The Working Memory Token Test (WMTT):

Preliminary Findings in Young Adults With and Without Dyslexia

Shimon Sapir, Ph.D. and Ravit Cohen Mimran, Ph.D.
Department of Communication Disorders
Faculty of Social Welfare & Health Studies, University of Haifa, Israel
sapir@research.haifa.ac.il

10.21437/Interspeech.2005-431

Abstract

The purpose of the study was to assess the validity of a new version of the Token Test (TT) which is aimed at detecting deficits in auditory verbal working memory (henceforth WMTT). Twenty eight young healthy adults, 14 with and 14 without dyslexia, all paid volunteers, took part in the study. The WMTT was found to be positively and significantly correlated with tests of digit span forward and backward, pseudoword reading performance, but not with real word reading, performance IQ, and the d-2 attention measure. The WMTT also appeared to be a more sensitive method than the original TT to detect differences in auditory verbal working memory between the two subject groups. These main findings, albeit preliminary, indicate that the WMTT is superior to the original Token Test (memory parts) in assessing auditory verbal working memory deficits.

1. Introduction

The Token Test (TT) [1] has been used extensively in research and clinical practice to screen individuals with deficits in auditory verbal comprehension and short term auditory verbal memory. According to the model of working memory proposed by Baddeley [cf. 2] verbal working memory depends on two sub-systems: a temporal verbal acoustic storage system, also called the phonological loop (PL), and a working memory capacity (WMC). The PL is assumed to be necessary for the immediate retention of sequences of strings of phonemes (words and nonwords). The WMC is assumed to be responsible for attention control of working memory (also called the central executive) where there is a need to combine simultaneous processing and storage [3].

We have revised and computerized the memory part of the TT, so that it can also help detect deficits in WMC in addition to PL deficits. We call this revision the Working Memory TT, or WMTT, which includes items that test both auditory verbal memory span and auditory working memory.

1.1. The WMTT

The original TT tested auditory verbal memory by having the subject point to tokens of 5 different colors, 2 shapes (circle, squares), and 2 sizes (small, large), in response to commands such as "touch the blue circle" (a one dimensional command, or 1-D, since the tokens differ only by color), or "touch the yellow square and the red circle" (a two dimensional command, or 2-D, since the tokens differ by color and/or shape), or "touch the small red circle and the large green square" (a three dimensional command, or 3-D, since the tokens differ by size, and/or color and/or shape). Such commands require that the subject stores the words in memory in the same order that they appear in the command. To include commands that test working memory, we used the words "before" and "after" at the beginning or between two parts of the command, so that the order in of the instructions within a command did not correspond to the order in which these instructions were to be performed by the subjects. For example, the command "before you touch the red circle touch the green square" or the command "touch the red circle after you touch the green square" both require that the subject first perform the second instruction and then the first instruction in the command. We call these types of commands non-serial, and their performance requires manipulation of the words stored in short term memory, as well as some form of inhibition and vigilance to ensure that the appropriate order of performance is achieved in spite of the difference in order of instructions. Thus, the performance of these commands requires use of executive functions, which are most characteristic of working memory [2]. In contrast, when the words "before" and "after" are inserted in the commands such that the execution of these commands is in the same order as the instructions (i.e. serial commands), the task is less likely to involve working memory, and most likely to involve auditory memory span.

1.2. Working memory and reading

There is considerable evidence that individuals with reading disabilities (RD) have an impaired working memory ability for verbal information [4, 5]. However, researchers do not agree where the working memory deficits of individuals with RD originate with respect to the model of working memory proposed by Baddeley and his colleagues. In particular, it is not clear whether the memory problems of individuals with RD include mainly deficits in temporary storage of verbal information [6,7] or are realized mainly when there is a need to activate the central executive [5]. The present study addresses this controversy by assessing the verbal working memory model as a whole rather than focusing on particular aspects of it, as has been done by much of the research to date.

The primary purpose of this investigation was to assess the validity and sensitivity of the WMTT, by correlating it with the forward and backward digit span tests (FDST and BDST, respectively), and by correlating it with pseudoword reading (PWR) (which has been shown to be strongly correlated with auditory working memory) and real word reading (RWR).

The inclusion of the dyslexics in the study is for two main purposes: to form part of a continuity of the level of performance on memory tests, and to assess the extent to which the WMTT can contribute to the differentiation between those individuals with DD and phonologic processing deficits (expressed primarily in terms of poor performance on the PWR) and those without.

In both the serial and nonserial commands, the auditory memory is most likely mediated via the phonologic loop, as argued by Baddeley [2]. Since pseudowords reading (PWR), much more so than real word reading (RWR), requires working memory and the use of the phonologic loop, deficits in working memory should impair PWR, and to a much less extent RWR [4]. With these in mind, we anticipated a stronger correlation between the PWR and the nonserial commands on the WMTT than between the PWR and the serial commands on the WMTT. Also, since the WMTT was designed to tackle working memory deficits, we anticipated that performance on the WMTT would be poorer than on the original TT. We also anticipated that the WMTT will be more effective than the original TT (memory commands) in detecting deficits in working memory and in PWR in individuals with RD, when compared to individuals without RD.

2. Method

The WMTT, FDST, BDST, PWR, and RWR were administered to university students (ages: 23-30) with (n=14; 8 females) and without dyslexia (n=14; 8 females), all paid volunteers. All participants were native Hebrew speakers. Participants were defined as reading disabled when their achievement on a reading diagnostic test [8] fell below two SDs. This test consists of 2 subtests in which respondents are required to read aloud lists of words and pseudowords, respectively, as quickly and as accurately as possible within 1 min. Each test comprises a list of 100 items. Scores are based on the number of word/pseudowords read aloud correctly. The dyslexic and non-dyslexic subjects did not differ significantly in their measure of attention (Test of Attention d2) [9] and Performance IQ [10] ($t(26)=1.57, p>.05$; $t(26)=.36, p>.05$, respectively).

The WMTT consists of 60 commands which are arranged in 2 different parts: two dimensions commands and three dimensions commands. In each part there are 10 serial commands, 10 non-serial commands and 10 of the original TT memory commands (henceforth OTT). The WMTT was administered via a computer, with the tokens presented on the screen immediately after the presentation of the auditory commands. This was done to ensure that the subjects did not rely on visual cues to remember the commands (in the original TT the tokens are placed in front of the participant throughout the testing). The commands were pre-recorded with a young female voice and with clear speech. The auditory commands were presented binaurally at a comfortable loudness level via Sony MDR-P180 stereo headphones. The participant responded by selecting the "tokens" through a mouse click. After the participant responded to a command, there was a one second pause before the next command was presented.

3. Results

3.1. The relationships between memory tasks and reading accuracy

Table 1 shows the results of Pearson inter-correlations between most of the different subtests of the WMTT, the forward and backward digit span tests (henceforth, FDST and BDST, respectively), performance IQ, d2 attention test, real word reading (RWR), and pseudoword reading (PWR). As can be seen, none of the memory subtests correlated significantly with Performance IQ and d2 attention tests. Significant correlations were found between the most of the WMTT subtests and both reading and Digit Span measures. The highest correlations were found between the PWR and the General (i.e., OTT+ serial + nonserial subtests scores combined), the 3D serial and 3D nonserial WMTT commands ($r= 0.65, 0.61, 0.61$, respectively).

Table 1. Pearson correlations between the different subtests of the WMTT, the IQ, d2, the reading tests, and the FDST and BDST. OTT: original Token Test (TT) memory subtests; WMTT – the revised working memory TT.; Perf. IQ-performance IQ; d2 – d2 attention test; RWR- regular word reading; PWR- pseudoword reading; DSF- digit span forward; DSB – digit span backward. 2D-two dimension, 3D-three dimension; * $p< 0.05$; ** $p <0.01$ (two tailed); n=28

	IQ	d2	RWR	PWR	FDS test	BDS test
WMTT General	.32	.26	.48*	.65**	.51**	.54**
2D-OTT	.20	.20	.44*	.47*	.29	.47*
2D-Serial	.27	.07	.18	.40*	.43*	.32
2D-Nonserial	.17	.28	.44*	.38*	.49*	.41*
3D-OTT	.31	.25	.38*	.43*	.18	.28
3D-Serial	.28	.08	.32	.61**	.57**	.51**
3D-Nonserial	.24	.21	.42*	.61**	.38*	.48**

To determine the special contribution (variance) of each of the subtests to both the PWR and RWR, we performed hierarchical regression analyses, whereby we determined the R^2 change. The results, shown in Table 2, indicate that none of the subtests contributed significantly to the RWR, whereas the BDST and the WMTT-3D (serial and nonserial), each contributed significantly to the PWR. The FDST and the WMTT-2D did not contribute significantly to the PWR. The WMTT-3D contributed a unique variance to the PWR task after both BDST and FDST measures were partialed out.

Table 2. Summaries of Hierarchical Regression Analyses. R^2 change. See abbreviations and levels of significance in table 1.

	Reading Accuracy	
	PWR	RWR
Step Variable		
1. FDST	.13	.13
2. BDST	.20*	.11
3. General	.16*	.06
3. 2D-OTT	.05	.06
3. 2D–Serial	.05	-
3. 2D-nonserial	.02	.06
3. 3D-OTT	.08	.06
3. 3D–Serial	.15*	-
3. 3D-nonserial	.15*	.04

3.2. Comparing between RD and control groups

Mean and SDs for the RD group and the control group are presented in Table 3. The RD group scored significantly below the control group on the FDST and BDST, as well as on the 2D non-serial and 3D serial WMTT.

Three factor RM-MANOVA examined main effects of Group (RD x Control) x Level of memory load (2 dimensions x 3 dimensions) and Type of working memory capacity (OTT x serial x non-serial). Results of this analysis showed a significant main effect of Group. Above all conditions, RD group showed significantly lower scores [$F(1,26)= 8.99$, $P<.01$] as compare to the Control group. Also, significant main effect of Level [$F(1,26)= 43.75$, $P<.001$] and Type [$F(2,25)= 16.31$, $P<.001$] were indicated. For both groups, 3 dimensions commands were harder than 2 dimensions commands and non-serial commands were harder than both OTT and serial commands. Additionally, a Group x level x Type interaction was indicated ($F(2, 25)=3.28$, $p=.05$). This interaction stemmed from larger differences between types (serial/OTT vs. nonserial) in the 2 dimensions level for the RD group as compared to the control group.

Table 3: Memory Tasks. Mean (SD). FDST = Forward Digit Span Test; BDST= Backward Digit Span Test; 2D=2 dimensions; 3D=3 dimensions; WMTT= Working Memory Token Test; *P<.05; **P<.01

	RD (N=14)	Control (N=14)	T- value
Digit Span			
FDST	8.93 (1.48)	10.21 (1.53)	2.28 *
BDST	6.73 (2.60)	9.35 (2.16)	2.94 **
WMTT			
2D OTT	8.93 (1.38)	9.64 (.49)	1.97 ns
2D serial	9.40 (.98)	9.86 (.36)	1.75 ns
2D non-serial	7.93 (1.62)	9.26 (1.14)	2.89 **
3D OTT	7.53 (1.59)	8.57 (1.22)	1.82 ns
3D serial	7.27 (2.49)	9.21 (1.05)	2.82 *
3D non-serial	5.80 (2.42)	7.50 (2.21)	1.96 ns

4. Discussion

The present findings are preliminary and their interpretation is tentative at this point. In general, it appears that the WMTT is a valid method for testing auditory verbal working memory, since it correlates significantly and positively with the FDST and BDST, and since it has a unique contribution to the PWR performance but not to the RWR. The highest correlation between the general WMTT scores and the PWR scores suggests that the WMTT indeed tackles working memory deficits.

Importantly, the scores of the 2D nonserial and the 3D serial subtests of the WMTT were significantly different (worse in the RD group) between the two groups of subjects (RD vs. Controls), whereas the scores of the 2D and 3D of the OTT did not differentiate between the groups. These findings suggest that the WMTT is a more sensitive method than the OTT for detecting deficits in auditory verbal working memory in individuals with RD.

In the present study we did not measure reaction time (RT). Since RT is an important index of WM processing [11, 12], it should be included in future studies. Also, in the present study we presented the auditory commands before the subject viewed the tokens. We did so in order to eliminate any visual cues that might help the subject remember the instructions. Thus, the subject had to rely only on his short term auditory memory to perform the task. However, we are not sure to what extent this method is superior to the original method of the TT, whereby the commands to the subject are provided as the subject views the tokens. For example, it is possible that viewing the tokens may confuse rather than aid the subject in remembering the instructions, inasmuch as the visual cues interfere or compete with the auditory cues. Such interference may be especially problematic in individuals with RD, who have been shown to have deficits in executive functions and inhibitory processing of conflicting information [13, 14, 15]. Certainly, a study comparing auditory commands presented alone vs. auditory commands presented simultaneously with the tokens would be helpful in resolving this issue.

The present findings are encouraging, but merit further investigation to assess the psychometric characteristics of the WMTT and the utility of the WMTT as a test of auditory verbal working memory.

5. References

[1] DeRenzi, E., & Vignolo, L. A. (1962). The Token Test: a sensitive test to detect receptive disturbances in aphasics. *Brain*, 85, 556-678.

[2] Baddeley,A.D. (2003). Working memory and language: an overview. *Journal of Communication Disorders*. 36, 189-208.

[3] Baddeley,A.D. (2002). Fractionating the central executive. In D. Stuss & R.T. Knight (Eds.), *Principles of frontal lobe function.* New York: Oxford University Press.

[4] Banai, K. & Ahissar, M. (2004). Poor frequency discrimination probes dyslexics with particularly impaired working memory. *Audiology and Neuro-Otology*, 9, 328-40

[5] De Jong, P.F. (1998). Working memory deficits of reading disabled children. *Journal of Experimental Child Psychology*, 70, 75-96.

[6] Brady, S.A. (1991). The role of working memory in reading disability. In S.A. Brady & Shankweiler (Eds.), *Phonological processes in litracy*: A tribute to Isabelle Y. Liberman. Hillsdale, NJ: Erlbaum.

[7] Morris, R.D., Stuebing, K.K., Fletcher, J.M., Shaywitz, S.E., Lyon, G.R., Shankweiler, D.P., Katz, L., Francis, D.J,. & Shaywitz, B.A. (1998). Subtypes of reading disability: variability around a phonological core. *Journal of Educational Psychology*, 90, 347-373.

[8] Shatil, E. (1995). *One-Minute Test for Words and Pseudowords*. Unpublished test. University of Haifa. Haifa.

[9] Brickenkamp, R., & Zillmer, E. (1998). *The d2 Test of Attention.* (1st US ed.). Seattle, WA: Hogrefe & Huber Publishers.

[10] Wechsler, D. (1991). *Wechsler Intelligence Scale for adults*. Cleveland, OH: Psychological Corp.

[11] Barnea A, Lamm O, Epstein R, & Pratt H. (1994). Brain potentials from dyslexic children recorded during short-term memory tasks. *International Journal of Neuroscience*, 74, 227-37.

[12] Mynatt BT. (1977). Reaction times in a bi-sensory task: implications for attention and speech perception. *Journal of Experimental Psychology and Human Perception and Performance* 3, 316-24.

[13] Brosnan M, Demetre J, Hamill S, Robson K, Shepherd H, Cody G. (2002). Executive functioning in adults and children with developmental dyslexia. *Neuropsychologia*, 40, 2144-55.

[14] Kelly MS, Best CT, & Kirk U. (1989). Cognitive processing deficits in reading disabilities: a prefrontal cortical hypothesis. *Brain and Cognitoin*, 11, 275-93.

[15] Van der Schoot M, Licht R, Horsley TM, Sergeant JA. (2002). Fronto-central dysfunctions in reading disability depend on subtype: guessers but not spellers. Developmental *Neuropsychology*, 22, 533-64.