Distinctions between Spatial and Verbal Working Memory: A Study Using Event-related Potentials

Yung-Nien Chen[1,2], MD, PhD; Suvobrata Mitra[2], PhD

Background: The main manifestation of dementia is a defect in working memory. N-back tasks are frequently used in research on working memory. Researchers can study differences between different loadings by controlling N factors. Furthermore, the interface of N-back tasks can be verbal or visual-spatial.

Methods: Event-related potentials under verbal and spatial tasks and different loadings were recorded using a digital electroencephalogram, and analyzed together with behavior results.

Results: The differences between spatial and verbal processing were found mainly inter-component, where P3 was enhanced in verbal tasks and P2a was enhanced in spatial tasks. Furthermore, P3 was only enhanced in the left hemisphere in the target stimulus. N2 was enhanced by verbal non-target with similar amplitudes. The lateralization was not significant between spatial and verbal tasks.

Conclusion: The difference between spatial and verbal N-back tasks is not only lateralization but also more complex presentations, including P2a (for spatial tasks), P3 (for verbal tasks), and N2 (for non-target detection in verbal tasks).
(Chang Gung Med J 2009;32:380-9)

Key words: N-back task, spatial, verbal, working memory, event-related potentials

The term working memory (WM) encompasses both temporary maintenance as well as manipulation of information required for an ongoing cognitive task. In multi-component models of WM,[1] a domain-general 'executive' subsystem retrieves, maintains and manipulates WM contents, and controls various domain-specific 'slave' subsystems that maintain specific types of information (e.g., verbal, spatial, and visual patterns). Impairment of working memory is thought to be a feature of dementia. There have been extensive investigations of the neural substrate of working memory systems using positron emission tomography (PET),[2] functional magnetic resonance image (fMRI),[3,4] and electrophysiological techniques.[5,6]

The emphasis in many imaging studies has been on identifying the cortical loci of the proposed functional (verbal and visuospatial) subsystems of working memory.[2,6] A commonly used task in these studies has been the N-back task.[2,4,5] In this task, the participant is shown a series of items (e.g., letters, words, location markers) and is asked to decide, on each presentation, whether a given property of the current item matches the same property of the item N presentations back. If N = 0, each new item is matched against the very first item in the series. If N

= 1, each new item is matched against the immediately preceding item, and if N = 2, the new item is matched against the item presented just before the preceding item, and so on.

Researchers' preference for N-back tasks is most likely based on the assumption that it taps into processes involved in manipulating as well as maintaining information in WM.[3,7] It has been suggested, however, that the literature in fact lacks a thorough task analysis of the N-back paradigm, and that this leaves room for seemingly reasonable assumptions that may prove to be unsustainable under closer inspection. For example, the task analysis carried out by Meegan et al. (2004) cast doubt on the commonly held assumption that spatial and verbal N-back tasks actually tap into spatial and verbal WM processes, respectively. Based on behavioral studies using letter and position N-back tasks, these authors concluded that irrespective of the actual stimulus material or task demands, N-back task performance always involves both spatial and verbal processing.[8] This conclusion fits with neuroimaging results[2] showing that in an N-back task where letters were presented at different positions, activity in both cerebral hemispheres was obtained both under verbal task instructions (matching letter identity) and under spatial task instructions (matching letter position). However, because activity was lateralized slightly to the left under verbal instructions, and slightly to the right – at least in some areas – under spatial instructions, these authors concluded that verbal and spatial WM are in fact mediated by different neural substrates.

Identifying the neural correlates of WM in the N-back task will become easier and more reliable if our assumptions about the cognitive operations required for manipulating information in this task are clearly delineated and tested. In the present study, we first propose such a step-by-step analysis of the subprocesses involved in the N-back task. We then use the excellent temporal resolution of event-related potentials (ERPs) to isolate and compare the (broadly localized) physiological correlates of these subprocesses under verbal versus spatial task instructions.

Goals of the present study

The present study aimed to provide a fine-grained analysis of the electrophysiological correlates of verbal versus spatial WM processes. To this purpose, we compared ERPs elicited under spatial and verbal task instructions in 0-, 1- and 2-back conditions. It has to be noted that any systematic differences in perceptual processing across the spatial and verbal versions of the task would cause corresponding differences in early ERP components, and might also produce corresponding follow-up differences in later components.[9] Therefore, we eliminated such perceptual differences by employing identical stimulus displays in both tasks, changing only the task instructions. Stimuli were drawn from lists of 20 words and were presented one at a time at one of eight different screen locations. Under *spatial task instructions,* participants were asked to match the screen location of items. Under *verbal task instructions,* they were asked to match word identity.

We examined the general pattern of ERP effects associated with the spatial and verbal tasks. We expected to replicate earlier findings such as an increased posterior P3 component for infrequent 'match' relative to frequent 'non-match' trials,[10] and reduced P3 amplitude under higher WM load conditions.[11] The question of interest was how task instructions (spatial versus verbal) would alter the processing of identical stimulus displays. Specifically, it was assumed that if task instructions elicited domain-specific processing, then the verbal task should be accompanied by neural activity predominantly in the left hemisphere, whereas the spatial task should be accompanied by neural activity predominantly in the right hemisphere.[12-14]

METHODS

Participants

Sixty paid volunteers (35 women), ranging in age from 18 to 40 years (mean = 21) participated in the experiment. According to self-report, all had normal or corrected-to-normal vision, and all except six participants were right-handed.

Stimuli and apparatus

Stimuli were 20 words (either 20 positive or 20 negative adjectives) with similar frequency and length. Words were presented in white on black on a 17" computer monitor, at one out of eight circularly arranged positions 4° from the screen centre. Words had a height of approximately 0.8° visual angle, and the width ranged from 3.2° to 6.4° (mean = 5°).

Stimulus presentation and data acquisition were managed by C-programs running under MS-DOS. Behavioral and electrophysiological data were saved on the hard disk.

Procedure

Participants were seated in an armchair in front of a computer screen at a distance of approximately 60 cm. They were told to keep a comfortable posture, and to avoid eye movements and eye blinks during the experimental trials. Participants completed the first half of the experiment, comprising six blocks of n-back tasks, followed by a break, during which they were encouraged to leave the experiment room. They then completed the second half of the experiment. Each half consisted of two 0-back blocks, two 1-back blocks, and two 2-back blocks in sequence. In the first half, each pair of blocks was preceded by a corresponding practice block, to familiarize participants with the changing task requirements. In the second half, no practice blocks were administered.

Experimental blocks consisted of 64 trials (20 match trials and 44 non-match trials). Each trial began with the presentation of a fixation cross in the centre of a screen for 350 ms, followed by 350 ms of a blank screen. Then a stimulus word was shown for 500 ms at one of the eight predefined screen locations. This was followed by another blank screen for 1500 ms (Fig.1). The identity and location of each stimulus were determined pseudo-randomly, to achieve an approximately even distribution of match trials and an approximately equal distribution of identities and locations. Practice blocks were constructed in the same way, but contained only 20 trials and provided additional feedback (the words "correct" or "wrong" presented in the centre of the screen) immediately after the participant's response. Data from practice blocks were not saved.

In the 0-back task, participants indicated whether or not each stimulus matched the first one of the block. For the more demanding levels of the N-back task, participants had to match the current stimulus with the previous one (1-back task), or with the stimulus before the previous one (2-back task). Participants pressed a "yes" key for a match and a "no" key or a non-match. Keys were "\" and "/" keys of a computer keyboard, which had to be pressed with the left and right index finger, respectively. Participants were asked to respond as quickly and accurately as possible, and assignment of keys to the "yes" and "no" responses was counterbalanced across participants.

Participants were randomly assigned to either the verbal task or the spatial task, with equal numbers of participants in each group. In the verbal task, participants had to match the identity of the stimulus words. In the spatial task, they had to match their location. Note that verbal and spatial versions of the experiment differed only with respect to the instruc-
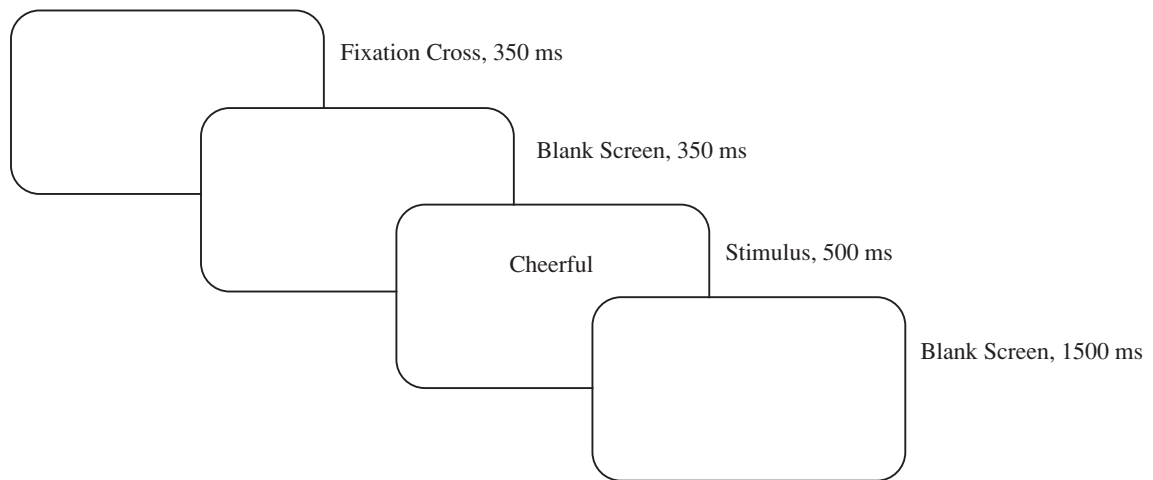
Fig. 1 Example of an experimental trial.

tion given to the participants, and were identical in all other respects. Within each group, half of the participants received positive words as stimuli, and half received negative words. However, this mood manipulation was of no relevance for the present research question and did not produce any systematic effects on the ERP components of interest. Therefore, it will not be discussed any further in the present study.

Electrophysiological recording and data processing

Using a BioSemi Active-Two amplifier system (BioSemi B.V., Amsterdam, Netherlands), continuous electroencephalogram (EEG) recordings were made with Ag/AgCl electrodes, mounted on a nylon cap, from 32 locations in the international 10-20 system (midline: FZ, CZ, PZ, OZ; left: Fp1, AF3, F7, F3, FC1, FC5, T7, C3, CP1, CP5, P7, P3, PO3, O1; and corresponding sites on the right). The sampling rate was 256 Hz. EEG signals were filtered off-line using a 0.01-30 Hz band-pass filter, and were re-referenced to linked earlobes.

Data analysis was conducted using EEGLAB software.[15] The EEG was averaged off-line for epochs of 900 ms, starting 100 ms prior to the stimulus onset, and ending 800 ms afterwards. Trials containing saccadic eye movement, eye blinks, or other movement artifacts (indicated by amplitudes beyond 3 standard deviations (SDs) in a single channel, or beyond 1.5 SDs in all channels), and trials where participants gave an incorrect response were excluded from analysis. The EEG on correct-response trials was averaged for each condition separately, relative to a 100-ms pre-stimulus baseline. Thus for each participant, 6 ERP waveforms were constructed: One match ERP and one non-match ERP from each of the 0-, 1-, and 2-back tasks.

Data analysis

Eight participants were excluded because of artifact rejection, because they had less than 25 EEG trials remaining in one or more conditions, or because they produced error rates of more than 2.5 SDs above the group's mean. No other data trimming procedures were employed.

Response times (RTs) and error rates (ERs) were analyzed using a repeated-measures analysis of variance (ANOVA) with the between-subject Task (spatial, verbal) and the within-subject factors Stimulus

(match, non-match) and N-Back (0, 1, 2).

Based on visual inspections of the grand mean waveforms (collapsed across N-Back conditions), four latency windows were selected for analysis: An early (150-250 ms) positive-negative shift in posterior areas (non-midline: P3/4, P7/8, O1/2, PO3/4; midline: Pz, Oz), further referred to as an early posterior complex (EPC); a positive peak between 200-300 ms in anterior areas (non-midline: FP1/2, AF3/4, F7/8, FC1/2, FC5/6; midline: Fz, Cz), further referred to as P2a; a negative-going shift at 300-400 ms in anterior areas (non-midline: FC5/6, F7/8, FC1/2, AF3/4, FP1/2; midline: Fz, Cz), further referred to as N2a; and a P3 component at 300-500 ms in central-posterior areas (non-midline: FC1, FC5, C3, T7, CP1, CP5, P3, P7, O1, PO3, and corresponding contralateral channels; midline: Cz, Pz, and Oz). ERP component amplitudes, which were defined as mean amplitudes within these time windows, were analyzed separately using a repeated-measures ANOVA with the between-subject factors and Task and the within-subject factors Stimulus and N-Back (0 / 1 / 2), and with the additional factor hemisphere (left / right) in the analysis of non-midline channels.

An α-level of .05 was applied for all statistical analyses. Greenhouse-Geisser corrections were applied and corrected p-values were reported where appropriate (Greenhouse-Geisser epsilon indicated in the Results section as ε).

RESULTS

Behavioral data

Behavioral results are presented in Fig. 2.

Response time (RT)

The RT to non-target stimuli increased in spatial tasks and decreased in verbal tasks in comparison with that to target stimuli, which was about the same in both tasks, as evidenced by a significant Stimulus x Task interaction, $F (1, 50) = 8.73$, $p = .005$. The RT increased with increasing memory loads, as evidenced by a significant N-Back effect, $F (1.3, 65.7) = 67.06$, $p < .001$. In 0- and 1-back tasks, the RT to non-target stimuli was shorter than that to target ones in 0-back tasks, almost the same in 1-back tasks, and longer in 2-back tasks, as evidenced by a significant Stimulus x N-Back effect, $F (1.8, 87.3) = 16.9$, $p < .001$. Other main effects or interactions were
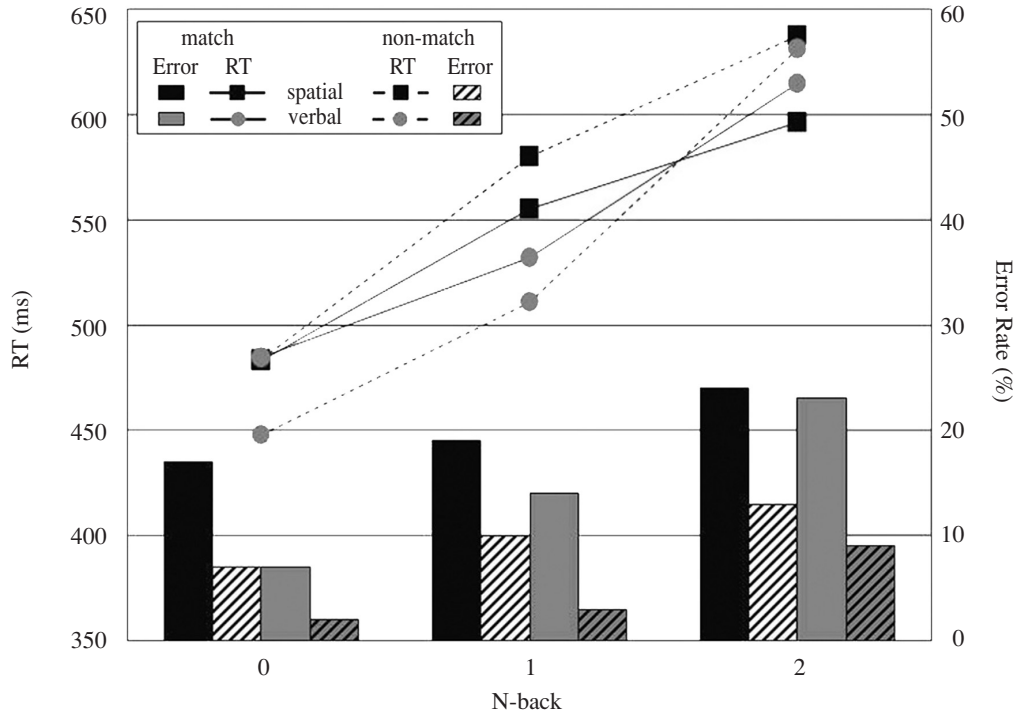
Fig. 2 Reaction times (lines) and error rates (bars) in 0-, 1-, and 2-back conditions, separately for spatial (thin lines) and verbal (thick lines) tasks, and separately for match trials (solid lines) and non-match trials (dashed lines).

non-significant for the RT, all $F < 2.18$, all $p > .138$.

Error rates (ERs)

ERs were higher in target stimuli than in non-target ones, as evidenced by a significant Stimulus effect, $F (1, 50) = 114.03$, $p < .001$, higher in spatial tasks than in verbal tasks, as evidenced by a significant Task effect, $F (1,50) = 8.02$, $p = .007$, and increased pertaining to memory loads, as evidenced by a significant N-Back effect, $F (1.6, 81.4) = 34.15$, $p < .001$. Other main effects or interactions were non-significant for the ERs, all $F < 3.13$, all $p > .059$.

Electrophysiological data: overall ERPs

The grand mean ERP waveforms, collapsed across the N-back factor, are presented in Fig. 3.

A main effect of Task ($p = .010$) was obtained only for the EPC: During this early latency window, ERPs elicited in the verbal task were generally more negative than ERPs elicited in the spatial task. In contrast, a main effect of Stimulus – with ERPs elicited by matching stimuli being more positive than ERPs elicited by non-matching stimuli – occurred

only in the three subsequent latency windows. This match effect was further modified by a Stimulus x Task interaction, which was found to change with time: In the earlier P2a latency window, the match effect ($p = .014$) was larger in the spatial than in the verbal task, particularly at midline and left-hemisphere sites. However, within the later two latency windows (N2, $p = .135$; and P3, $p = .006$) this relationship reversed, and the match effect became larger in the verbal tasks, again particularly in the left hemisphere (at midlines sites, the Stimulus x Task interaction was significant only in the P3 latency range, $p < 0.001$). Finally, the match effect was found to increase with increasing N in the N2 latency window, but only for the verbal, not for the spatial task, as evidenced by a significant N-Back x Stimulus x Task interaction ($p = 0.006$).

A main effect of N-Back – with ERP amplitudes becoming increasingly positive with increasing N – was evident at midline sites in all three earlier latency windows (EPC, $p = .003$; P2a, $p = .005$; and N2, $p = .017$), and at lateral sites during the EPC- ($p = .023$) and N2- ($p = .021$) latency windows. Initially
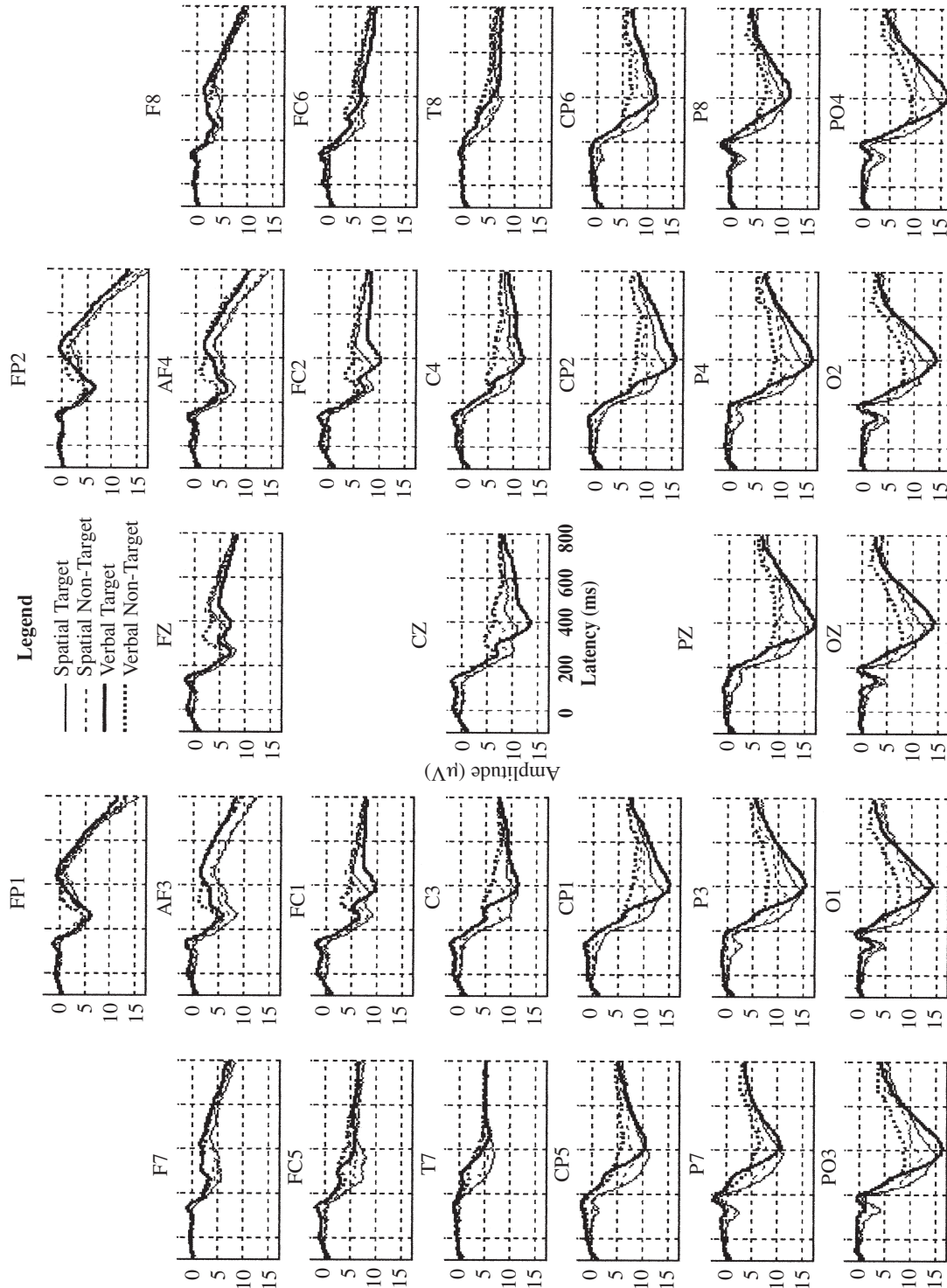
Fig. 3 Grand mean ERP waveforms, collapsed across the N-back factor, elicited during spatial (thin lines) and verbal (thick lines) tasks. ERPs elicited by matching items are indicated by solid lines, and ERPs elicited by non-matching items are indicated by dashed lines.

(i.e., within the EPC latency window), this effect was lateralized to the right hemisphere. However, an N-Back x Task interaction in the P3 latency range ($p = .045$) was due to the fact that within this time window, the N-Back effect reversed its direction (i.e., amplitudes were less positive for higher values of N) under verbal task instructions. Furthermore, a three-way interaction of N-Back x Task x Hemisphere was observed in the P2a ($p = .11$) and N2 ($p = .31$) latency windows, as the N-Back effect was particularly pronounced in the right hemisphere under spatial task instructions, but in the left hemisphere under verbal task instructions. Finally, N-Back effects were larger for matching than for non-matching stimuli (N-Back x Stimulus interaction), and this difference was significant in the N2 latency range ($p = .031$).

DISCUSSION

The present experiment investigated the electrophysiological correlates of verbal and spatial WM in the N-back task with varying information processing loads. We will first discuss the overall ERP effects of task instruction and memory load. Subsequently, we will consider the ERP evidence for distinct subprocesses in the N-back task and their possible differences depending on task instructions.

Electrophysiological correlates of verbal and spatial WM

The only main effect of task instructions was obtained in the EPC-latency range (150 – 250 ms after stimulus onset) at posterior electrode sites, where mean ERP amplitudes were more negative under verbal than under spatial instructions. There were no main effects of task instruction in any of the subsequent time windows, nor was there any evidence for differential lateralization of verbal and spatial tasks. These results contrast with earlier findings of task-specific lateralization even when stimulus material is held constant,[2] and appear to be more in line with the assumption that WM is a unitary mechanism which is not subdivided into modality-specific subsystems.[16]

However, this picture was complicated by the analysis of matching versus non-matching stimuli. As expected, infrequent matching stimuli elicited more positive-going ERPs than frequent non-match-

ing stimuli from approximately 200 ms post-stimulus onwards at anterior and particularly at centro-parietal sites, in line with results from studies using the oddball paradigm.[17] Unexpectedly, though, this 'matching' effect was initially larger under spatial task instructions at anterior sites in the P2a-latency window (in fact, it was virtually absent under verbal task instructions in this latency window). Subsequently, it became larger under verbal than under spatial task instructions both at anterior and at central-posterior sites in the N2- and P3-latency windows. This latter result seems to correspond to the findings of McEvoy et al. (1998),[11] although these authors do not report statistically significant Stimulus Type x Task instruction effects in this latency range. Taken together, these results suggest systematic processing differences under verbal versus spatial task instructions.

Unlike McEvoy et al. (1998), who observed a left-lateralized P3 component, the present study did not find any general amplitude differences between hemispheres. Match effects, however, were found to be larger in the left than in the right hemisphere, both under spatial task instructions (P2a latency window at anterior sites) and under verbal task instructions (subsequent latency windows at anterior and central-posterior sites). Given that in the McEvoy et al. (1998) study, participants always responded with the right hand, whereas in the present experiment, match- and nonmatch-responses were assigned to left- and right-hand responses counterbalanced across participants, one might argue that the former effect is influenced by lateralized motor processes, whereas the present effect indicates that the left hemisphere was more strongly involved in the match/non-match decision process than the right hemisphere. This might be taken as evidence that matching is a modality-unspecific WM process mediated by left-hemispheric structures. However, this is not the only possible interpretation. In particular, it seems more likely that the left-lateralized match effect indicates that participants employed verbal encoding strategies both under verbal and spatial task instructions. It is well known from the so-called Stroop effect that words will be read and understood even if such semantic processing is not only not helpful, but in fact interferes with task performance.[18] Therefore, it seems safe to assume that participants in the spatial task conditions did process

the stimulus words semantically – thereby engaging left-hemispheric structures – even though word meaning was not task relevant. Additionally, the possibility can not be ruled out that under spatial task instructions, participants encoded the stimulus location verbally ("top," "bottom," "top left," etc.), which also would favor left-hemisphere processing.

Of particular interest is the finding that in the N2 latency window, the anterior match effect increased with increasing N under verbal, but not under spatial task instructions. As can be seen from Fig. 4, this was due to the fact that whereas in all other conditions, amplitudes increased with increasing N, N2-amplitudes triggered by verbal non-matching stimuli remained constant.

These data suggest that the rejection of a non-matching stimulus word under verbal task instructions is not a load-sensitive process, whereas rejecting a non-matching stimulus under spatial task instruction, and recognition of a matching stimulus under either instruction, are load-sensitive processes. However, the behavioral data (Fig. 2) do not confirm this conclusion, as RTs and error rates for verbal non-matching stimuli increased with increasing N at least as much as in the other conditions. Therefore, it has to be concluded that the anterior N2 reflects a specific sub-process which does not contribute directly to response times, and which is different in verbal and spatial tasks.

Overall, ERP amplitudes increased with increasing N in the earlier three latency windows (particularly for matching items), and decreased with

increasing N in the P3 latency window. The former effect was particularly pronounced in the right hemisphere under spatial task instructions, whereas it was of approximately equal size in both hemispheres under verbal task instructions. Furthermore, the reduced P3 amplitude with increasing N replicates previous findings,[11] with the notable exception that here, it was observed only under verbal, but not under spatial task instructions. This fits with the assumption that spatial features of the stimulus material are processes whether or not they are response relevant.[8]

Taken together, these results indicate systematic differences between WM processes concerning spatial and verbal aspects of identical stimulus displays. However, these differences are considerably more subtle than general lateralization of activity to the right under spatial and to the left under verbal task instructions.[12-14]

REFERENCES

1. Baddeley AD, Hitch GJ. Developments in the concept of working memory. Neuropsychology 1994;8:485-93.
2. Smith EE, Jonides J. Working memory: a view from neuroimaging. Cognit Psychol 1997;33:5-42.
3. Ragland JD, Turetsky BI, Gur RC, Gunning-Dixon F, Turner T, Schroeder L, Chan R, Gur RE. Working memory for complex figures: an fMRI comparison of letter and fractal n-back tasks. Neuropsychology 2002;16:370-9.
4. Jansma JM, Ramsey NF, Coppola R, Kahn RS. Specific versus nonspecific brain activity in a parametric n-back task. Neuroimage 2000;12:688-97.
5. Gevins A, Cutillo B. Spatiotemporal dynamics of component processes in human working memory. Electroencephalogr Clin Neurophysiol 1993;87:128-43.
6. Ruchkin DS, Johnson R, Grafman J, Canoune H, Ritter W. Distinctions and similarities among working memory processes: an event-related potential study. Brain Res Cogn Brain Res 1992;1:53-66.
7. Meegan DV, Purc-Stephenson R, Honsberger MJM, Topan M. Task analysis complements neuroimaging: an example from working memory research. Neuroimage 2004;21:1026-36.
8. Meegan DV, Honsberger MJM. Spatial information is processed even when it is task-irrelevant: Implications for neuroimaging task design. Neuroimage 2005;25:1043-55.
9. Schendan HE, Ganis G, Kutas M. Neurophysiological evidence for visual perceptual categorization of words and faces within 150 ms. Psychophysiology 1998;35:240-51.
10. Katayama J, Polich J. Auditory and visual P300 topogra-
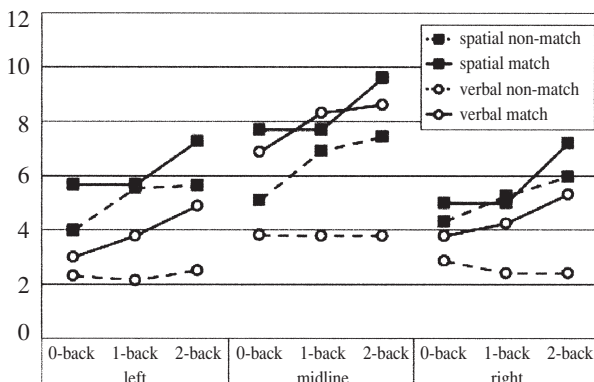
Fig. 4 Mean N2 amplitudes at left, midline, and right anterior electrodes, plotted separately for spatial and verbal match and non-match stimuli, and separately for 0-back, 1-back, and 2-back tasks.

phy from a three-stimulus paradigm. Clin Neurophysiol 1999;110:463-8.

11. McEvoy LK, Smith ME, Gevins A. Dynamic cortical networks of verbal and spatial working memory: Effects of memory load and task practice. Cereb Cortex 1998;8:563-74.

12. Beauregard M, Chertkow H, Bub D, Murtha S, Dixon R, Evans A. The neural substrate for concrete, abstract, and emotional word lexica: A positron emission tomography study. J Cogn Neurosci 1997;9:441-61.

13. Deutsch G, Bourbon WT, Papanicolaou AC, Eisenberg HM. Visuospatial tasks compared via activation of regional cerebral blood flow. Neuropsychologia 1988;26:445-52.

14. Petersen SE, Fox PT, Snyder AZ, Raichle ME. Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. Science 1990;249:1041-4.

15. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 2004;134:9-21.

16. Ross P, Segalowitz SJ. An EEG coherence test of the frontal dorsal versus ventral hypothesis in n-back working memory. Brain Cogn 2000;43:375-9.

17. Bennington JY, Polich J. Comparison of P300 from passive and active tasks for auditory and visual stimuli. Int J Psychophysiol 1999;34:171-7.

18. MacLeod CM. Half a century of research on the Stroop effect: an integrative review. Psychol Bull 1991;109:163-203.

文字與空間工作記憶的不同：使用事件關連電位的研究

陳勇年[1,2] Suvobrata Mitra[2]

背 景： 失智症主要表現爲工作記憶 (working memory) 的缺失。N-back task 在工作記憶研究上常用。藉由控制 N 這個參數，實驗心理學者可以用來研究不同工作負擔 (loading) 下心理機制的差別。同時N-back task 也很容易便能改裝以文字或圖形介面測試。

方 法： 我們使用文字或圖形介面供不同的受試者辨認，並給予不同的工作負擔，同時記錄其事件關聯電位(ERPs)。

結 果： 實驗結果顯示文字及空間處理上左右腦的差異並不明顯，其區別主要在不同的組成波 (component)，文字增強 P3，空間處理則增強 P2a。目標／非目標的辨別則明顯爲左腦主導，僅左腦的 P3 會被引發。N2 波會被語言的非目標所引發，而且波幅固定。

結 論： 文字和圖形介面下，以及不同的工作負擔下的 N-back task，並非是一般所認爲的左腦語言／右腦圖像差異，而是有更加細微的區別表現。

(長庚醫誌 2009;32:380-9)

關鍵詞： N-back 測驗，空間，語言，工作記憶，事件關聯電位