The inner workings of working memory: Preliminary data from unimpaired populations

Abstract

Wright et al. (2007) tested Persons With Aphasia (PWA) using three N-Back tasks featuring different types of linguistic information – phonological, semantic, and syntactic -- to determine whether Verbal Working Memory (VWM) is a single, united resource. The current study tested two groups of cognitively normal individuals with the same tasks, as well as an additional vision-focused task, to expand on this previous research and provide a baseline for future studies of WM in PWA. Results indicated no effects of aging outside of Reaction Times, and significant differences in performance across all types of information except phonological and visual cues.

Introduction

Past studies have provided evidence that Verbal Working Memory [VWM] is a separate resource from Spatial Working Memory, and that VWM loading can be used to predict how difficult processing a sentence will be (Nation et al., 1999; Gibson, 2000). Some studies suggest that VWM is one unified resource, concerned with both processing and storing verbal information (Just & Carpenter, 1992); others argue that storage and processing are controlled by separate resources (Caplan & Waters, 1999). It is possible that further divisions in VWM exist, dedicated to different types of linguistic information (Wright et al., 2007). Persons With Aphasia (PWA) have difficulty processing syntactically complex sentences, but the exact mechanism causing this difficulty is unknown. PWA could have diminished VWM capacities, which could be syntax-specific or system-wide.

The N-Back task is a common measure of VWM (Chen et al., 2006; Wright et al., 2007). VWM measures are usually preliminary tests in research, used to divide participants into high- and low-capacity groups or to establish correlations between WM span and performance on some other task. Wright et al., for example, correlated verbal N-Back tasks to performance on established language performance tasks. Through these correlations, the authors sought to examine the question of smaller divisions inside VWM. Wright et al.'s study, however, used a small sample of exclusively PWA, and tasks that were exclusively language-related. The current study is part of a larger research effort to expand on Wright et al.'s previous work with the N-Back paradigm, by adding a nonverbal ShapeBack task and comparing aphasic performances to a baseline of cognitively intact subjects. The current study sought to establish this baseline by including two groups of cognitively intact adults who completed Wright et al.'s three language tasks as well as the new ShapeBack task. The results should indicate whether aging has any effect on individuals' performance on N-Back tasks, which will contribute to the interpretation of the performance of the predominantly-older PWA population.

Methods

Two subject groups were included in this study: a younger group, consisting of adults aged 18-35 years, and an older group, aged 50-90 years. All participants were native English speakers with normal or corrected-to-normal vision and hearing and no history of speech-language or cognitive deficits.

As part of screening procedures, participants completed a questionnaire asking about their personal medical history, handedness, language status (i.e. whether they are native speakers of English), and vision status. All participants also underwent a short hearing screening, wherein an audiometer with over-the-ear headphones was used to test hearing of pure tones at 500, 1000, 2000, and 4000 Hz at 30 dB. Participants then completed Raven's Coloured Progressive Matrices, a standardized test of cognitive function.

Once the screening procedures were complete, the participants began the experimental tests. Each of the four n-back tasks began with a short practice section, after which the experimental trials were presented. For both the practice and experimental tasks, a string of words or sentences were played over headphones or a string of visual displays appeared on a computer screen, and the participants responded to stimuli meeting pre-established criteria by pressing a button on a computer keyboard.

Results

Three measures of performance were calculated for each task: a Criterion score (C), a D Prime sensitivity level (D'), and a Reaction Time (RT, in milliseconds). A repeated-measures Analysis of Variance (ANOVA) was performed to analyze the effects of task type, difficulty, and age group on these values. Age group showed a significant main effect for RT ($F = 6.967$, $p = .014$), but not for D' ($F = .589$, $p = .450$) or C ($F = 3.668$, $p = .067$). Task type showed a significant main effect across all three of these measurements:

C: $F = 30.340$, $p < .001$
D': $F = 76.449$, $p < .001$
RT: $F = 228.295$, $p < .001$

A paired-samples t-test revealed that, for younger participants, there was no significant difference between D' measures for PhonoBack and ShapeBack at the 1-back level ($t = -.819$, $p = .427$). For both participant groups, PhonoBack and ShapeBack at the 2-back level showed no significant difference (Younger: $t = -.601$, $p = .557$; Older: $t = -.024$, $p = .981$).

No significant interactions were found between task type and age group or between difficulty level and age group. A three-way analysis of task type, difficulty level, and age group also revealed no significant interactions. Task type and difficulty level showed an interaction in RT ($F = 6.440$, $p = .002$), but not in D' (F = `.432, $p = .243$) or C ($F = 1.424$, $p = .247$).

Discussion

Wright et al.'s original study reported only raw accuracy scores for target items, and showed significant decreases in performance for all tasks from 1-back to 2-back difficulty. The current study, however, found that raw accuracy scores *increased* for the SynBack task from the 1- to 2-back conditions for both the younger and older subject groups (Fig. 1, Fig. 2). This outcome suggested that simple accuracy scores are not the best measure of performance for n-back tasks, because a participant's tendency to respond or to abstain can affect measures of accuracy. A better measure for such tasks is a combination of Criterion scores, which measure how strong the stimulus needs to be before the subject will respond, and D Prime sensitivity scores, which measure how well the subject detects the presence or absence of the target stimulus. With these new measures, subjects showed the expected decrease in performance from the 1- to 2-back conditions across all tasks, including SynBack.

After the original misleading measures were corrected, results fell into the expected patterns. SemBack was consistently the "easiest" of the tasks, as found in Wright et al.'s study, and SynBack was consistently the "hardest." Reaction time showed a supra-additive effect for different tasks as the difficulty level changed. Older subjects demonstrated consistently longer reaction times than the younger subjects, but showed no difference in sensitivity or criterion.

Significant departures from these patterns by aphasic subjects would seem to indicate a qualitative difference in approach to the task.

The surprising absence in variation between Phono- and Shape-Back could indicate that sensory information is processed similarly regardless of modality, in contradiction to Baddeley's model of a phonological loop and a separate visuo-spatial scratchpad. This theory may be supported or disputed by data from PWA – if aphasic participants show a significant deficit in PhonoBack, but not in ShapeBack, Baddeley's theory would be restored. If, however, aphasic participants show a significant deficit in both Phono- and ShapeBack equally, a united theory of WM would be further indicated.

Overall, this data provides a strong baseline for comparison to aphasic participants' performance. Results for normally-functioning adults were consistent across age groups in every aspect except RT, meaning any differences seen in PWA can be more confidently attributed to their condition. Furthermore, this data emphasizes the importance of choosing the correct measurements to take when dealing with the n-back task. The raw accuracy scores used by Wright et al, and initially recorded in the current study, provide a misleading view of subjects' performance and could ultimately lead to mistaken conclusions concerning the underlying mechanisms at work during completion of the n-back task.

References

Caplan, D. & Waters, G. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*, 77-126.

Chen, Y., Mitra, S., & Schlaghecken, F. (2006). Sub-processes of working memory in the N-back task:An investigation using ERPs. *Clinical Neurophysiology, 119,* 1546-59.

Just, M. A. & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122-149.

Gibson, E. (2000). The Dependency Locality Theory: A distance-based theory of linguistic complexity. In Miyashita, Y., Marantz, A. & O'Neil, W. (Eds.) Image, Language, Brain. MIT Press: Cambridge, MA. 95-126.

Nation, K., Adams, J. & Snowling, M. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology, 73*, 139-158.

Wright, H. H., Downey, R. A., Gravier, M., Love, T. & Shapiro, L. P. (2007). Processing distinct linguistic information types in working memory in aphasia. *Aphasiology, 21,* 802-13.

Fig. 1

Younger Mean ACC scores:

		Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12	Y13	Y14	Y15	Mean	SD
Phono1	ACC	0.938462	0.84615	0.98462	0.98462	0.95385	0.86154	0.98462	0.95	0.97	1	1	0.8	0.91	0.91	1	0.94	0.062294
Phono2	ACC	0.934211	0.81579	0.97368	1	0.93421	0.90789	0.86842	0.93	0.96	0.97	0.97	0.78	0.87	0.89	0.99	0.92	0.064337
Sem1	ACC	1	0.98462	1	1	1	1	0.96923	0.98	1	0.98	0.98	0.88	0.98	0.97	1	0.98	0.030472
Sem2	ACC	0.973684	0.93421	1	1	0.97368	1	0.92105	1	1	1	1	0.75	0.97	0.95	0.99	0.96	0.06475
Syn1	ACC	0.753846	0.75385	0.75385	0.75385	0.75385	0.75385	0.73846	0.75	0.75	0.75	0.74	0.71	0.75	0.74	0.72	0.74	0.013395
Syn2	ACC	0.802632	0.84211	0.81579	0.77632	0.73684	0.76316	0.81579	0.74	0.76	0.78	0.78	0.78	0.79	0.75	0.76	0.78	0.029866
Shape1	ACC	0.923077	0.95385	0.93846	0.98462	0.98462	0.96923	0.90769	0.95	1	0.97	0.94	0.86	1	0.94	0.94	0.95	0.037053
Shape2	ACC	0.881579	0.90789	0.86842	0.93421	0.96053	0.96053	0.86842	0.96	0.99	1	0.96	0.79	0.97	0.95	0.95	0.93	0.056442

Fig. 2
Older Mean ACC Scores:

		Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12	Mean	SD
Phono1	ACC	0.94	1.00	0.82	0.861538	0.953846	0.815385	0.861538	0.907692	0.87692	0.98462	0.953846	1	0.91	0.0673668
Phono2	ACC	0.84	0.99	0.92	0.868421	0.855263	0.723684	0.881579	0.828947	0.78947	0.96053	0.986842	0.97	0.88	0.0836848
Sem1	ACC	0.98	1.00	1	1	1	0.984615	1	0.969231	0.96923	1	1	1	0.99	0.0126035
Sem2	ACC	0.97	1.00	0.96	0.960526	0.947368	0.868421	1	0.868421	0.88158	0.96053	1	1	0.95	0.0510762
Syn1	ACC	0.75	0.75	0.72	0.753846	0.753846	0.753846	0.753846	0.723077	0.72308	0.75385	0.753846	0.75	0.74	0.0139176
Syn2	ACC	0.79	0.78	0.75	0.802632	0.802632	0.710526	0.763158	0.75	0.82895	0.76316	0.776316	0.78	0.77	0.0307495
Shape1	ACC	0.98	1.00	0.95	0.984615	0.984615	0.969231	0.953846	0.938462	0.92308	0.95385	1	0.98	0.97	0.0243173
Shape2	ACC	0.95	1.00	0.87	0.842105	0.894737	0.881579	0.894737	0.921053	0.84211	0.93421	0.921053	0.86	0.90	0.0468499