JOURNAL OF CLINICAL AND EXPERIMENTAL NEUROPSYCHOLOGY
2011, 33 (7), 721–734

Psychology Press
Taylor & Francis Group

Computerized analysis of error patterns in digit span recall

David L. Woods[1,2,3,4], T. J. Herron[1], E. W. Yund[1], R. F. Hink[1], M. M. Kishiyama[1], and Bruce Reed[2,5]

[1]Human Cognitive Neurophysiology Laboratory, Veterans Affairs Northern California Health Care System, Martinez, CA, USA
[2]Department of Neurology, University of California at Davis, Davis, CA, USA
[3]Center for Neurosciences, University of California at Davis, Davis, CA, USA
[4]Center for Mind and Brain, University of California at Davis, Davis, CA, USA
[5]Alzheimer's Disease Center, Veterans Affairs Northern California Health Care System, Martinez, CA,USA

We analyzed error patterns during digit span (DS) testing in four experiments. In Experiment 1, error patterns analyzed from a community sample of 427 subjects revealed strong primacy and recency effects. Subjects with shorter DSs showed an increased incidence of transposition errors in comparison with other error types and a greater incidence of multiple errors on incorrect trials. Experiment 2 investigated 46 young subjects in three test sessions. The results replicated those of Experiment 1 and demonstrated that error patterns of individual subjects were consistent across repeated test administrations. Experiment 3 investigated 40 subjects from Experiment 2 who feigned symptoms of traumatic brain injury (TBI) with 80% of malingering subjects producing digit spans in the abnormal range. A digit span malingering index (DSMI) was developed to detect atypical error patterns in malingering subjects. Overall, 59% of malingering subjects with abnormal digit spans showed DSMIs in the abnormal range and DSMI values correlated significantly with the magnitude of malingering. Experiment 4 compared 29 patients with TBI with a new group of 38 control subjects. The TBI group showed significant reductions in digit span. Overall, 32% of the TBI patients showed DS abnormalities and 11% showed abnormal DSMIs. Computerized error-pattern analysis improves the sensitivity of DS assessment and can assist in the detection of malingering.

Keywords: Verbal; Memory; Short-term; Working; Malingering; Wechsler Adult Intelligence Scale; Adaptive; Digit span; Traumatic brain injury; Concussion.

INTRODUCTION

Digit span (DS) measures are among the oldest and most widely used neuropsychological assessments of verbal working memory (Richardson, 2007). In DS testing, all digits must be reported in the correct order for the trial to be scored as correct. Although measurements are typically restricted to tallying the total number of correct trials or the maximal span (Wechsler, 1997a, 1997b), recent reports suggest that more detailed error analysis may enhance the clinical sensitivity of DS testing (Kramer et al., 2003; Lamar et al., 2007). The characterization of DS errors requires both an analysis of the types of errors that occur and their serial positions within the digit list. Errors can be categorized into five types. Omission errors occur when a subject forgets to report a digit, but reports the remaining digits in correct order, for example when the string "1–2–3–4" is reported as "1–2–4." Substitution errors occur when a

This research was supported by the VA Research Service and Award R01ES014038 from the National Institute of Environmental Health Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. We would also like to acknowledge the assistance of the neuropsychology data collection staff: Nerene Lynskey, Kataraina George, and John McKeogh in New Zealand and Ben Edwards, Zoe Doss, David Doss, and Oren Poliva in Martinez.

Address correspondence to David L. Woods, Neurology Service (127E), VANCHCS, 150 Muir Rd., Martinez, CA 95553, USA. (E-mail: dlwoods@ucdavis.edu).

http://www.psypress.com/jcen
DOI: 10.1080/13803395.2010.550602

subject substitutes one digit for another, (e.g., "1–2–3–4" is reported as "1–2–5–4"). Intrusion errors occur when a subject reports extra digits in the digit string (e.g., "1–2–3–4" is reported as "1–2–3–4–5"). Omissions, substitutions, and intrusions collectively constitute *item errors*—that is, the digit string presented and the string reported contain different digits. There are also two types of *order error*, where the list presented and list reported contain identical digits in different orders. Transposition errors occur when a subject transposes a pair of digits. For example, when the string "1–2–3–4" is reported as "1–3–2–4." Permutation errors occur when digit order is incorrect in a manner that cannot be explained by transpositions, for example when the string "1–2–3–4" is reported as "1–4–2–3."

Serial-position effects in list recall are also well established for a variety of verbal list learning tests: Initial and final items in the list are recalled more accurately than items occurring in the middle of the list, due to so-called primacy and recency effects (Henson, 1998; Page & Norris, 1998). Primacy and recency effects have also been previously reported in DS tests (Bunting, Cowan, & Colflesh, 2008; Gillam, Cowan, & Day, 1995), and recency effects are reportedly reduced in children with specific language impairment (Gillam, Cowan, & Marler, 1998).

The current manuscript describes DS errors that occurred in four experiments that incorporated adaptive computerized delivery of randomized digit sequences and improved procedures for quantifying mean digit span (MS; Woods et al., 2010). In Experiment 1, DS errors were analyzed from a single test performed on a large community sample of 427 subjects. The goal of Experiment 1 was to characterize the serial-position functions of DS recall, categorize the types of errors that occurred, and evaluate correlations between error types and MS. In Experiment 2, error patterns were analyzed in 47 young volunteers who underwent three separate test sessions with the goal of determining whether error patterns were consistent across test sessions in individual subjects. In Experiment 3, a subset of subjects from Experiment 2 ($N = 40$) underwent an additional DS test while simulating impairments produced by traumatic brain injury (TBI). The goal of Experiment 3 was to evaluate the effects of malingering on digit span performance (Axelrod, Fichtenberg, Millis, & Wertheimer, 2006; Heinly, Greve, Bianchini, Love, & Brennan, 2005) and to determine whether the patterns of errors made by malingering subjects could be distinguished from the error patterns observed in subjects performing with full effort. Experiment 4 analyzed DS test results from 29 patients diagnosed with TBI in comparison with a demographically matched group of 38 control subjects.

EXPERIMENT 1

Method

Subjects

A 10-list DS test was performed by 763 community volunteers in Rotorua, New Zealand, who participated in a study investigating the effects of hydrogen sulfide exposure on health and mental performance. Due to incomplete scoring by one examiner, error analysis was performed on a subset of these subjects ($N = 427$). Subjects in this subset ranged in age from 18 to 65 years (mean age = 46.9 years) with an average of 12.5 years of education. The large age range also enabled us to examine age-related differences in error patterns including the incidence of transposition errors that reportedly increases in older subjects (Ardila, 2007; Jurden, Laipple, & Jones, 1993; Kemtes & Allen, 2008; Kessels, van den Berg, Ruis, & Brands, 2008). All subjects signed written consent forms approved by the institutional review boards (IRBs) in Rotorua and at the Veterans Affairs Northern California Health Care System (VANCHCS).

Apparatus and stimuli

Forward and backward digit span testing was performed midway through a brief 30-min computerized assessment battery that included six tests from the California Cognitive Assessment Battery (CCAB). Testing was performed in a quiet testing room using a standard PC controlled by Presentation software (Pebler, 2011). Responses were recorded by the experimenter using a PC-gaming keyboard. First, the forward digit span testing procedure was explained to the subject. Then, spoken digits (1–9) that had been digitally recorded (44.1 kHz, 16 bits) and normalized in mean intensity (70 dB SPL) were delivered at the rate of 1/s through headphones at an intensity of 70 dB SPL. Digits were randomly sampled without replacement up to list lengths of 9 digits, with the constraint that successive digit pairs could not occur in regular ascending or descending sequence (e.g., 1–2–3), or in ascending or descending sequences of odd or even numbers (e.g., 2–4–6).

A warning cue followed the final digit at an interval of 1.0 s, and the subjects then repeated the digit string in forward order. The digit sequence was displayed on the examiner's monitor prior to list presentation. Responses were transcribed by the examiner using the computer keyboard. The experiment log file included the identity and timing of each digit presented and the identity and timing of each response as transcribed by the examiner.

The subject received 10 trials with list lengths adaptively adjusted to reflect subject performance. Forward testing began at a list length of five digits. The list length on successive trials was controlled by a 1:2 staircase: A single correct response increased the length of the next list by one digit, while two incorrect responses reduced list length by one digit. Following forward span testing, the subject received 10 trials of backward span testing with the digit sequence reported in backward order. Backward span testing began at a list length of four digits, but was otherwise identical to forward span testing.

Error scoring

Two stages of error analysis were used. First, all common digits in the stimulus and response were extracted in their relative orders, and the minimum number of transpositions and permutations needed to transform the stimulus digit ordering into the response digit ordering were determined (Cameron, 1999). Second, after correcting for permutations and transpositions, we applied the Levenshtein distance algorithm (Gusfield, 1997) to find the minimum number of intrusions, omissions, and substitutions to match the order-corrected response to the stimulus. This two-stage algorithm produces a unique set of errors for every response that does not contain improperly repeated digits. For incorrect responses that contained repeats, the algorithm selected the repeat closest to its stimulus list position for use in the first stage of error analysis.

Item errors (omissions, substitutions, and intrusions) and order errors (transpositions and permutations) were tallied on each incorrect trial using computerized scoring metrics. Transpositions were further subdivided into first-order (between adjacent digits), second-order, and third-order transpositions (e.g., “1234” reported as “4231”). We also quantified permutations, where the digit sequences were reported in incorrect order that could not be corrected by a single digit transposition. In order to examine serial-position effects, total errors and errors of each type were tallied at each serial position for digit strings of different lengths.

An error severity score (ESS) was also calculated on each incorrect trial using the equation by summing the total number of errors after multiplying the number of permutations by 1.5 and dividing by the total number of digits presented. Thus, the ESS reflected the percentage of digits incorrectly reported on each incorrect trial. Table 1 shows an example of scoring a DS forward test sequence for one subject. This subject had 13 total errors on 7 trials missed and produced an average ESS of 23.1% on incorrect trials.

Results

Error analysis

We estimated DS using a mean span (MS) metric based on psychophysical procedures. In comparison with traditional DS measurements, MS scoring improves test–retest reliability, provides more reliable estimates of forward–backward span differences, and enhances correlations with demographic variables and other neuropsychological test scores (Woods et al., 2010). The average MS for forward span was 6.52 ($SD = 1.01$), and the average MS for backward span was 4.95 ($SD = 1.11$). On average, subjects missed 17.5% of digits on forward span trials and 19.9% of digits on backward span trials. Subjects who made more errors on incorrect trials had shorter spans, producing a negative correlations between ESS scores and MS for both forward ($r = -.34$, $p < .0001$) and backward ($r = -.50$, $p < .0001$) span.

Errors were classified as either item errors, in which digits were deleted, added, or substituted, or order errors, where digit order was transposed or permuted. There were more item than order errors in both forward (7.29 vs. 4.49) and backward (6.48 vs. 3.98) span tests. Forward MS was not correlated with either overall item or order errors, but it showed negative correlations with first-order transpositions ($r = -.26$, $p < .0001$) and total transposition errors ($r = -.21$, $p < .0001$). For backward span, MS showed a surprising positive correlation with omissions ($r = .33$, $p < .0001$). This was due to the fact that subjects often stopped reporting digits midway through long backward span lists. As a result, subjects with longer backward spans omitted more digits than those with shorter spans. Backward DS was negatively correlated with order errors ($r = -.22$, $p < .0001$) and particularly

TABLE 1
Test results from a single test for one subject on digit span forward testing

Trial	*Length*	*Presented*	*Response*	*Errors*	*OE*	*IE*	*SE*	*TE1*	*TE2*	*PE3*
1	5	62748	62748	0	0	0	0	0	0	0
2	6	478**6**31	487381	3	1	1	0	1	0	0
3	6	953748	953748	0	0	0	0	0	0	0
4	7	6359**1**72	635792X	2	1	0	0	1	0	0
5	7	12**8**9653	126913X	3	1	0	1	1	0	0
6	6	976285	972865	1	0	0	0	0	0	1
7	6	692784	697824	1	0	0	0	0	0	1
8	5	53617	536173	1	0	1	0	0	0	0
9	5	19348	13943	2	0	0	1	1	0	0
10	4	2814	2814	0	0	0	0	0	0	0

Note. Ten trials were presented, with list length (column 2) increasing after correct trials and decreasing after two successive incorrect trials at the same list length. The list presented is shown in column 3, the response in column 4. The number of errors is shown in column 5, with the error types indicated in columns 6–10 (OE = omission errors, IE = intrusion errors, SE = substitution errors, TE1 = first-order transposition errors, TE2 = second-order transposition errors, and PE3 = third-order permutation errors). For example, on Trial 5, the subject omitted the “8,” made a TE1 error reversing the digit pair “96,” and substituted a “1” for a “5.” Digits not reported are indicated with X, as in Trial 4. Scoring was performed automatically by computer from responses typed by the examiner.

with first-order transpositions ($r = -.37$, $p < .0001$) and total transposition errors ($r = -.38$, $p < .0001$).

Serial-position effects

Serial-position functions for forward lists of length 7 and backward lists of length 6 are shown in Figure 1. Both forward and backward spans showed clear primacy and recency effects, with greatest error probabilities observed in the penultimate two digits in the list. Omission errors increased for digits reported later in the string and hence occurred predominantly among the last digits presented in forward span testing and among the first digits presented (but last reported) during backward span. In contrast, transposition errors were most frequent for digits presented near midlist during both forward and backward span testing.

Figure 2 shows the percentage of errors of different types as a function of trial length. Transposition errors predominated in short lists and showed a decline in relative incidence in longer lists, while omission errors were infrequent in short lists but increased in relative incidence with increasing list length. Permutation errors also increased somewhat with list length, although the increase was less than the increase in the number of possible permutations (a factorial function).

Correlations with age, education, and other neuropsychological tests

Error patterns were little affected by age. When tested at the $p < .005$ level (for multiple comparisons), there were no significant correlations between age and error type for either forward span (FS) or backward span (BS). Education exerted a marginal effect, with the only significant correlation observed between years of education and BS ESSs ($r = -.21$, $p < .0001$).

However, ESSs correlated significantly with scores on other neuropsychological tests. For example, errors in the National Adult Reading Test (NART; O'Caroll & Gilleard, 1986) were positively correlated FS and BS ESSs ($r = .26$ and $r = .31$, $p < .0001$, for both comparisons). Similarly, scores in the Benton Visual Retention Test (BVRT; Benton, 1962) and the Hopkins Verbal Learning Test–Revised (HVLT–R) encoding and delayed recall scores (Shapiro, Benedict, Schretlen, & Brandt, 1999) were negatively correlated with BS ESSs ($r = -.21$ to $r = -.32$, $p < .0001$, for all comparisons). There were also trends indicating that an increased frequency of transposition errors (TEs) correlated with poor performance on tests of verbal memory. For example, NART errors correlated with BS total TEs ($r = .15$, $p < .01$), while HVLT total recall and delayed recall scores were

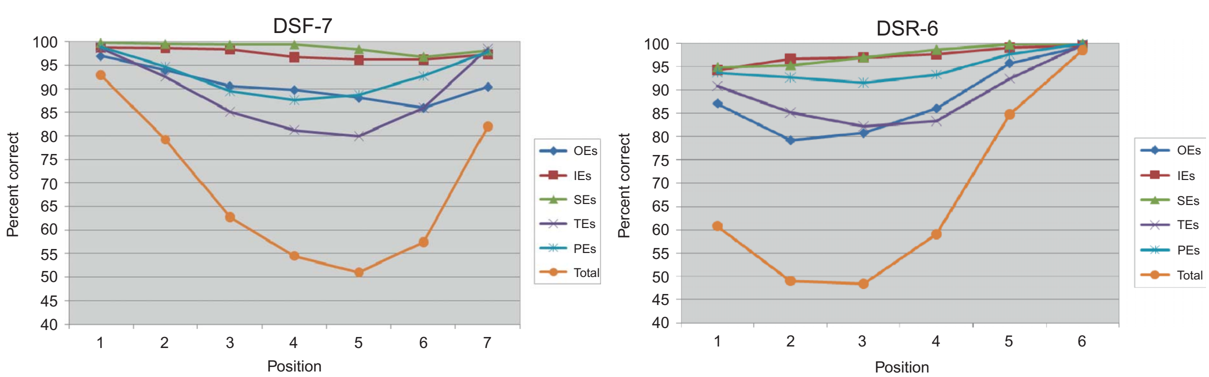

Figure 1. Experiment 1 serial-position functions for total errors and errors of different types for digit span forward (DSF) list length of 7 digits and digit span backward (DSR) list length of 6 digits. OEs = omission errors, IEs = intrusion errors, SEs = substitution errors, TEs = transposition errors, PEs = permutation errors, and Total = total errors. Note that in the digits backwards condition, Position 6 is the last presented, first to be recalled item. To view a color version of this figure, please see the online issue of the Journal.

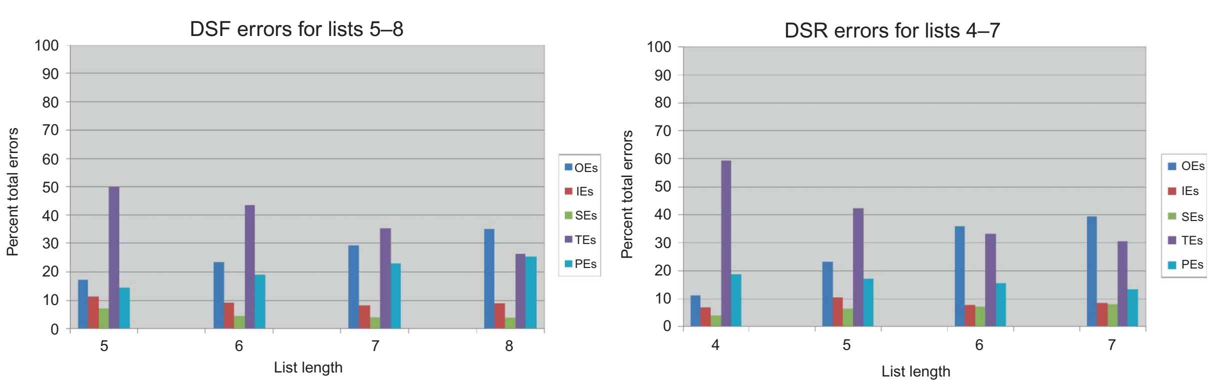

Figure 2. Experiment 1 percentage of total errors of different types as a function of error type for list lengths 5–8 for digit span forward (DSF) and 4–7 for digit span backward (DSR). OEs = omission errors, IEs = intrusion errors, SEs = substitution errors, TEs = transposition errors, PEs = permutation errors. To view a color version of this figure, please see the online issue of the Journal.

negatively correlated with BS total TEs ($r = -.15$ and $r = -.14$, $p < .01$).

Discussion

Correlations between error type and digit span

MS performance in forward span was negatively correlated with ESS and with first-order and total transpositions. In backward span, MS was negatively correlated with ESS and with first-order and total transpositions. Paradoxically, subjects with greater numbers of omission errors showed longer reverse digit spans. This was likely the consequence of the fact that subjects frequently abandoned backward digit report in midlist, producing greater numbers of omission errors for subjects with longer spans.

List length and serial-position effects

Primacy and recency effects were seen in both forward and backward span testing. However, the primacy and recency pattern depended on the interaction of two underlying types of error. Omission errors increased in relative frequency at longer list lengths and tended to occur late in digit report sequence. In contrast, order errors, particularly transpositions, decreased in relative frequency with increasing list length and tended to occur in the middle of digit lists. As a result, error composition changed with list length: Transposition errors predominated in short lists while omission errors predominated in long lists, due in part to the tendency of subjects to halt digit report for lists that exceeded their spans.

Verbal memory and error patterns

Even though digit list lengths were adapted to performance, subjects with greater ESSs on incorrect trials had shorter digit spans. This likely reflects the fact that an increase of one digit in the length of the digit string had greater impact on subjects with short spans. For example, when a subject with a MS of 4.0 digits is tested with a digit list of 5 digits, the list exceeds DS capacity by 25%. In contrast, when a subject with a MS of 8.0 digits is tested with a digit list of 9 digits, the list exceeds DS capacity by only 12.5%.

The pattern of errors also correlated with DS performance. In particular, first-order and total transpositions showed negative correlations with both forward and backward MS and predicted poor performance in the NART and HVLT–R. This suggests that transposition errors may be a particularly sensitive sign of short-term verbal memory problems, consistent with their increased incidence in certain clinical syndromes (Lamar et al., 2007).

EXPERIMENT 2

In Experiment 2, subjects underwent three successive DS test sessions to examine the consistency of error patterns over repeated testing. In addition, subjects in Experiment 2 were encouraged to fully report all digits in the digit list, guessing when they lacked confidence in their accuracy of their report. This modification was designed to permit a more accurate analysis of error types as a function of increasing list lengths than was possible in Experiment 1, where subjects frequently abandoned report when lists exceeded their spans.

Method

Subjects

Forty-seven subjects participated in Experiment 2 after giving written informed consent following IRB regulations of the VANCHCS. The subjects were mainly students and included 23 men and 24 women between the ages of 18 and 46 years (mean age = 26.5 years), with an average 14.8 years of completed education. One female subject, who used a mnemonic strategy and had a backward span of 13 digits, was excluded from the analysis.

Procedure

Forward and backward digit span testing was performed midway through the 1.5-hour-long version of the California Cognitive Assessment Battery (CCAB), which contained 17 computerized tests and three adaptive questionnaires.[1] In order to evaluate test–retest reliability, each subject underwent three test sessions at intervals ranging from 2 to 11 days. The methods were similar to those used in Experiment 1 except that 14 random digit lists were presented, and forward digit span testing began at 3 digits, while backward digit span began at 2 digits. Subjects were encouraged to report the same number of digits as in the test list, guessing if necessary.

Results

Error analysis

The average MS for forward span was 7.43 ($SD = 1.05$), and the average MS for backward span was 5.81 ($SD = 1.23$). ESS indicated that subjects missed an average of 12.5% of digits on incorrect forward span trials and

[1]The CCAB includes the following computerized tests and questionnaires: finger tapping, simple reaction time, symbol–digit, Stroop, digit span forward and backward, phonemic and semantic verbal fluency, card sorting, verbal list learning, spatial span, trail making, symmetry detection, design fluency, the Wechsler Test of Adult Reading (WTAR), visual feature conjunction, the Paced Auditory Serial Addition Task (PASAT), the Cognitive Failures Questionnaire (CFQ), the posttraumatic stress disorder (PTSD) symptoms checklist, and a traumatic brain injury (TBI) questionnaire.

15.2% of digits on incorrect backward span trials. As in Experiment 1, there were strong significant correlations between ESS and MS for both forward ($r = -.41$, $p < .005$) and backward span ($r = -.78$, $p < .0001$).

There were more item errors than order errors in both forward (7.14 vs. 5.95) and backward span (8.23 vs. 4.57). Correlation analysis revealed that both forward and backward MS showed a negative correlation with first-order transpositions ($r = -.38$, $p < .01$, and $r = -.58$, $p < .0001$, respectively). In contrast, MS showed no significant correlations with the most common type of error, omissions, in either forward or backward span.

Serial-position effects

Serial-position effects were similar to those observed in Experiment 1. Figure 3 shows the incidence of total errors and errors of different types at different serial positions in forward lists containing 7 digits and backward lists containing 6 digits. Errors showed strong primacy and recency effects for both forward and backward span. The percentage of digits correctly reported declined from the first to the penultimate digit and then increased for the final digit reported. The differences in accuracy as a function of digit position were substantial. For example, backward span error rates varied by 30-fold as a function of serial position: from 1.5% for the last digit presented and first digit reported to 45% for the penultimate digit. Similarly, forward span error rates increased 6-fold between the first and penultimate digit reported. As in Experiment 1, error patterns also changed as a function of list length (Figure 4). Transposition errors predominated at short list lengths, but showed a decline in relative frequency as list lengths increased. In contrast, omissions were infrequent in short lists, but increased with list length.

Consistency of error patterns

To examine the tendency of individual subjects to produce item or order errors, we examined the ratio of item errors to total errors for each subject. For forward span, subjects had an average item/total error ratio of .53 and showed significant positive correlations ($r \geq .30$, $p < .05$) across two of three pairwise comparisons of test sessions. For backward span, subjects had an average item/total error ratio of .62, with significant positive correlations observed across all three test sessions ($r \geq .30$, $p < .05$).

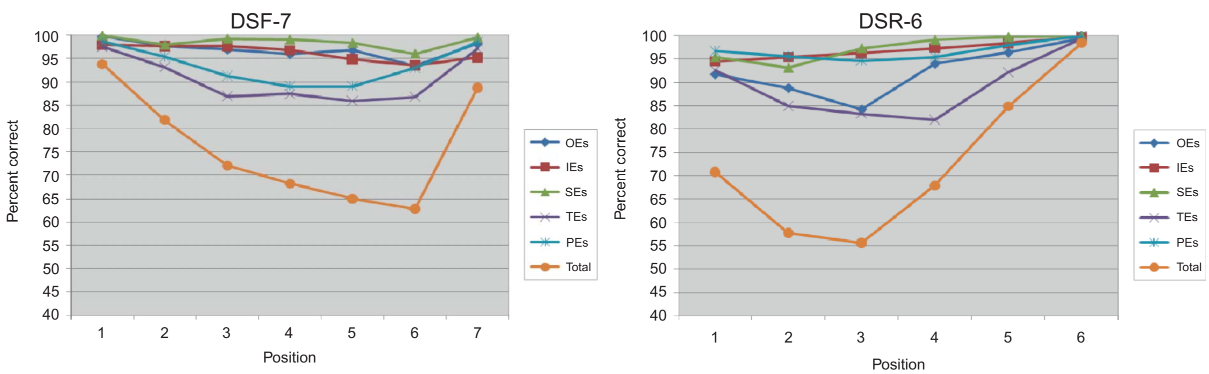

Figure 3. Experiment 2 serial-position functions for total errors and errors of different types for digit span forward (DSF) list length of 7 digits and digit span backward (DSR) list length of 6 digits. OEs = omission errors, IEs = intrusion errors, SEs = substitution errors, TEs = transposition errors, PEs = permutation errors. To view a color version of this figure, please see the online issue of the Journal.

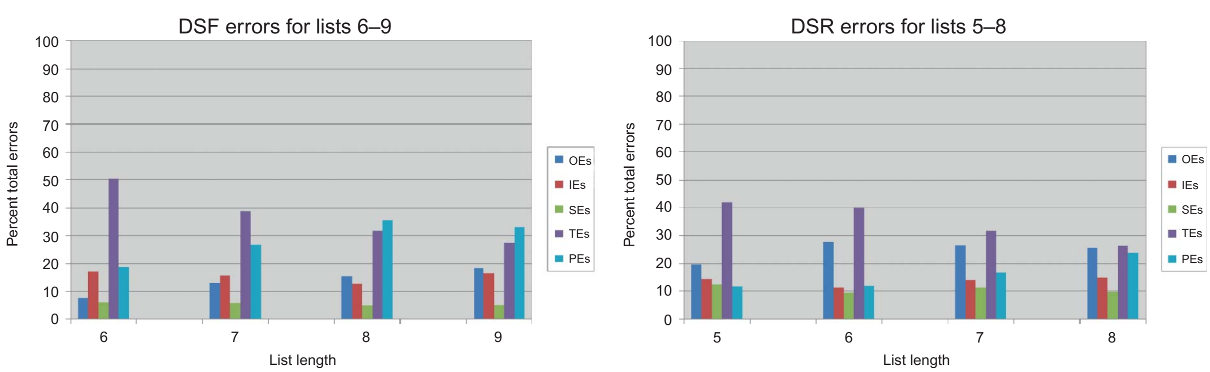

Figure 4. Percentage of total errors of different types as a function of error type for list lengths 6–9 for digit span forward (DSF) and 5–8 for digit span backward (DSR) in Experiment 2. OEs = omission errors, IEs = intrusion errors, SEs = substitution errors, TEs = transposition errors, PEs = permutation errors. To view a color version of this figure, please see the online issue of the Journal.

Discussion

Correlations between error type and digit span

As in Experiment 1, we found that the pattern of errors correlated with overall MS performance. Surprisingly, the incidence of the most frequent type of error, omissions, was not significantly correlated with overall digit span. In contrast, as in Experiment 1, the incidence of first-order transpositions showed significant negative correlations with MS.

Serial-position effects

We found primacy and recency effects for digit span that resembled the serial-position functions reported in other serial list learning tasks (Henson, 1998; Page & Norris, 1998). The poorest performance was found for the penultimate digit reported in forward span and in the digit prior to the penultimate digit in backward span. The serial-position functions cannot be accounted for by the interval between digit delivery and digit report. Since the average interval between the report of successive digits (0.91 s) was similar to the average interval of digit delivery (1 s/digit), the interval between digit presentation and report changed only slightly for forward span testing. Backward span showed near-perfect performance for the final digit presented and degraded performance for digits reported later. The near-perfect recall of the final digit in backward span likely reflected both recency effects and the minimal delay between digit delivery and digit report.

List length and types of error

As in Experiment 1, transposition errors predominated in short lists but their relative incidence diminished as list lengths increased. Thus, subjects who produce a high incidence of transposition errors would be expected to make more errors at short list lengths and hence have shorter spans. In contrast, the relative incidence of omission errors increased with lists of increasing length, accounting for the poor overall correlation between omission errors and MS. These results suggest that order information, while apparently fragile in short lists, did not deteriorate as rapidly with increasing list lengths as did item information. This pattern was observed in each test session for both forward and backward span.

Subjects showed some consistency in error patterns with significant correlations in the ratio of item/total errors found across test sessions. However, the correlations in error patterns were modest in magnitude, as would be expected because of the small number of incorrect trials produced in a typical test (mean of approximately 6 trials in FS and approximately 7 trials in BS).

EXPERIMENT 3

Neuropsychologists are often faced with the challenge of distinguishing malingering subjects from patients with impairments due to clinical disability. Because DS performance is generally well preserved in patients with brain dysfunction, including amnesic patients (Greiffenstein, Baker, & Gola, 1994; Iverson & Franzen, 1996; Iverson & Tulsky, 2003), unexpectedly low DS performance has itself been used as a potential index of malingering (Dean, Victor, Boone, Philpott, & Hess, 2009; Fisher & Rose, 2005; Greve et al., 2007; Heinly et al., 2005; Iverson & Franzen, 1994, 1996; Schwarz, Gfeller, & Oliveri, 2006; Shum, O'Gorman, & Alpar, 2004; Trueblood, 1994; Vagnini et al., 2006). Different metrics have also been proposed for the detection of malingering in DS tests, including the reliable digit span, the longest string of digits repeated without error over two trials (Babikian, Boone, Lu, & Arnold, 2006; Greiffenstein & Baker, 2008; Greiffenstein et al., 1994; Larrabee, 2003; Ruocco et al., 2008; Strauss et al., 2002), age-corrected scaled scores (Axelrod et al., 2006; Greve et al., 2007; Iverson & Tulsky, 2003; Whitney, Davis, Shepard, Bertram, & Adams, 2009), and DS performance relative to the Vocabulary subtest (V-DS) of the Wechsler Adult Intelligence Scale–Revised (WAIS–R) or the Wechsler Adult Intelligence Scale–Third Edition (WAIS–III; Dean et al., 2009; Iverson & Tulsky, 2003; Miller, Ryan, Carruthers, & Cluff, 2004; Mittenberg, Theroux-Fichera, Zielinski, & Heilbronner, 1995). However, none of these metrics analyze the pattern of errors in digit lists that are incorrectly recalled.

Malingering subjects often adopt a simple strategy (for example, reporting the first several digits in the list) which can be distinguished from the error patterns that are observed in subjects performing with full effort. Indeed, malingering subjects who attempt to duplicate normal error patterns are faced with several difficult challenges. First, they must produce errors with appropriate serial-position functions: Normal errors occur in the late-middle portion of digit strings, so that malingering subjects must be careful to avoid errors among initial and final digits. Second, malingering subjects should produce a mixture of item and order errors, with order errors predominating in short lists and item errors predominating in longer lists. Thus, to reproduce normal error patterns, a malingering subject would need to keep track of both the serial positions and the types of error made on each trial. Finally, a malingering subject would need to produce deficits of appropriate magnitude during forward and backward span testing, to produce shorter BS than FS scores. We therefore investigated whether a DS malinger index (DSMI) that incorporated error serial-position functions, error-type variability, and the difference between FS and BS scores could assist in the identification of subjects who were malingering.

Method

Subjects

Forty subjects from Experiment 2 participated in Experiment 3. These subjects included 20 women and 20 men between the ages of 18 and 46 years (mean age = 26.8 years) with an average of 14.7 years of education. All subjects gave written informed consent following IRB regulations of the VANCHCS. Following the third test session in Experiment 2, these subjects were given a description of cognitive impairments that result from TBI, including difficulty concentrating, poor attentional focusing, fatigue, and impaired memory. They were instructed to develop a plan for simulating these impairments during a CCAB assessment in the following week.

Procedure

DS testing in Experiment 3 occurred 5–10 days after the final test session in Experiment 2. As in the test sessions of Experiment 2, forward and backward digit span testing was performed midway through the complete version of the CCAB.

Analysis

As in previous studies of malingering (Schwarz et al., 2006; Trueblood, 1994), we first summed FS and BS scores and compared subjects' average spans in Experiment 3 with their spans in Experiment 2. We also extracted three additional response-profile measures for each test session in Experiment 2 and Experiment 3: (a) the ratio of hit rates for the initial and final digit versus the hit rates for digits $N-1$ and $N-2$ of report; this metric reflected how well the malingering subject maintained appropriate serial-position functions while malingering; (b) the ratio of item/total errors; this metric reflected the extent to which subjects produced an appropriate mix of item and order errors when malingering; (c) the FS–BS difference score; this score reflected whether malingering subjects maintained similar simulated impairments in the two DS tests. Metrics b and c were gathered from both FS and BS test results.

Normative distributions for these response-pattern metrics were estimated from the 141 test sessions of Experiment 2. The probability of DSMI scores was estimated from this normative distribution after combining the probabilities of all five metrics using the Fisher method (Brown, 1975). The DSMI score for each subject in Experiment 3 was then compared to the normative DSMI distribution defined in Experiment 2. The results were evaluated with a chi-squared distribution adjusted to account for correlations observed between different malingering metrics. In order to evaluate DSMI reliability, this process was repeated 1,000 times using different random samples of 95% of Experiment 2 control data to estimate distributions.

Results

Of the 40 subjects in Experiment 3, 32 had abnormal spans in the malingering session (i.e., within the bottom 5% of average forward + backward MSs in Experiment 2). These subjects were categorized as abnormal malingering subjects. The average FS of these subjects was 4.96 ($SD = 0.96$), and the average BS was 3.93 ($SD = 1.02$): 2.47 digits and 1.88 digits, respectively, lower than comparable FS and BS in Experiment 2. In addition, six subjects produced average MSs that were reduced by at least 0.5 digits in comparison with their average results from Experiment 2, but which nevertheless remained within the normal range of scores. In addition, two subjects produced longer spans in Experiment 3 than in Experiment 2.

Serial-position functions of malingering subjects

Figure 5 shows the serial-position functions of malingering subjects. Although errors in forward span showed both primacy and recency effects, overall serial-position functions were flatter than those in control conditions. For example, first-position errors increasing by approximately 3-fold, and final-position errors increasing by approximately 2-fold in comparison with the results of Experiment 2 (cf. Figure 3). In BS testing, changes in serial-position functions were even larger: Final-position errors increased by approximately 9-fold (from 1.5% to 14%), and first-position errors approximately doubled.

Digit span malingering indices

Figure 6 shows average MS and DSMI scores for the 141 individual DS tests in Experiment 2. Random

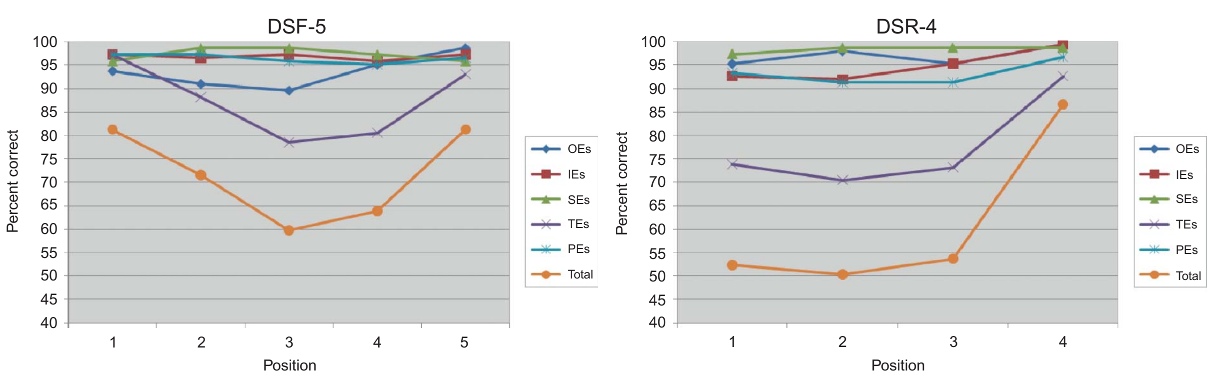

Figure 5. Experiment 3 (malingering) serial-position functions for total errors and errors of different types for digit span forward (DSF) list length of 5 digits and digit span backward (DSR) list length of 4 digits. OEs = omission errors, IEs = intrusion errors, SEs = substitution errors, TEs = transposition errors, PEs = permutation errors. To view a color version of this figure, please see the online issue of the Journal.

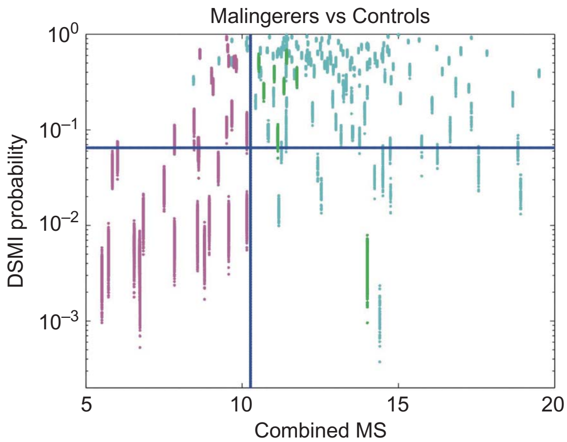

Figure 6. Comparison of combined mean span (MS; forward span, FS + backward span, BS) and DSMI (digit span malingering index) scores in Experiment 2 (light, cyan) and Experiment 3 (dark, red). The vertical crosshair shows the 5% cutoff for combined MS, and the horizontal crosshair represents the 7% cutoff of DSMI scores from Experiment 2 data. Purple shows DMSI and combined MS scores for malingering subjects with abnormal combined MSs. Green: malingering subjects with MSs that remained within the normal range. Data smearing reflects the range of probabilities calculated from 1,000 different random samples of control data. To view a color version of this figure, please see the online issue of the Journal.

permutations were used to estimate DMSI distributions in Experiment 2 to set a relatively strict DSMI criterion ($p < .07$) as shown in Figure 6 (horizontal line). An examination of the control data revealed that 15 tests (11%) exceeded this threshold on at least some of the 1,000 permutations. In contrast, among malingering subjects with abnormal average DSs, 19 of 32 tests (59%) equaled or exceeded the DSMI threshold on some permutations. Neither of the two subjects whose average MSs increased during “malingering” showed DSMI scores suggestive of malingering. Of the remaining six malingering subjects with shortened spans that still fell within the normal range, 33% produced DSMIs in the abnormal range.

Further analysis showed that there was a significant negative correlation between the magnitude of malingering (i.e., the difference in mean span between Experiment 2 and Experiment 3) and the DSMI p value ($r = -.42$, $p < .001$)—that is, the greater the magnitude of malingering the greater was the magnitude of DSMI abnormalities. Because subjects who malingered also produced short average MS scores, we also investigated whether the DSMIs correlated with average MS scores in subjects who were not malingering. There was no significant correlation between DSMI p-values and average MS scores either in the broad community sample tested in Experiment 1 ($r = .07$) or in Experiment 2 ($r = .04$).

Discussion

Subjects in Experiment 3 had been given one week to prepare a malingering strategy and were familiar both with the symptoms of TBI and with the procedures of digit span testing. As a result, more than 80% produced scores in the abnormal range. In previous studies using WAIS digit span, Schwartz and colleagues (2006) found that 86–90% of normal subjects who had been coached to malinger produced DS scores in the abnormal range.

Error pattern analysis of malingering subjects

Average serial-position functions were altered by malingering, with reduced primacy and recency effects. Error-type distributions were also affected, with some malingering subjects producing exclusively item or order errors, rather than the mixture of item and order errors normally seen in control subjects. Finally, some malingering subjects showed reduced differences between forward and backward MS scores.

DSMIs based on these parameters detected 59% of malingering subjects with abnormal spans. This suggests that DSMI abnormalities would assist in the confirmation of suspected malingering in the majority of cases where malingering subjects produce DS scores in the abnormal range. We also found that DSMI abnormalities increased as a function of the magnitude of malingering, making it particularly sensitive to large differences between observed and expected DS scores. Most importantly, DSMI p-values were uncorrelated with digit span performance in Experiment 2, or in the large population sample examined in Experiment 1. This suggests that subjects with intrinsically low digit spans nevertheless retain normal error pattern characteristics.

The sensitivity of the DSMI is surprising, given the small number of error trials (typically less than eight) that were subject to analysis. DSMI sensitivity reflects the fact that most malingering subjects adopt simple strategies: they produce low total spans by making consistent errors on each incorrect trial (e.g., always omitting a digit), they make frequent errors in reporting the first or last digit in the string, and they generally show smaller than normal differences between forward and backward span. Indeed, successful malingering would be a substantial cognitive challenge even for a well-trained malingerer. To avoid detection, the malingerer would have to err while mixing error types from trial to trial, producing errors in appropriate serial positions, and maintaining appropriate differences between FS and BS.

These results were obtained using the CCAB digit span test, which differs methodologically from WAIS DS testing in that each subject receives 14 digit lists regardless of performance. Because the digit list length increases with each correct response and is reduced after two failures, errors are typically produced on 5–9 trials. In the WAIS, DS testing ceases after two successive errors at any digit list length. Thus, DSMI analysis of WAIS digit span would likely show limited sensitivity because of the small sample of error trials available for analysis.

EXPERIMENT 4

The Traumatic Brain Injury (TBI) Clinical Trials Network has recommended digit span tests be included

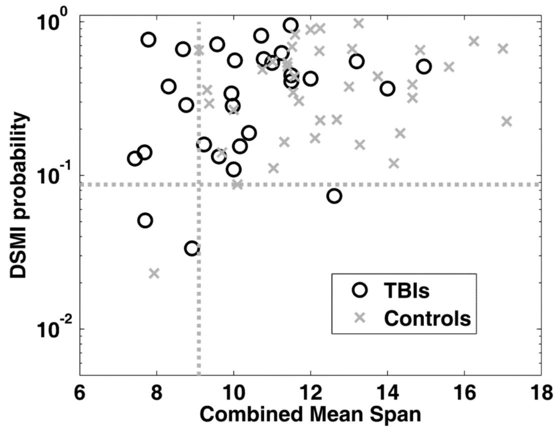

Figure 7. Comparison of combined mean span (MS; forward span, FS + backward span, BS) and DSMI (digit span malingering index) scores for 29 TBI (traumatic brain injury) patients (**o**) and 39 matched control subjects (×) in Experiment 4. Dashed vertical line: abnormality criteria ($p = .05$ of controls) for total digit span. Dashed horizontal line: abnormality criteria (7%) for the DSMI.

in evaluating clinical outcome (Bagiella et al., 2010), and DS abnormalities are seen in a significant percentage of TBI patients (Cicerone & Azulay, 2002; Curtiss, Vanderploeg, Spencer, & Salazar, 2001; Kersel, Marsh, Havill, & Sleigh, 2001), particularly in backward span testing (Chan, 2002; Conklin, Salorio, & Slomine, 2008; Fork et al., 2005; Wilson, Watson, Baddeley, Emslie, & Evans, 2000). However, many investigators have noted that digit span abnormalities in mild TBI are often associated with abnormal scores on symptom validity tests (Axelrod et al., 2006; Bailey, Echemendia, & Arnett, 2006; Greiffenstein & Baker, 2008; Heinly et al., 2005; West, Curtis, Greve, & Bianchini, 2010). In Experiment 4, we evaluated digit span performance and DSMI measures in 29 TBI patients with mild or severe TBI, who were tested at least one year post injury. The results were compared from data from 38 new control subjects matched in age and education level.

Method

Subjects

29 veterans with a history of TBI were recruited from the local VANCHCS patient population. The patients included 28 males and 1 female between the ages of 20 and 61 years (mean age = 36.7 years) with an average of 13.9 years of education. The patients had suffered TBIs of varying severity and etiology (Table 2) and were all tested at least one year post injury. All subjects had suffered head injuries and transient impairment of consciousness. Twenty-four of the patients had suffered one or more combat-related incidents with a cumulative loss of consciousness of less than 30 min, no hospitalization, and no evidence of brain lesions on clinical magnetic resonance imaging (MRI) scans. These patients were categorized as mild TBI. The remaining 5 patients had suffered severe accidents with hospitalization, coma duration exceeding eight hours, and posttraumatic amnesia exceeding 72 hours. These patients were categorized as severe TBI. Their results of TBI patients were compared with those of a control group of 38 newly recruited control subjects (12 females) matched in average age (range 18 to 66 years, mean 35.9 years) and education level (14.1 years) with the patients. All subjects gave written informed consent following IRB regulations of the VANCHCS.

Procedure

Subjects performed in a single CCAB test session with forward and backward digit span testing performed midway during testing, following methods described in Experiment 2.

Analysis

The analysis of digit span performance used the mean span measure described in Woods et al. (2010) and DSMI measures calculated as in Experiment 3.

Results

Group comparisons showed significant reductions in the TBI group for measures of forward span (controls = 7.0 vs. 6.1), $F(1, 65) = 10.1$, $p < .0011$, backward span (controls = 5.3 vs. 4.2), $F(1, 65) = 16.7$, $p < .00005$, and total span (controls = 12.3 vs. 10.3), $F(1, 65) = 15.8$, $p < .0001$. Although impairments were somewhat greater for backward than forward digit span for TBI patients, group differences in forward–backward difference scores failed to reach significance. Additional analysis showed no significant alterations in span-adjusted serial-position functions or the types of error produced by control and TBI patient groups.

Figure 7 shows combined span measures and DSMIs for the patients and control subjects.[2] There were no significant differences in DSMI measures between the two groups, $F(1, 65) = 0.40$, *ns*. Nine TBI patients (31%)

[2] The difference of 0.8 digits in the criterion ranges of Figure 6 and Figure 7 reflects learning effects due to repeated digit span testing in Experiment 3 as well as the use of a slightly older and less well educated control population in Experiment 4. See Woods et al. (2010) for further details.

TABLE 2
TBI patient characteristics

ID	*Age*	*EDU*	*ETIOL*	*TBI*	*TS*
P29	35	12	Veh. acc.	Severe	*7.70*
P28	24	12	Blast	Mild	11.25
P27	28	12	Blast	Mild	10.40
P26	46	12	Veh. acc.	Severe	13.20
P25	43	14	Blast[a]	Mild	*8.77*
P24	57	14	Veh. acc.	Severe	*7.79*
P23	39	17	Fall	Severe	14.00
P22	30	14	Veh. acc.	Mild	10.17
P21	52	14	Veh. acc.	Mild	11.00
P20	41	14	Blast[a]	Mild	11.48
P19	20	14	Blast[a]	Mild	12.62
P18	46	14	Veh. acc.	Severe	11.50
P17	25	15	Fall	Mild	9.94
P16	28	13	Blast	Mild	10.03
P15	25	12	Blast	Mild	*8.68*
P14	29	12	Blast	Mild	11.50
P13	47	14	Blast[a]	Mild	10.72
P12	28	14	Fall	Mild	14.95
P11	29	13	Blast	Mild	10.00
P10	61	18	Veh. acc.[a]	Mild	9.57
P09	27	15	Blast	Mild	9.98
P08	48	13	Blast	Mild	*9.23*
P07	50	20	Veh. acc.[a]	Mild	*7.43*
P06	49	12	Fall	Mild	*8.92*
P05	28	14	Fall	Mild	12.00
P04	39	13	Veh. acc.	Mild	10.78
P03	25	12	Blast[a]	Mild	*7.68*
P02	45	14	Veh. acc.	Mild	*8.31*
P01	23	14	Fall	Mild	9.62

Note. TBI = traumatic brain injury. Age in years. EDU = years of education. ETIOL = etiology. Veh. acc. = vehicle accident. TS = total span (forward + backward). Patients with TS scores in italics performed in the abnormal range.
[a]Multiple TBIs.

showed abnormal total spans relative to the $p = .05$ level of the control population, and 3 patients (11%) produced abnormal DSMIs.

Discussion

While digit span abnormalities are frequently encountered in patients with severe TBI (Kersel et al., 2001), in the current study we also found a high incidence of abnormalities (29%) among patients in the mild TBI group. One possible explanation is that the sensitivity of the computerized digit span test was improved in comparison with standard digit span testing because of its improved metric properties (Woods et al., 2010). However, the comorbidity of posttraumatic stress disorder (PTSD) may also have contributed to DS impairments. Overall, 62% of the TBI patients showed elevated scores on the posttraumatic stress disorder questionnaire (PTSD Checklist, PCL; Blanchard, Jones-Alexander, Buckley, & Forneris, 1996), and PCL scores correlated negatively with digit span in the TBI patient group, $r = -.41$, $t(25) = 2.25$, $p < .04$.

In contrast to previous reports of high levels of malingering in civilian TBI patient groups (Bailey et al., 2006; Curtis, Greve, & Bianchini, 2009; Greiffenstein & Baker, 2008; Whitney et al., 2009), we found minimal differences in DSMI scores of control subjects and the TBI patient group. Two factors may have been responsible for the low incidence of apparent malingering in the TBI patient group. First, the patients who underwent study were veterans who had volunteered for research studies to assist in TBI-related research. Second, it was made clear to the patients that the test results were for research purposes only and would not affect their medical or pension benefits.

GENERAL DISCUSSION

List length and serial-position effects

In Experiments 1 and 2, the types of error changed with list length. In short lists, transposition errors predominated, but their relative incidence diminished as list lengths increased, while the incidence of omission errors increased. It has been suggested that order errors are dependent on item errors (Conrad, 1965), but the findings from Experiments 1 and 2 indicate that order and item errors may be reflect mnemonic processes with different time courses (Aaronson, 1968; Bjork & Healy, 1974; Fuchs, 1969; Healy, 1974; McNicol, 1971). The serial-position functions of transposition and permutation errors in Experiments 1 and 2 produced bow-shaped curves with maximal error frequencies in midlist, similar to the order-error curves observed in prior studies (Aaronson, 1968; Bjork & Healy, 1974; McNicol, 1971). One explanation is that midlist items are vulnerable to a greater number of possible order confusions, including first- and second-order transpositions and third- and higher order permutations (Estes, 1972). As either end of the list is approached, the number of higher order transpositions and permutations is reduced so that end-of-list items can only perturb in one direction (Lee & Estes, 1977). Thus, the fact that transpositions and permutations occur more frequently in midlist suggests that transpositions and permutations compete with the correct digit ordering.

Serial-position effects were also observed for item errors. However, unlike the functions observed for order errors, item errors increased with report delay, with the poorest performance generally observed for the penultimate digit in the list. Possible explanations for this pattern include fewer rehearsals for items late in the list (Rundus & Atkinson, 1970) and a decline in activation strength for successive list items. Page and Norris (1998) suggested that item errors reflect a gradual decline in item accessibility. Once items fall below an accessibility threshold they cannot be retrieved, leading to omissions (Burgess & Hitch, 1992; Rundus & Atkinson, 1970). Our results suggest that an important component of accessibility may derive from accurate list-order information that can facilitate item recall.

Error analysis

Error analysis in Experiments 1 and 2 showed that transposition errors predominated at short list lengths, but decreased in relative incidence as list lengths increased. In contrast, omission errors were infrequent in short lists but increased rapidly as list length increased. This result is somewhat surprising, given the fact that the number of order transpositions increases as a factorial of list length. Because the number of possible omissions increases only linearly with list length, the ratio of possible order errors to possible omission errors increases markedly with list length. The fact that order errors increased much more slowly than the possible number of order errors as a function of list length suggests that list ordering confusions occur in small subsets of the digit sequence, most likely 2–3 digits in length. Our results suggested that the accuracy of pairwise digit ordering did not decrease with list length, suggesting that pairwise ordering information is relatively insensitive to temporal decay. This is also consistent with the serial-position functions for order information. In contrast, digit omissions increased markedly with the increasing temporal delay associated with longer lists. To summarize, digit order information appears to be error prone (i.e., TE1 errors are seen even in short lists) but decays slowly over time, while digit item information is nearly error free in short list lengths, but is subject to rapid temporal decay as lists lengthen.

Individual differences in digit span

The frequency of transposition errors showed an inverse correlation with mean digit span in Experiments 1 and 2. This was due both to the fact that transposition errors occurred more frequently in shorter lists and to the fact that subjects with shorter spans produced relatively more transposition errors, even when list lengths were adjusted to their MS. In contrast, subjects with longer spans had better preserved order information. Because of its slow temporal decay, preserved order information might provide a contextual framework to facilitate digit recall though order associations between preceding and following digits in the sequence. In contrast, subjects with poor order information would have difficulties in maintaining list integrity and would therefore benefit less from order-based recall cues.

Digit span error pattern analysis may also improve the clinical sensitivity of DS testing for certain neurological disorders (Helland & Asbjornsen, 2004), including leukoaraiosis (Lamar et al., 2007), language impairments in children (Gillam et al., 1995), and TBI in children (Warschausky, Kewman, & Selim, 1996). Serial-position functions may be altered by Alzheimer's (Burkart, Heun, & Benkert, 1998; Capitani, Della Sala, Logie, & Spinnler, 1992; Gibson, 1981) and Huntington's disease (Massman, Delis, & Butters, 1993). However, in Experiment 4 we found no significant abnormalities in either error patterns or serial-position functions in adult TBI patients.

Malingering detection

Within-assessment malingering measures, such as the DSMI, can supplement symptom-validity tests (Green, Allen, & Astner, 1997; Iverson, Franzen, & McCracken, 1991; Slick, Hopp, Strauss, & Thompson, 1997; Tombaugh, 1996). Malingering may be suspected if subjects show unexpectedly low DS performance (Axelrod et al., 2006; Bianchini, Greve, & Love, 2003; Iverson & Tulsky, 2003; Mittenberg et al., 1995). The DSMI supplements DS measures with the additional analysis of response patterns observed on incorrect trials. DSMI measures were abnormal in the majority of malingering subjects with abnormal DS scores, and DSMI abnormalities correlated with the magnitude of malingering. In contrast, DSMI scores were uncorrelated with DS scores in Experiment 1 and Experiment 2, and DSMI scores did not differ significantly between TBI and control groups in Experiment 4. Thus, the DSMI improves the robustness of malingering detection even among well-prepared and motivated malingerers and provides increased assurance that abnormal scores in TBI patients are not the result of malingering.

Original manuscript received 12 August 2010
Revised manuscript accepted 5 December 2010

REFERENCES

Aaronson, D. (1968). Temporal courses of perception in an immediate recall task. *Journal of Experimental Psychology*, *76*, 129–140.

Ardila, A. (2007). Normal aging increases cognitive heterogeneity: Analysis of dispersion in WAIS–III scores across age. *Archives of Clinical Neuropsychology*, *22*(8), 1003–1011.

Axelrod, B. N., Fichtenberg, N. L., Millis, S. R., & Wertheimer, J. C. (2006). Detecting incomplete effort with Digit Span from the Wechsler Adult Intelligence Scale–Third Edition. *The Clinical Neuropsychologist*, *20*(3), 513–523.

Babikian, T., Boone, K. B., Lu, P., & Arnold, G. (2006). Sensitivity and specificity of various digit span scores in the detection of suspect effort. *The Clinical Neuropsychologist*, *20*(1), 145–159.

Bagiella, E., Novack, T. A., Ansel, B., Diaz-Arrastia, R., Dikmen, S., Hart, T., et al. (2010). Measuring outcome in traumatic brain injury treatment trials: Recommendations from the traumatic brain injury clinical trials network. *The Journal of Head Trauma Rehabilitation*, *25*(5), 375–382.

Bailey, C. M., Echemendia, R. J., & Arnett, P. A. (2006). The impact of motivation on neuropsychological performance in sports-related mild traumatic brain injury. *Journal of the International Neuropsychological Society*, *12*(4), 475–484.

Benton, A. L. (1962). The visual retention test as a constructional praxis task. *Confinia Neurologica*, *22*, 141–155.

Bianchini, K. J., Greve, K. W., & Love, J. M. (2003). Definite malingered neurocognitive dysfunction in moderate/severe traumatic brain injury. *The Clinical Neuropsychologist*, *17*(4), 574–580.

Bjork, E. L., & Healy, A. F. (1974). Short-term order and item retention. *Journal of Verbal Learning and Verbal Behavior*, *13*, 80–97.

Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., & Forneris, C. A. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy*, *34*(8), 669–673.

Brown, M. B. (1975). A method for combining non-independent, one-sided tests of significance. *Biometrics*, *31*(4), 987–992.

Bunting, M. F., Cowan, N., & Colflesh, G. H. (2008). The deployment of attention in short-term memory tasks: Trade-offs between immediate and delayed deployment. *Memory & Cognition*, *36*(4), 799–812.

Burgess, A., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory & Language*, *31*, 429–460.

Burkart, M., Heun, R., & Benkert, O. (1998). Serial position effects in dementia of the Alzheimer type. *Dementia and Geriatric Cognitive Disorders*, *9*(3), 130–136.

Cameron, P. J. (1999). *Permutation groups*. Cambridge, UK: Cambridge University Press.

Capitani, E., Della Sala, S., Logie, R. H., & Spinnler, H. (1992). Recency, primacy, and memory: Reappraising and standardising the serial position curve. *Cortex*, *28*(3), 315–342.

Chan, R. C. (2002). Attentional deficits in patients with persisting postconcussive complaints: A general deficit or specific component deficit? *Journal of Clinical and Experimental Neuropsychology*, *24*(8), 1081–1093.

Cicerone, K. D., & Azulay, J. (2002). Diagnostic utility of attention measures in postconcussion syndrome. *The Clinical Neuropsychologist*, *16*(3), 280–289.

Conklin, H. M., Salorio, C. F., & Slomine, B. S. (2008). Working memory performance following paediatric traumatic brain injury. *Brain Injury*, *22*(11), 847–857.

Conrad, R. (1965). Order error in immediate recall of sequences. *Journal of Verbal Learning and Verbal Behavior*, *4*, 161–169.

Curtis, K. L., Greve, K. W., & Bianchini, K. J. (2009). The Wechsler Adult Intelligence Scale–III and malingering in traumatic brain injury: Classification accuracy in known groups. *Assessment*, *16*(4), 401–414.

Curtiss, G., Vanderploeg, R. D., Spencer, J., & Salazar, A. M. (2001). Patterns of verbal learning and memory in traumatic brain injury. *Journal of the International Neuropsychological Society*, *7*(5), 574–585.

Dean, A. C., Victor, T. L., Boone, K. B., Philpott, L. M., & Hess, R. A. (2009). Dementia and effort test performance. *The Clinical Neuropsychologist*, *23*(1), 133–152.

Estes, W. K. (1972). An associative basis for coding and organisation in memory. In A. W. Metlon & E. Martin (Eds.), *Coding processes in human memory* (pp. 161–190). Washington, DC: Winston & Sons.

Fisher, H. L., & Rose, D. (2005). Comparison of the effectiveness of two versions of the Rey memory test in discriminating between actual and simulated memory impairment, with and without the addition of a standard memory test. *Journal of Clinical and Experimental Neuropsychology*, *27*(7), 840–858.

Fork, M., Bartels, C., Ebert, A. D., Grubich, C., Synowitz, H., & Wallesch, C. W. (2005). Neuropsychological sequelae of diffuse traumatic brain injury. *Brain Injury*, *19*(2), 101–108.

Fuchs, A. (1969). Recall for order and content of serial word lists in short-term memory. *Journal of Experimental Psychology*, *82*, 14–21.

Gibson, A. J. (1981). A further analysis of memory loss in dementia and depression in the elderly. *The British Journal of Clinical Psychology*, *20*(3), 179–185.

Gillam, R. B., Cowan, N., & Day, L. S. (1995). Sequential memory in children with and without language impairment. *Journal of Speech and Hearing Research*, *38*(2), 393–402.

Gillam, R. B., Cowan, N., & Marler, J. A. (1998). Information processing by school-age children with specific language impairment: Evidence from a modality effect paradigm. *Journal of Speech, Language, and Hearing Research*, *41*(4), 913–926.

Green, P., Allen, L., & Astner, K. (1997). *The Word Memory Test: A manual for the oral and computerized forms*. Durham, NC: CogniSyst.

Greiffenstein, M. F., & Baker, W. J. (2008). Validity testing in dually diagnosed post-traumatic stress disorder and mild closed head injury. *The Clinical Neuropsychologist*, *22*(3), 565–582.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*, *6*(3), 218–224.

Greve, K. W., Springer, S., Bianchini, K. J., Black, F. W., Heinly, M. T., Love, J. M., et al. (2007). Malingering in toxic exposure: Classification accuracy of reliable digit span and WAIS–III Digit Span scaled scores. *Assessment*, *14*(1), 12–21.

Gusfield, D. (1997). *Algorithms on strings, trees, and sequences: Computer science and computational biology*. Cambridge, UK: Cambridge University Press.

Healy, A. F. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *13*, 644–655.

Heinly, M. T., Greve, K. W., Bianchini, K. J., Love, J. M., & Brennan, A. (2005). WAIS digit span-based indicators of malingered neurocognitive dysfunction: Classification accuracy in traumatic brain injury. *Assessment*, *12*(4), 429–444.

Helland, T., & Asbjornsen, A. (2004). Digit span in dyslexia: Variations according to language comprehension and mathematics skills. *Journal of Clinical and Experimental Neuropsychology*, *26*(1), 31–42.

Henson, R. N. (1998). Short-term memory for serial order: The Start–End Model. *Cognitive Psychology*, *36*(2), 73–137.

Iverson, G. L., & Franzen, M. D. (1994). The Recognition Memory Test, Digit Span, and Knox Cube Test as markers of malingered memory impairment. *Assessment*, *1*, 323–334.

Iverson, G. L., & Franzen, M. D. (1996). Using multiple objective memory procedures to detect simulated malingering. *Journal of Clinical & Experimental Neuropsychology*, *18*(1), 38–51.

Iverson, G. L., Franzen, M. D., & McCracken, L. M. (1991). Evaluation of a standardized instrument for the detection of malingered memory deficits. *Law and Human Behavior*, *15*, 667–676.

Iverson, G. L., & Tulsky, D. S. (2003). Detecting malingering on the WAIS–III. Unusual Digit Span performance patterns in the normal population and in clinical groups. *Archives of Clinical Neuropsychology*, *18*(1), 1–9.

Jurden, F. H., Laipple, J. S., & Jones, K. T. (1993). Age differences in memory-span errors: Speed or inhibitory mechanisms? *The Journal of Genetic Psychology*, *154*(2), 249–257.

Kemtes, K. A., & Allen, D. N. (2008). Presentation modality influences WAIS Digit Span performance in younger and older adults. *Journal of Clinical and Experimental Neuropsychology*, *30*(6), 661–665.

Kersel, D. A., Marsh, N. V., Havill, J. H., & Sleigh, J. W. (2001). Neuropsychological functioning during the year following severe traumatic brain injury. *Brain Injury*, *15*(4), 283–296.

Kessels, R. P., van den Berg, E., Ruis, C., & Brands, A. M. (2008). The backward span of the Corsi Block-Tapping Task and its association with the WAIS–III Digit Span. *Assessment*, *15*(4), 426–434.

Kramer, J. H., Jurik, J., Sha, S. J., Rankin, K. P., Rosen, H. J., Johnson, J. K., et al. (2003). Distinctive neuropsychological patterns in frontotemporal dementia, semantic dementia, and Alzheimer disease. *Cognitive and Behavioral Neurology*, *16*(4), 211–218.

Lamar, M., Price, C. C., Libon, D. J., Penney, D. L., Kaplan, E., Grossman, M., et al. (2007). Alterations in working memory as a function of leukoaraiosis in dementia. *Neuropsychologia*, *45*(2), 245–254.

Larrabee, G. J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist*, *17*(3), 410–425.

Lee, C. L., & Estes, W. K. (1977). Order and position information in primary memory for letter strings. *Journal of Verbal Learning and Verbal Behavior*, *16*, 395–418.

Massman, P. J., Delis, D. C., & Butters, N. (1993). Does impaired primacy recall equal impaired long term storage? Serial position effects in Huntington's disease and Alzheimer's disease. *Developmental Neuropsychology*, *9*, 1–15.

McNicol, D. (1971). The confusion of order in short-term memory. *Australian Journal of Psychology*, *23* (1), 77–84.

Miller, L. J., Ryan, J. J., Carruthers, C. A., & Cluff, R. B. (2004). Brief screening indexes for malingering: A confirmation of Vocabulary minus Digit Span from the WAIS–III and the Rarely Missed Index from the WMS–III. *The Clinical Neuropsychologist*, *18*(2), 327–333.

Mittenberg, W., Theroux-Fichera, S., Zielinski, R. E., & Heilbronner, R. (1995). Identification of malingered head injury on the Wechsler Adult Intelligence Scale–Revised. *Professional Psychology: Research and Practice*, *26*, 491–498.

O'Caroll, R. E., & Gilleard, C. J. (1986). Estimation of premorbid intelligence in dementia. *British Journal of Clinical Psychology*, *25*(2), 157–158.

Page, M. P., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, *105*(4), 761–781.

Pebler, P. (2011). Presentation software. Albany, CA: NeuroBehavioral Systems, Inc.

Richardson, J. T. (2007). Measures of short-term memory: A historical review. *Cortex*, *43*(5), 635–650.

Rundus, D., & Atkinson, R. C. (1970). Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior*, *9*, 99–105.

Ruocco, A. C., Swirsky-Sacchetti, T., Chute, D. L., Mandel, S., Platek, S. M., & Zillmer, E. A. (2008). Distinguishing between neuropsychological malingering and exaggerated psychiatric symptoms in a neuropsychological setting. *The Clinical Neuropsychologist*, *22*(3), 547–564.

Schwarz, L. R., Gfeller, J. D., & Oliveri, M. V. (2006). Detecting feigned impairment with the digit span and vocabulary subtests of the Wechsler Adult Intelligence Scale–Third Edition. *The Clinical Neuropsychologist*, *20*(4), 741–753.

Shapiro, A. M., Benedict, R. H., Schretlen, D., & Brandt, J. (1999). Construct and current validity of the Hopkins Verbal Learning Test–Revised. *Clinical Neuropsychology*, *13*(3), 348–358.

Shum, D. H., O'Gorman, J. G., & Alpar, A. (2004). Effects of incentive and preparation time on performance and classification accuracy of standard and malingering-specific memory tests. *Archives of Clinical Neuropsychology*, *19*(6), 817–823.

Slick, D. J., Hopp, G., Strauss, E., & Thompson, G. (1997). *The Victoria Symptom Validity Test*. Odessa, FL: PAR.

Strauss, E., Slick, D. J., Levy-Bencheton, J., Hunter, M., MacDonald, S. W., & Hultsch, D. F. (2002). Intraindividual variability as an indicator of malingering in head injury. *Archives of Clinical Neuropsychology*, *17*(5), 423–444.

Tombaugh, T. N. (1996). *Test of Memory Malingering (TOMM)*. San Antonio, TX: PsychCorp.

Trueblood, W. (1994). Qualitative and quantitative characteristics of malingered and other invalid WAIS–R and clinical memory data. *Journal of Clinical and Experimental Neuropsychology*, *16*(4), 597–607.

Vagnini, V. L., Sollman, M. J., Berry, D. T., Granacher, R. P., Clark, J. A., Burton, R., et al. (2006). Known-groups cross-validation of the letter memory test in a compensation-seeking mixed neurologic sample. *The Clinical Neuropsychologist*, *20*(2), 289–304.

Warschausky, S., Kewman, D. G., & Selim, A. (1996). Attentional performance of children with traumatic brain injury: A quantitative and qualitative analysis of digit span. *Archives of Clinical Neuropsychology*, *11*(2), 147–153.

Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale–administration and scoring manual* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale–administration and scoring manual* (3rd ed.). San Antonio, TX: The Psychological Corporation.

West, L. K., Curtis, K. L., Greve, K. W., & Bianchini, K. J. (2010). Memory in traumatic brain injury: The effects of injury severity and effort on the Wechsler Memory Scale–III. *Journal of Neuropsychology*.

Whitney, K. A., Davis, J. J., Shepard, P. H., Bertram, D. M., & Adams, K. M. (2009). Digit span age scaled score in middle-aged military veterans: Is it more closely associated with TOMM failure than reliable digit span? *Archives of Clinical Neuropsychology*, *24*(3), 263–272.

Wilson, B. A., Watson, P. C., Baddeley, A. D., Emslie, H., & Evans, J. J. (2000). Improvement or simply practice? The effects of twenty repeated assessments on people with and without brain injury. *Journal of the International Neuropsychological Society*, *6*(4), 469–479.

Woods, D. L., Kishiyama, M. M., Yund, E. W., Herron, T. J., Edwards, B., Poliva, O., et al. (2010). Improving digit span assessment of short-term verbal memory. *Journal of Clinical and Experimental Neuropsychology 33*, 101–111.