 10.1016/j.cortex.2016.04.019
- Ivanova MV, Kuptsova SV, & Dronkers NF (2017). A comparison of two working memory tasks in aphasia. *Aphasiology*, 31(3), 265–281. 10.1080/02687038.2016.1172699
- Ivanova MV, & Hallowell B (2014). A new modified listening span task to enhance validity of working memory assessment for people with and without aphasia. *Journal of Communication Disorders*, 52, 78–98. 10.1016/j.jcomdis.2014.06.001 [PubMed: 24986153]
- Jaeggi SM, Buschkuhl M, Perrig WJ, & Meier B (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, 18(4), 394–412. 10.1080/09658211003702171 [PubMed: 20408039]
- Just MA, & Carpenter PA (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, 99(1), 122–49. [PubMed: 1546114]
- Just MA, Carpenter PA, & Keller TA (1996). The capacity theory of comprehension: New frontiers of evidence and arguments. *Psychological Review*, 103(4), 773–780. 10.1037/0033-295X.103.4.773 [PubMed: 8888654]
- Kane MJ, Conway ARA, Miura TK, & Colflesh GJH (2007). Working Memory, Attention Control, and the N-Back Task: A Question of Construct Validity, 33(3), 615–622. 10.1037/0278-7393.33.3.615
- Kane MJ, Hambrick DZ, Tuholski SW, Wilhelm O, Payne TW, & Engle RW (2004). The Generality of Working Memory Capacity: A Latent-Variable Approach to Verbal and Visuospatial Memory Span and Reasoning, 133(2), 189–217. 10.1037/00963445.133.2.189
- Karnath H-O, Sperber C, & Rorden C (2018). Mapping human brain lesions and their functional consequences. *NeuroImage*, 165, 180–189. 10.1016/j.neuroimage.2017.10.028 [PubMed: 29042216]
- Kimberg DY, Coslett HB, & Schwartz MF (2007). Power in Voxel-based Lesion – Symptom Mapping. *Journal of Cognitive Neuroscience*, 19, 1067–1080. 10.1162/jocn.2007.19.7.1067 [PubMed: 17583984]

- Kondo H, Morishita M, Osaka N, Osaka M, Fukuyama H, & Shibasaki H (2004). Functional roles of the cingulo-frontal network in performance on working memory. *NeuroImage*, 21(1), 2–14. 10.1016/j.neuroimage.2003.09.046 [PubMed: 14741637]
- Levitt JJ, McCarley RW, Dickey CC, Voglmaier MM, Niznikiewics MA, Seidman LJ, & et al. (2002). MRI study of caudate nucleus volume and its cognitive correlates in neuroleptic-naïve patients with schizotypal personality disorder. *American Journal of Psychiatry*, 159(7), 1190–1197. 10.1038/jid.2014.371 [PubMed: 12091198]
- Luria AR (1980). *Higher cortical functions in man* (2nd ed.). New York, NY: Basic Books.
- Ma WJ, Husain M, & Bays PM (2014). Changing concepts of working memory. *Nature Neuroscience*, 17(4), 347–356. 10.1038/nn.3655. [PubMed: 24569831]
- Mayer JF, & Murray LL (2012). Measuring working memory deficits in aphasia. *Journal of Communication Disorders*, 45(5), 325–39. 10.1016/j.jcomdis.2012.06.002 [PubMed: 22771135]
- McCloskey M, & Schubert T (2014). Shared versus separate processes for letter and digit identification. *Cognitive Neuropsychology*, 31(5–6), 437–460. 10.1080/02643294.2013.869202 [PubMed: 24404944]
- McNab F, & Klingberg T (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*, 11(1), 103–107. 10.1038/nn2024 [PubMed: 18066057]
- Mesulam MM, Wieneke C, Hurley R, Rademaker A, Thompson CK, Weintraub S, & Rogalski EJ (2013). Words and objects at the tip of the left temporal lobe in primary progressive aphasia. *Brain*, 136(2), 601–618. 10.1093/brain/aws336 [PubMed: 23361063]
- Mirman D, Chen Q, Zhang Y, Wang Z, Faseyitan OK, Coslett HB, & Schwartz MF (2015). Neural organization of spoken language revealed by lesion-symptom mapping. *Nature Communications*, 6, 1–9. 10.1038/ncomms7762
- Mirman D, Landrigan J, Kokolis S, Verillo S, Ferrara C, & Pustina D (2017). Corrections for multiple comparisons in voxel-based lesion-symptom mapping. *Neuropsychologia*, (August), 1–12. 10.1016/j.neuropsychologia.2017.08.025
- Miyake A, & Friedman NP (2012). The nature and organisation of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 21(1), 8–14. 10.1177/0963721411429458 [PubMed: 22773897]
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, & Wager TD (2000). The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. 10.1006/cogp.1999.0734 [PubMed: 10945922]
- Müller NG, & Knight RT (2006). The functional neuroanatomy of working memory: Contributions of human brain lesion studies. *Neuroscience*, 139(1), 51–58. 10.1016/j.neuroscience.2005.09.018 [PubMed: 16352402]
- Müller NG, Machado L, & Knight RT (2002). Contributions of Subregions of the Prefrontal Cortex to Working Memory: Evidence from Brain Lesions in Humans. *Journal of Cognitive Neuroscience*, 14(5), 673–686. 10.1162/08989290260138582 [PubMed: 12167253]
- Nestor A, Behrmann M, & Plaut DC (2013). The neural basis of visual word form processing: A multivariate investigation. *Cerebral Cortex*, 23(7), 1673–1684. 10.1093/cercor/bhs158 [PubMed: 22693338]
- Nichols TE, & Holmes AP (2001). Nonparametric permutation tests for functional neuroimaging experiments: A primer with examples. *Human Brain Mapping*, 15(1), 1–25. 10.1002/hbm.1058
- Novick JM, Trueswell JC, & Thompson-Schill SL (2005). Cognitive control and parsing: Reexamining the role of Broca’s area in sentence comprehension. *Cognitive, Affective and Behavioral Neuroscience*, 5(3), 263–281. 10.3758/CABN.5.3.263
- O’Reilly RC, & Frank MJ (2006). Making Working Memory Work: A Computational Model of Learning in the Prefrontal Cortex and Basal Ganglia. *Neural Computation*, 18(2), 283–328. 10.1162/089976606775093909 [PubMed: 16378516]
- Osaka M, Osaka N, Kondo H, Morishita M, Fukuyama H, Aso T, & Shibasaki H (2003). The neural basis of individual differences in working memory capacity: An fMRI study. *NeuroImage*, 18(3), 789–797. 10.1016/S1053-8119(02)00032-0 [PubMed: 12667855]

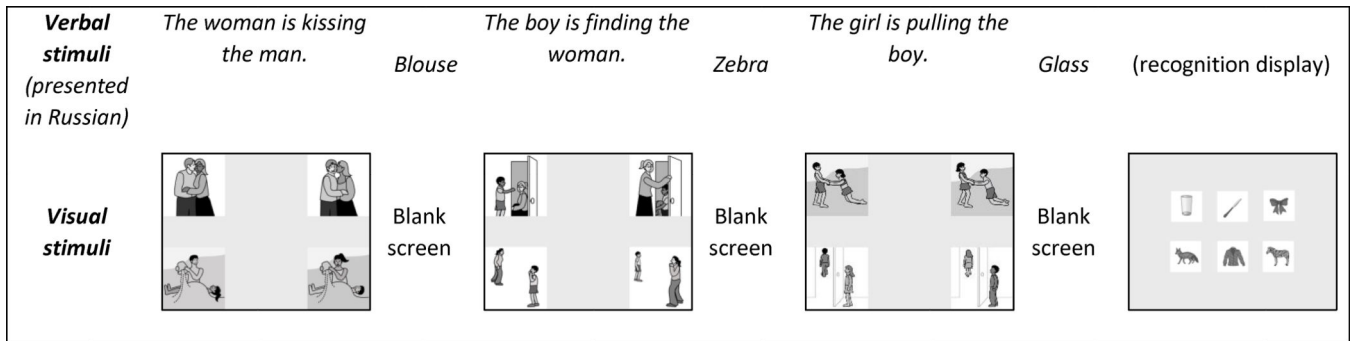
- Osaka N, Osaka M, Kondo H, Morishita M, Fukuyama H, & Shibasaki H (2004). The neural basis of executive function in working memory: An fMRI study based on individual differences. *NeuroImage*, 21(2), 623–631. 10.1016/j.neuroimage.2003.09.069 [PubMed: 14980565]
- Ottenhausen M, Krieg SM, Meyer B, & Ringel F (2015). Functional preoperative and intraoperative mapping and monitoring: Increasing safety and efficacy in glioma surgery. *Neurosurgical Focus*, 38(January), 1–13. 10.3171/2014.10.FOCUS14611.Disclosure
- Owen AM, McMillan KM, Laird AR, & Bullmore E (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, 25(1), 46–59. 10.1002/hbm.20131 [PubMed: 15846822]
- Pillay SB, Stengel BC, Humphries C, Book DS, & Binder JR (2014). Cerebral localization of impaired phonological retrieval during rhyme judgment. *Annals of Neurology*, 76(5), 738–746. 10.1002/ana.24266 [PubMed: 25164766]
- Redick TS, & Lindsey DRB (2013). Complex span and n-back measures of working memory: A meta-analysis. *Psychonomic Bulletin and Review*, 20(6), 1102–1113. 10.3758/s13423-013-0453-9 [PubMed: 23733330]
- Reynolds JR, West R, & Braver T (2009). Distinct neural circuits support transient and sustained processes in prospective memory and working memory. *Cerebral Cortex*, 19(5), 1208–1221. 10.1093/cercor/bhn164 [PubMed: 18854581]
- Rogalsky C, & Hickok G (2011). The role of Broca's area in sentence comprehension. *J Cogn Neurosci*, 23(7), 1664–1680. 10.1162/jocn.2010.21530 [PubMed: 20617890]
- Rojkova K, Volle E, Urbanski M, Humbert F, Acqua FD, & Schotten MT De. (2016). Atlasing the frontal lobe connections and their variability due to age and education: a spherical deconvolution tractography study. *Brain Structure and Function*, 1751–1766. 10.1007/s00429-015-1001-3 [PubMed: 25682261]
- Rorden C, & Brett M (2000). Stereotaxic Display of Brain Lesions. *Behavioural Neurology*, 12(4), 191–200. 10.1155/2000/421719 [PubMed: 11568431]
- Rorden C, & Karnath H-O (2004). Using human brain lesions to infer function: a relic from a past era in the fMRI age? *Nature Reviews. Neuroscience*, 5(10), 813–9. 10.1038/nrn1521 [PubMed: 15378041]
- Rottschy C, Langner R, Dogan I, Reetz K, Laird AR, & Schulz JB (2012). NeuroImage Modelling neural correlates of working memory: A coordinate-based meta-analysis. *NeuroImage*, 60(1), 830–846. 10.1016/j.neuroimage.2011.11.050 [PubMed: 22178808]
- Sefcsik T, Nemeth D, Janacek K, Hoffmann I, Scialabba J, Klivenyi P, ... Vecsei L (2009). The role of the putamen in cognitive functions — A case study. *Learning & Perception*, 1(2), 215–227. 10.1556/LP.1.2009.2.4
- Simos PGP, Breier JI, Fletcher JM, Foorman BR, Castillo EM, & Papanicolaou AC (2002). Brain mechanisms for reading words and pseudowords: an integrated approach. *Cerebral Cortex*, 12(3), 297–305. 10.1093/cercor/12.3.297 [PubMed: 11839603]
- Smith EE, Geva a, Jonides J, Miller a, Reuter-Lorenz P, & Koeppel R. a. (2001). The neural basis of task-switching in working memory: effects of performance and aging. *Proceedings of the National Academy of Sciences of the United States of America*, 98(4), 2095–2100. 10.1073/pnas.98.4.2095 [PubMed: 11172081]
- Stretton J, Sidhu MK, Winston GP, Bartlett P, McEvoy AW, Symms MR, ... Duncan JS (2014). Working memory network plasticity after anterior temporal lobe resection: A longitudinal functional magnetic resonance imaging study. *Brain*, 137(5), 1439–1453. 10.1093/brain/awu061 [PubMed: 24691395]
- Szczepanski SM, & Knight RT (2014). Review Insights into Human Behavior from Lesions to the Prefrontal Cortex. *Neuron*. 10.1016/j.neuron.2014.08.011
- Thiebaut de Schotten M, ffytche Dominic, H., Bizzi A, Dell F, Allin M, Walshe M, ... Catani M (2011). Atlasing location, asymmetry and inter-subject variability of white matter tracts in the human brain with MR diffusion tractography. *NeuroImage*, 54(1), 49–59. 10.1016/j.neuroimage.2010.07.055 [PubMed: 20682348]
- Thiebaut de Schotten M, Tomaiuolo F, Aiello M, Merola S, Silvetti M, Lecce F, ... Doricchi F (2014). Damage to White Matter Pathways in Subacute and Chronic Spatial Neglect: A Group Study and 2



- Single-Case Studies with Complete Virtual “ In Vivo ” Tractography Dissection. *Cerebral Cortex*, 24(3), 691–706. 10.1093/cercor/bhs351 [PubMed: 23162045]
- Tsapkini K, Frangakis CE, & Hillis AE (2011). The function of the left anterior temporal pole: Evidence from acute stroke and infarct volume. *Brain*, 134(10), 3094–3105. 10.1093/brain/awr050 [PubMed: 21685458]
- Tsuchida A, & Fellows LK (2009). Lesion evidence that two distinct regions within prefrontal cortex are critical for n-back performance in humans. *Journal of Cognitive Neuroscience*, 21(12), 2263–2275. 10.1162/jocn.2008.21172 [PubMed: 19199405]
- Tsvetkova L, Akhutina T, & Pylaeva N (1981). *Kolichestvennaya otsenka rechi u bol'nykh s aphasiej*. Moscow: Izdatelstvo Moskovskogo Gosudarstvennogo Universiteta.
- Turken AU, D'Esposito M, & Dronkers NF (2010). Normalization of stroke patient images for white matter lesion-symptom mapping analysis. In *Organization for Human Brain Mapping*.
- Unsworth N, & Engle RW (2007). The Nature of Individual Differences in Working Memory Capacity: Active Maintenance in Primary Memory and Controlled Search From Secondary Memory, 114(1), 104–132. 10.1037/0033-295X.114.1.104
- Volle E, Kinkingnéhun S, Pochon JB, Mondon K, Thiebaut De Schotten M, Seassau M, ... Levy R (2008). The functional architecture of the left posterior and lateral prefrontal cortex in humans. *Cerebral Cortex*, 18(10), 2460–2469. 10.1093/cercor/bhn010 [PubMed: 18308710]
- Voytek B, & Knight RT (2010). Prefrontal cortex and basal ganglia contributions to visual working memory. *Proceedings of the National Academy of Sciences*, 107(42), 18167–18172. 10.1073/pnas.1007277107
- Wager TD, & Smith EE (2003). Neuroimaging studies of working memory: a meta-analysis. *Cognitive, Affective & Behavioral Neuroscience*, 3(4), 255–74.
- Wilson SM, Demarco AT, Henry ML, Gesierich B, Babiak M, Mandelli ML, ... Gornotempini ML (2014). What Role Does the Anterior Temporal Lobe Play in Sentence-level Processing? Neural Correlates of Syntactic Processing in Semantic Variant Primary Progressive Aphasia, 970–985. 10.1162/jocn
- Wright HH, Downey RA, Gravier M, Love T, & Shapiro LP (2007). Processing distinct linguistic information types in working memory in aphasia. *Aphasiology*, 21(6–8), 802–813. 10.1080/02687030701192414 [PubMed: 19554209]
- Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, & Gerig G (2006). User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*, 31(3), 1116–1128. 10.1016/j.neuroimage.2006.01.015 [PubMed: 16545965]

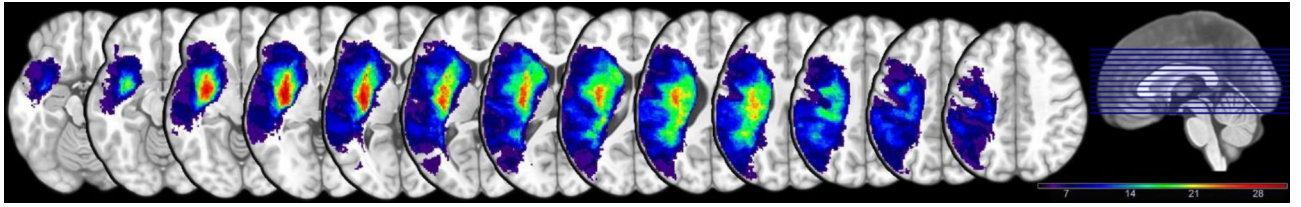
### Highlights

- We determined brain areas required for performance on two common WM tasks using VLSM.
- The left inferior frontal gyrus was critical for performance on the complex span task.
- The left superior and middle temporal gyri were crucial for performance on the N-back task.
- The two WM tasks depended on distinct neural and cognitive mechanisms.
- Results emphasize the multi-dimensional nature of WM.

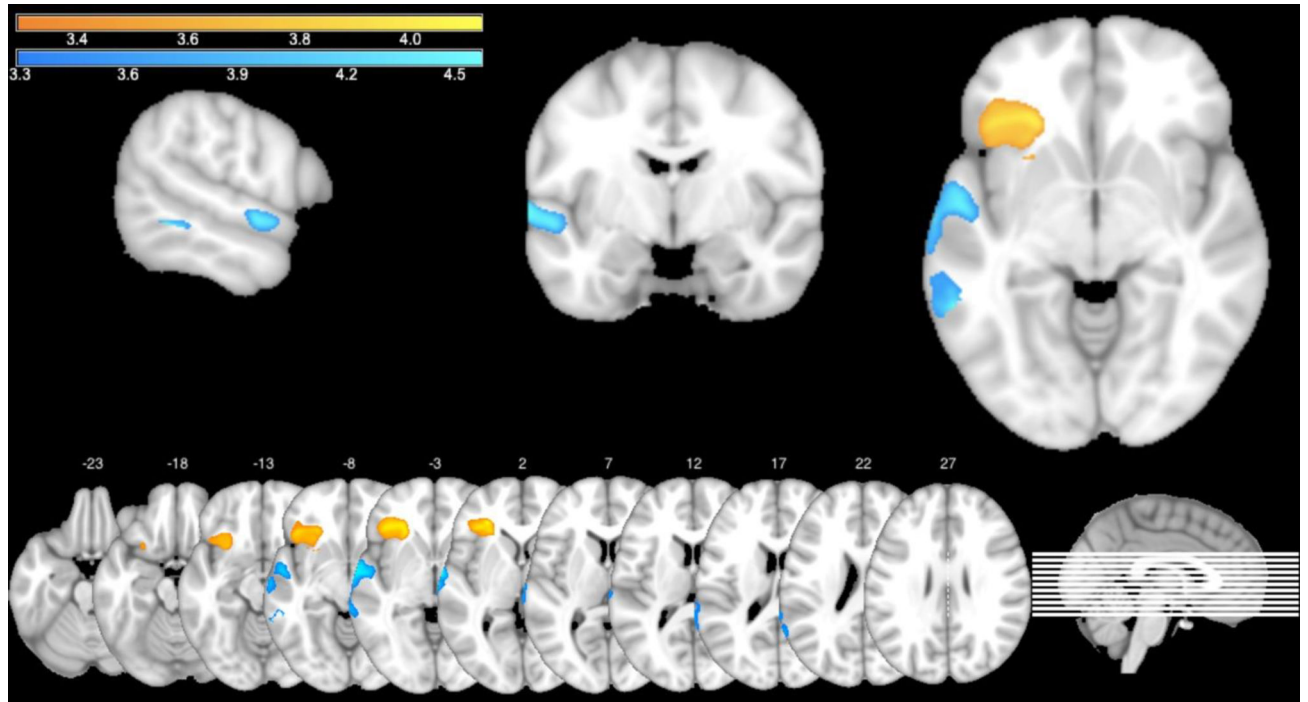


**Figure 1.**

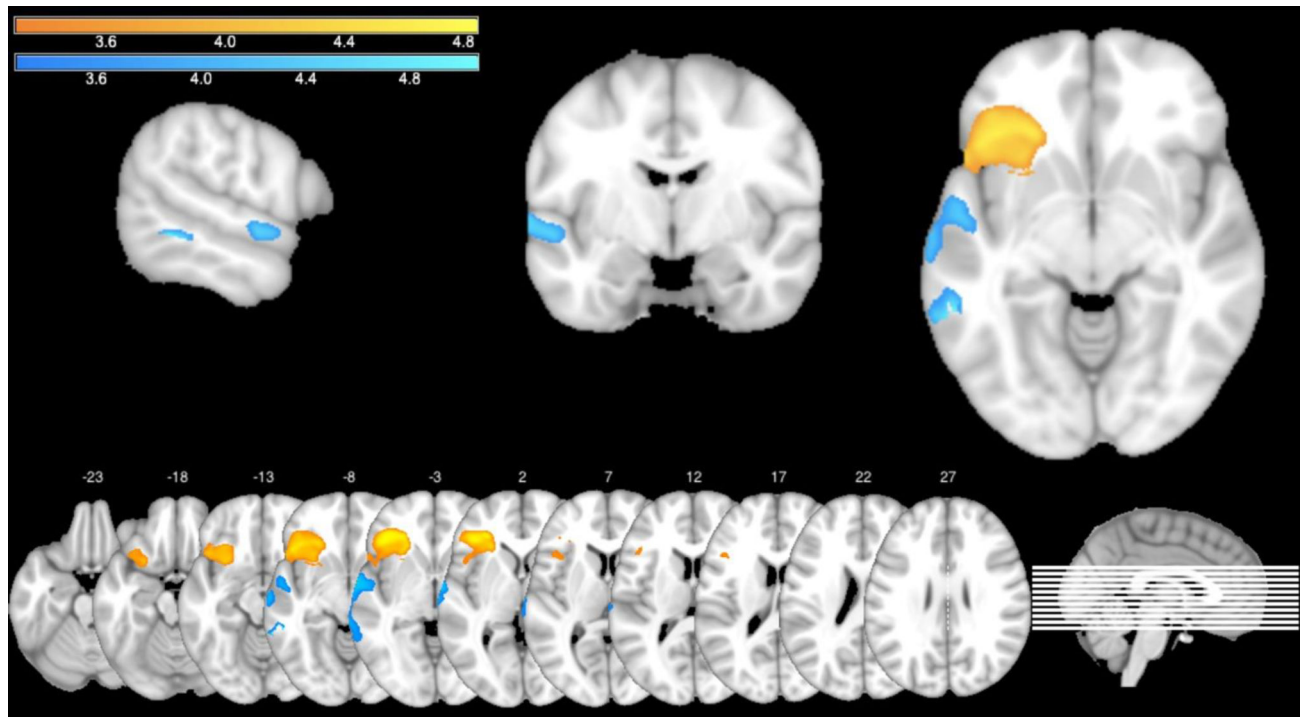
Example of a set from the complex span task (set size three). All words for recognition are disyllabic in Russian.



**Figure 2.** Lesion overlay demonstrating overlap across participants' lesions, with a minimum of 4 participants' lesions in each voxel and a maximum of 31.



**Figure 3.**  
VLSM map of performance on the complex span task (orange) and on the 2-back task (blue). Brighter colors indicate increasing t-values.



**Figure 4.** VLSM map of performance on the complex span task (orange) and on the 2-back task (blue) while taking language processing requirements of the two tasks into account. Brighter colors indicate increasing t-values.

**Table 1.**

List of anatomical regions covered by the VLSM maps of the complex span and 2-back tasks (all regions are located in the left hemisphere).

	Complex span task		2-back task	
	Number of voxels covered	Percent of area covered	Number of voxels covered	Percent of area covered
<b>AAL atlas</b>				
Inferior frontal triangular	837	4.2%	0	0%
Inferior frontal orbital	2776	20.4%	0	0%
Insula	1767	11.8%	0	0%
Superior temporal	0	0%	2001	10.9%
Middle temporal	0	0%	2080	5.3%
<b>Natbrainlab atlas</b>				
Inferior occipital frontal fasciculus	702	6.4%	0	0%

**Table 2.**

List of anatomical regions covered by the VLSM maps of the complex span and 2-back tasks with behavioral covariates (all regions are located in the left hemisphere).

	Complex span task		2-back task	
	Number of voxels covered	Percent of area covered	Number of voxels covered	Percent of area covered
<b>AAL atlas</b>				
Inferior frontal triangular	1974	9.8%	0	0%
Inferior frontal orbital	4820	35.5%	0	0%
Insula	3117	20.7%	0	0%
Superior temporal	0	0%	1492	8.1%
Middle temporal	0	0%	221	5.6%
<b>Natbrainlab atlas</b>				
Inferior occipital frontal fasciculus	983	8.9%	0	0%
Uncinate fasciculus	432	5.9%	0	0%