Running head: SVT IN CHRONIC PAIN

Symptom Validity Testing in Chronic Pain Patients Presenting for Disability Assessments

Roger O. Gervais, Paul Green

Neurobehavioral Associates, Edmonton, Alberta

Lyle M. Allen III

Durham, North Carolina

Grant L. Iverson

University of British Columbia and Riverview Hospital

Author Notes: Address correspondence to: Roger Gervais, Ph.D., 201, 17107-107 Ave. Edmonton, Alberta, T5S 1G3. Ph. (780) 486-6633, Fax: (780) 484-5552.

Abstract

Complaints of impaired memory and concentration are common in people with chronic pain. The purpose of this study was to examine one of the many factors that can make assessment difficult in this population, namely suboptimal effort on ability tests. A total of 118 chronic pain patients seen for a disability-related psychological evaluation were given two tests designed to detect suboptimal effort, the Computerized Assessment of Response Bias (CARB; Conder, Allen, & Cox, 1992; Allen, Conder, Green, & Cox, 1997) and the Word Memory Test (WMT; Green, Allen, & Astner, 1996). When standard instructions were given, failure rates were high on both tests, suggesting that a substantial number of pain patients were making insufficient effort to produce valid results. Both their effort test scores and their verbal memory test scores were significantly lower than those of patients with severe traumatic brain injury and neurological diseases. When subsequent patients were given additional information that CARB is unaffected by pain or emotional distress, the failure rates on the CARB dropped very significantly. However, being told that the WMT contains measures of effort did not significantly alter failure rates on any of the effort or memory measures in this test. This study illustrates the importance of assessing for biased responding in disability-related pain evaluations. It also shows that being given prior information that tests measure effort is likely to influence performance on some effort tests more than others.

Key Words: Symptom Validity, Exaggeration, Pain, Disability, Neuropsychology

Symptom Validity Testing in Chronic Pain Patients Presenting for Disability Assessments

Complaints of cognitive dysfunction are widespread in people with chronic pain (Iverson & McCracken, 1997) and the assessment of various cognitive abilities, such as memory, is often part of a comprehensive pain evaluation. Such assessments in pain patients are complicated by the possibility that pain, depression, fatigue, or medication could erode concentration (Kewman, Vaishampayan, Zald, & Han, 1991; Eccleston, 1995, 1997; Grigsby, 1995; Taylor, 1996; Branca, 1996; Di Stefano & Radnov, 1995; Smed, 1997; Karlsborg, Smed, Jespersen, Stephensen, Cortsen, Jennum, Herning, Korfitsen, & Werdelin, 1997). The assessment process can be challenging when the client is in litigation or is pursuing disability benefits, because some patients might exaggerate their pain and disability. Kay and Morris-Jones (1998) reported a very high rate of exaggeration of disability in litigating patients with chronic pain. Schnurr and MacDonald (1995) found that chronic pain patients with medico-legal “incentives” reported significantly more memory impairment than medical patients or people in psychotherapy.

In a meta-analysis of the literature, entitled “Money matters”, Rohling and colleagues (1995) concluded that compensation incentives are significantly related to symptom presentation in patients with chronic pain. In a review of cognition and chronic pain, Hart, Martelli and Zasler (2000) concluded that the use of specific methods to detect invalid test results has a valuable role to play in the clinical assessment of chronic pain-related disability.

Within neuropsychological evaluations, two-alternative, forced-choice "symptom validity tests" (SVTs) or "effort tests" are widely recommended to detect possible exaggeration of cognitive deficits in the results of ability testing (Binder & Willis, 1991; Hiscock & Hiscock, 1989; Green & Iverson, in press; Green, Iverson, & Allen, 1999; Slick, Hopp, Strauss, Hunter, & Pinch, 1994). These SVTs, many of which have been reviewed recently by Iverson and Binder (2000), Sweet (1999), and Reynolds (1998), are based on the principle that patients with known severe impairment, resulting from traumatic brain injury or neurological disease, can easily score above a certain level. Therefore, test results from people with no neurological compromise who score below these levels are of questionable validity. For example, Green et al., (1999) reported that patients with mild head trauma scored significantly lower than patients with well-established moderate to severe brain injury on both the effort measures and the memory subtests of the Word Memory Test (WMT; Green, Allen, & Astner, 1996). This finding was contrary to well established knowledge that severe brain injury produces more severe cognitive impairment than mild head trauma and, therefore, the test results from the mild head trauma group were of doubtful validity.

Schmand, Lindeboom, Schagen, Heijt, Koene, and Hamburger (1998), reported that cognitive exaggeration on the Amsterdam Short-term Memory Test was evident in 61% of litigating post-whiplash patients. Similar results were documented by Vickery, Ranseen, Cooley, and Berry (1999) who administered three SVTs and response bias tests (Digit Memory Test, Hiscock & Hiscock, 1989; Rey 15-Item Test and Rey Word List, Lezak, 1983) to 34 adults complaining of pain and emotional distress resulting from traumatic injuries not involving the head. The authors found that 53% of the sample fell below the cutoff on at least one test, and 27% fell below the cutoffs on two or more effort tests. They recommended the use of tests to rule out exaggeration of cognitive symptoms in the neuropsychological evaluation of patients presenting with no central nervous system injury. Allen, Conder, Green, and Cox (1997) reported that patients with DSM-IV chronic pain disorder had the highest rate of failure on the WMT and the Computerized Assessment of Response Bias (CARB; Conder, Allen, & Cox, 1992), among seven separate diagnostic groups of disability claimants. Gervais, Green, and Allen (1999) administered the CARB and two other SVTs, the Test of Memory Malingering (TOMM; Tombaugh, 1996) and the WMT to 150 consecutive patients undergoing psychological assessment for chronic pain disability, personal injury or vocational guidance. Using the cutoffs for failure defined in the respective test manuals, they found that 27% of the chronic pain cases failed all three effort tests. In contrast, no more than six percent failed any effort test in the vocational assessment group, consisting of people seeking financial support for post-secondary education. In the latter studies, SVT failure indicated performance falling below the lower limits of scores from patients with documented brain dysfunction due to traumatic injury or neurological conditions (Pankratz & Binder, 1997; Tombaugh, 1996) and SVT failure was not based on worse than chance responding. Iverson and Binder (2000) have argued convincingly that reliance upon a worse-than-chance criterion on any test will lead to an unacceptably high rate of false negative classifications of incomplete effort.

The current study began in response to an ethical and practical dilemma, arising from unexpectedly high rates of failure on SVTs in patients with chronic pain. On reviewing 57 cases assessed for chronic pain before any experimental question had arisen, it was noted that chronic pain patients were failing the effort tests (CARB and WMT) more than any other diagnostic group, including compensation claimants with mild head injuries. The conventional interpretation of SVT failure would be that these pain patients exerted inadequate effort to produce valid test results and, thus, the test results could not be used to support their claims of cognitive impairment. However, because the patients' disability claims could be adversely affected by such conclusions, it was felt that there was a need to rule out the alternative possibility, which was that, hypothetically, their pain could have prevented them from being able to perform well on these tests. To resolve this dilemma, we decided to test the next series of chronic pain patients referred for disability assessments in such a way that we could determine, in each case, whether SVT failure was under voluntary control or whether it was determined by pain-related factors outside of the patient's control.

Previous research had shown that coaching prior to testing increased scores on both the WMT and the CARB in an experimental study with simulated malingerers, who were asked to perform like someone with a brain injury (Dunn, Shear, Howe, & Ris, 1999). When given general guidance on methods for defeating SVTs, the simulators' scores on the CARB and WMT were significantly higher than those given no advice about the tests. Therefore, we decided to use selective coaching to test whether patients were able to pass the effort tests when coached. In the baseline clinical testing phase, we had already employed two SVTs using standard test instructions. For the remaining patients, we used one SVT to measure effort and the other as a control for the possibility that pain or distress had caused failure on the other SVT. We gave the first new group of patients additional information about the CARB before it was administered but the WMT was given under standard conditions. In the next phase, another new group of pain patients was given additional information about the WMT before testing but the CARB was given under standard conditions. If such selective coaching had no effect on test scores, it would suggest that some chronic pain patients were genuinely unable to pass these tests and this would have significantly altered our interpretation of the clinical SVT results. Alternatively, if we found a coaching-related improvement on either SVT, this would suggest that failure on the SVTs was under voluntary control and it would further support the utility of effort tests in chronic pain disability evaluations.

To provide an additional perspective in this study, we compared the test results from pain patients with those of a group of previously tested patients with verified brain damage from severe traumatic brain injury or neurological disease (Green & Allen, 1999; Allen & Green, 1999). On the assumption that pain patients did not have cerebral dysfunction, we hypothesized that the pain patients' scores on the SVTs and on independent measures of memory would be significantly greater than those of organically impaired patients. The finding of equal or significantly lower SVT or memory scores in the pain patients would be difficult to explain on a neurological basis. If such an implausible pattern of results were found, in conjunction with failed effort tests, it would suggest suboptimal effort in some chronic pain patients.

Method

Participants

The participants included 118 patients claiming total disability benefits due to chronic pain arising out of work-related injuries or motor-vehicle accidents. The independent medical consensus prior to psychological evaluation was that they all met the criteria for DSM-IV Pain Disorder. The reported pain severity and duration of disability were deemed in excess of objective findings and disproportionate to their accident injuries. The patients were referred for a standardized psychological assessment to determine whether their pain was causing significant psychological and psychosocial distress, as determined by objective psychological testing, self-report inventories, and a structured clinical interview. In this clinical practice, both the CARB and WMT have been administered routinely to all patients since 1996, irrespective of diagnosis. They were included in a comprehensive test battery developed for chronic pain assessments, consisting of specific tests requested by the referral agency (including at least one SVT), as well as a number of other standardized instruments selected to measure self-reported symptoms (e.g. MMPI-2 and the SCL-90-R).

The mean age of the chronic pain patient sample was 47 years (SD = 9.1), their mean education was 10.4 years (SD = 2.8) and 59% were male. On a six point scale, with zero representing “totally pain free” and five representing “worst imaginable pain”, the patients obtained an average self-rating of 3.9 (SD = .77) at the time of taking the CARB. They also reported a mean of 2.9 (SD = 1.2) different body sites for their pain, the most common sites of pain being arm, leg or shoulder pain (87%), followed by back pain (67%) and neck pain (57%).

For comparison with the chronic pain patients' results on the CARB and the WMT, we used data from 97 patients with known neurological compromise previously described in detail by Green and Allen (1999) and Allen and Green (1999). This sample was comprised of 57 moderately to severely brain injured patients and 40 patients with neurological diseases, including brain tumors, ruptured aneurysms, strokes and multiple sclerosis. Fifty-three of the 57 brain injured patients had CT or MRI data available and in 85% there was evidence of intracranial damage. Their mean GCS was 9.64 (SD = 4.15) and their mean duration of PTA was 406.2 hours (SD = 598.4). The diagnoses represented in the 40 neurological patients included: brain tumor (n = 5), ruptured aneurysm (n = 8), stroke (n = 10), multiple sclerosis (n = 6), epilepsy (n = 3), and a variety of other diagnoses (n = 8). These patients had all been given the CARB and the WMT as part of a comprehensive clinical neuropsychological assessment related to workers' compensation, medical disability or personal injury claims. In the present study, the patient group made up of people with moderate to severe brain injury and neurological patients will be referred to as the "neurological-STBI" group (N-STBI; N = 97). The 97 patients shown in Tables 1, 2, and 3 were assumed to be making a full effort, which is consistent with their almost perfect scores on the CARB and WMT effort measures. The N-STBI sample was 73% male, was an average age of 41.8 years of age (SD = 11.5) and had an average education of 12.2 years (SD = 2.8). Their self-reported pain rating was .75 (SD = 1.1) on the six point pain scale. They reported a mean of 1.3

(SD = 1.3) body sites for their pain, with the most common sites of pain being the neck (33%), back (33%), temporomandibular joint (25%), and arm, leg, or shoulder (25%). The N-STBI sample reported significantly less pain than the chronic pain patients (t = -22.99, df = 171.35, $p < .0005$) and fewer body sites for their pain (t = -5.24, df = 77.62, $p < .0005$). Compared with the N-STBI sample, the chronic pain group more frequently reported head pain (X^2 = 26.56, df = 1, $p < .0005$), neck pain (X^2 = 8.25, df = 1, $p < .01$), back pain (X^2 = 17.43, df = 1, $p < .0005$), and shoulder, arm, or leg pain (X^2 = 43.16, df = 1, $p < .0005$).

Compared with all chronic pain patients combined, the N-STBI sample was significantly younger (t = -3.59, df = 181, $p < .0005$), contained significantly more males (X^2 = 5.07, df = 1, $p < .05$), had significantly more years of education (t = 4.92, df = 209, $p < .0005$), and had a significantly lower proportion of patients for whom English was a second language but who, like many Canadians, were bilingual as opposed to nonfluent (ESL; X^2 = 8.75, df = 1, $p < .01$). These groups did not differ in conversational English competency, which was judged to be present if the person could easily be interviewed and tested without the need for an interpreter (X^2 = .80, df = 1, $p > .37$). Age and years of education were unrelated to performance on the CARB and the WMT in a heterogeneous sample of 307 compensation patients (Allen et al., 1997). These authors examined a set of 1752 CARB test results from compensation evaluations conducted in 13 sites in the US and Canada and found an overall CARB failure rate of 30% (range 21% to 76% per site) but no significant correlations were found between total CARB score and age (r = -.0016) or sex (r = -.0037). A very weak positive relationship was found between CARB performance and years of education (r = .065; p < .01), but this accounted for less than one half of one percent of the total variance.

Assessment Methods

In the course of their chronic pain assessment, all the pain patients underwent detailed structured clinical interviews regarding how pain affected their lives and they were administered the CARB, the WMT and a number of psychological tests, inventories, and neuropsychological tests. The total time required for completion of the psychological test battery varied from four to six hours. The Computerized Assessment of Response Bias (CARB) is a two-alternative forced-choice digit recognition task, which is designed to detect probable invalid test results arising from inconsistent or suboptimal effort. Iverson and Binder (2000) described the CARB as one of the best validated forced-choice symptom validity tests. The patient is presented with three sets of 37 five-digit numbers on the computer screen. Each five-digit number is followed by a message to count backwards from 20. After a brief delay, two five-digit numbers, one of which was the original number, appear on the left and right sides of the computer screen. The patient must choose which of the two numbers was originally displayed by selecting either the left or the right shift key. For test security purposes, it is conventional not to specify the cut-offs for SVTs in publications. A total score below the cut-off score described in the test manual defined failure on the CARB. This cutoff score is over 3.8 SD below the mean score observed in 62 patients with moderate to severe brain injury presumed to be making a good effort (Green et al., 1999; Allen & Green, 1999; Conder et al., 1992). The

CARB's

early termination feature for good performance was used in this study, as recommended by Allen et al. (1997).

The CARB required less than five minutes to complete for a patient demonstrating full effort. Otherwise, the administration could take up to 20 minutes. Similarly, the WMT required approximately 20 minutes of the patient's working time. The routine use of the CARB and the WMT is standard practice with all patients in this clinical setting. Consistent with general clinical methods at that time and prior to undertaking the assessment, the pain patients were informed both verbally and in writing that the assessment contained measures of effort. They were requested to advise the assistant if they were unable to provide their best effort for any reason and they signed a consent form acknowledging that they understood and accepted these conditions.

The Word Memory Test (WMT) measures verbal learning and memory and it was designed with built-in indices of biased responding. It is a computerized list-learning task, in which the subject is presented with 20 semantically linked word pairs (e.g. pig--bacon, man--woman), each pair appearing for six seconds. After the entire list is presented twice, there is an immediate recognition trial (IR), in which the person is shown new word pairs containing only one of the words from the original list and must select the words belonging to the original list. Without advance warning, a similar delayed recognition procedure (DR) is administered after 30 minutes, using different foil words paired with the original words. Scoring of the computerized WMT includes a calculation of consistency of responding from IR to DR (WMT-CONS), as does its predecessor, the orally administered WMT. The WMT response consistency measure has been shown to be sensitive to patient effort (Green et al., 1996). Response consistency as a measure of patient compliance was extensively discussed by Frederick and Foster (1991) and was incorporated into the design of the Validity Indicator Profile (VIP; Frederick, 1997; Frederick & Crosby, 2000). The DR subtest is followed by a series of tests of gradually increasing difficulty, measuring verbal memory, including multiple choice (MC), paired associates (PA), delayed free recall (DFR) and long delayed free recall (LDFR) procedures. Based on the results of patients with severe traumatic brain injuries or neurological disorders (Allen & Green, 1999; Green & Allen, 1999), failure on the WMT in the current study was defined as a score on the WMT-DR subtest below the cut-off recommended in the test manual (Green et al., 1996). In a recent study of 298 consecutive head injury patients, 64 cases with moderate to severe brain injury obtained a mean score well above the cut-offs on all of the WMT effort measures (Green et al., 1999). When response bias is evident in failure on either the CARB or WMT effort measures, it raises questions about the validity of the patient's test results and self-reported symptoms and, therefore, about the validity of the claimed disability (Green et al., 1996).

The Memory Complaints Inventory (MCI; Green & Allen, 1997), a 58-item computer administered self-report inventory of memory problems, was incorporated into the test battery in 1997 and was completed by the last 59 consecutive pain patients in this study. The instrument contains nine scales designed to identify specific types of reported memory problems ranging from common to implausible. The first six scales involve the most common and plausible memory complaints. The final three scales describe relatively implausible memory problems, which are rarely found in patients with bona fide memory impairment of organic origin. These include complaints such as: "There are big gaps in my memory of my childhood" (remote amnesia), "Minutes or hours pass by and I have no idea of what I have been doing" (amnesia for complex behavior), and "I have hit someone and had no memory of doing it" (amnesia for antisocial behaviour). Please see Table 1 for more details of the scales.

Conditions

The pain patient baseline group (phase 1) was derived from archival data from 57 consecutive pain patients seen for disability evaluations by the first and second authors. Under standard test instructions, these patients showed very high failure rates on effort tests, with 41% failing the CARB and 42% failing the WMT effort measures. The next 67 patients referred with Pain Disorder were selectively coached on one SVT or the other. A total of 6 patients, who had previously been exposed to the CARB in an earlier assessment and were re-referred, were excluded from the study, leaving 61 patients. In the first coaching condition, we informed 32 patients that we expected that they would easily be able to pass the CARB. After the standard instructions, they were told:

> "In your case, we are almost certain that your pain will not have any effect on this test because, in people with pain, it really just tests effort. You will see that the test is extremely easy and we have no doubt that you will be able to score almost 100% correct. If we found that you scored much lower than this, we would have to question whether you were making an effort. I

strongly recommend that, on this test, you make your best effort."

The patients were also told that the test would last longer if they did not make a good effort. These patients completed the CARB at or near the end of their evaluation. No information was given to these patients regarding the WMT, which was administered earlier in the day with standard instructions.

The second warning condition involved the next 29 pain patients, who constituted the WMT-Coached group. Following the standard WMT instructions and just before the first part of the test began, they were given the same additional information as the CARB-Coached patients. They were not told that there would be several other subtests after a half-hour delay and nor were they told which WMT subtests measured effort and which measured memory ability. No warning was given before starting the delayed recognition subtest (DR), 30 minutes after the immediate recognition subtest (IR). In the WMT-Coached group, no information was given regarding the effort measures in the CARB, which was administered with standard instructions earlier in the day, before the WMT.

There were no differences between any of the pain patient groups in the three phases of the study (baseline, CARB coaching and WMT coaching) on age ($\underline{F}$ (2,115) = 1.65, $\underline{p}$ >.19), gender ($\underline{X}^2$ = 1.95, $\underline{df}$ = 2, $\underline{p}$ > .37), years of education ($\underline{F}$ (2,111) = 1.86, $\underline{p}$ > .88), handedness ($\underline{X}^2$ = 2.47, $\underline{df}$ = 2 $\underline{p}$ > .29), or incidence of English as a second language (ESL), ($\underline{X}^2$ = 4.88, $\underline{df}$ = 2, $\underline{p}$ > .08). The pain patients in the baseline group and in the two informed conditions did not differ from each other in their reported number of pain sites ($\underline{F}$ (2,82) = 2.08, $\underline{p}$ > .13), subjective ratings of pain at the time of the assessment ($\underline{F}$ (2,90) = .62, $\underline{p}$ > .54) or on the Pain Severity scale of the Multidimensional Pain

Inventory (F (2,88) = .59, p > .55) (MPI; Kerns, Turk, & Rudy, 1985). There was no significant association between ESL and number of reported pain sites in the total sample (r = -0.19, p > .07).

Results

The MCI scores for the chronic pain and N-STBI groups are presented in Table 1, which shows that the pain patients in this study complained of more memory problems than patients with impairment from neurological disease or severe brain injury, as shown by their scores on the MCI (Green & Allen, 1997; Green et al., 1999). The chronic pain patients obtained a mean score of 64% of the maximum possible score (SD = 27.0) on the Pain Interferes with Memory (PIM) scale, denoting significant endorsement of pain-related memory complaints (e.g. "My pain makes it impossible to think clearly or remember things" and "Pain interferes with my concentration and memory."), whereas the N-STBI mean score on this scale was 28.1 (SD = 30.7). A MANOVA comparing the 59 current pain patients with 51 of the N-STBI for whom MCI results were also available, found a significant between group difference only on the Pain Interferes with Memory (PIM) scale (F = 42.54, p < .0005), and the mean overall MCI score (F = 4.40, p < .05).

Insert Table 1 about here

Failure rates, as well as means and standard deviations of CARB and each of the three WMT effort measures are shown in Table 2 for the neurological-STBI (N-STBI) patients and for each coached pain patient group in the study. Because the scores on

CARB and the effort components of the WMT are not normally distributed, nonparametric Kruskal-Wallis one-way ANOVAs for independent samples were employed to compare group performances on the main dependent measures. Whenever these models yielded significant differences, the Mann-Whitney U test was employed for comparisons between groups. To control for the possibility of spurious results, Bonferroni corrections were employed for the post-hoc Mann-Whitney U tests based on the formula .05 divided by the number of comparisons.

Insert Table 2 about here

The effects of coaching on the CARB were assessed by comparing total CARB scores across each of the three pain groups, which were Baseline, CARB-Coached, and WMT-Coached. A significant overall Kruskal-Wallis ANOVA model was produced ($X^2 = 23.4$, $df = 2$, $p < .0005$). A subsequent series of paired Mann-Whitney comparisons for mean total CARB scores between the three groups was undertaken using .016 as the criterion for significance (.05 / 3 comparisons). The CARB-Coached group obtained significantly higher CARB scores than the Baseline ($U = 369.0$, $z = -4.65$, $p < .0005$) and WMT-Coached groups ($U = 210.0$, $z = -3.81$, $p < .0005$), indicating a strong coaching effect. The CARB failure rate dropped from 41% in the Baseline group to only 6% in the CARB-Coached group ($X^2 = 12.52$, $df = 2$, $p < .01$).

A Kruskal-Wallis one-way ANOVA comparison for performance on WMT-DR across the three pain conditions (Baseline, CARB-Coached and WMT-Coached) failed to reach significance ($\underline{X}^2 = .48$, $\underline{df} = 2$, $\underline{p} > .78$), demonstrating that coaching before the start of the WMT had no effect on the WMT 30 minute delayed recognition scores (WMT-DR). The 41% WMT failure rate in the WMT-Coached group was virtually unchanged from the 42% rate of failure in the Baseline group. There was also no significant difference between conditions on WMT immediate recognition (WMT-IR, $\underline{X}^2 = .49$, $\underline{df} = 2$, $\underline{p} > .78$).

To investigate the discrepancy observed between the performance of pain patients on CARB and the WMT, a series of eight comparisons were undertaken employing the N-STBI sample (significance criterion = .05 / 8 = .006). These patients not only had well-established neurological damage but, as reported by Allen and Green (1999) and Green and Allen (1999), their memory was in the impaired range on the California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Ober, 1987). To reduce the total number of comparisons undertaken, the chronic pain disorder patients from all conditions were combined together into a single group ($\underline{N} = 118$). In spite of the fact that this group contained many cases coached on CARB, their average total CARB score was only 88.7% ($\underline{SD} = 16.8$), which was significantly lower than that of the 97 N-STBI patients (mean 97.8%, $\underline{SD} = 3.6$; $\underline{U} = 3776.0$, $\underline{p} < .0005$). The 118 pain patients also obtained significantly lower mean scores than the N-STBI group on each WMT effort measure and also on every one of the verbal memory measures contained within the WMT. Complete results for these comparisons employing the Mann-Whitney U test are contained in Table 3.

Insert table 3 about here

Discussion

The results from this study are compelling. Over 40% of patients seen initially for pain-related psychological disability evaluations scored below the cutoffs on each of two separate SVTs. These performance-based measures were designed to identify patients making insufficient effort to produce valid results on measures of actual cognitive ability. Therefore, incomplete effort and probable cognitive exaggeration in the pain group were suggested by their high failure rate on the CARB and WMT.

One interpretation might have been that these patients failed the SVTs or effort tests because of pain or emotional distress and it was this clinically important hypothesis that prompted the present study. However, when similar pain patients were told that the CARB was extremely easy, that it just tested effort in their cases and that their test scores would not be affected by their pain, they performed very significantly better than when they were tested with standard instructions. The coaching effect for the CARB was demonstrated within a three-phase design. The CARB failure rate was 41% in the baseline group. The failure rate dropped to only 6% in the CARB-Coached group, even though 34% of these cases failed the effort measures of the WMT, about which they had not been coached. In the next and final patient series, CARB was administered without any coaching and 38% of cases failed it.

The additional information provided about the CARB would not be expected to relieve pain, fatigue, or emotional distress. Therefore, the fact that coaching patients on the CARB caused a significant upward shift in test scores and a highly significant decline in overall failure rate suggests that it is very unlikely that the patients who failed the CARB did so because of pain, fatigue, or emotional distress. If pain and emotional distress exerted a major influence on effort test scores, it would be difficult to explain why the CARB-Coached patients performed so much better on the CARB but showed no significant change in their rate of failure on the WMT. They reported the same amount of pain as the baseline patients. Moreover, if fatigue had a systematic influence on these results, then the CARB performance of the coached subjects should have been the most affected because, for these patients, it was administered at the end of the day of testing, when fatigue would have been the greatest.

While pain, fatigue, medication, or stress might affect cognitive functioning, poor performance would not be expected on cognitively undemanding and objectively extremely easy SVTs, such as the CARB and WMT. In particular, pain and fatigue would not be expected to cause significantly lower scores in chronic pain patients than in cases of severe traumatic brain injury or neurological disease. Yet, the chronic pain patients in the current study scored significantly lower than the neurological and severe traumatic brain injury group on all of the WMT effort and memory measures, as well as on the CARB. The worse performance of the pain patients than the N-STBI group on the WMT effort and ability measures suggests that their results are implausible. The CARB and

WMT-DR

were failed by less than 2% of the 97 N-STBI patients used as a reference group in this study.

The CARB and WMT effort measures are of approximately equivalent difficulty and very easy. Nevertheless, the pain patients' scores on the WMT effort measures were low. The pain patients' WMT scores did not improve after being warned before the first subtest that the test is extremely easy. This is probably best explained by its design features, which are outlined in the test manual (Green et al., 1996), including the fact that the recognition memory tasks seem far more difficult than they are. When the recognition of 40 words is required after the list of 20 word pairs has been presented twice, it is very hard for the person taking the test to believe that patients with impaired memory would score as high as they do on this measure. The perceived difficulty is an illusion, as shown by the fact that only one percent of the neurological patients and severe traumatic brain injury patients in the current study failed the WMT effort measures. The WMT might also seem much harder than the CARB because word recognition is required after the whole list has been presented and then after a delay of 30 minutes, whereas digit recognition is required after only a few seconds on the CARB.

The WMT was designed with the intention of making it very difficult for a person who does not have memory impairment to produce a profile characteristic of people with genuine impairment, including normal range scores on the effort measures (IR, DR and consistency), but impaired and plausible scores on one or more of the memory measures

(MC, PA, DFR and LDFR; Green et al., 1996). Iverson, Green and Gervais (1999) told very sophisticated subjects, mainly psychologists, that the WMT contains measures of effort and memory. Although each of these sophisticated subjects tried to appear like patients with memory impairment, none were able to produce a valid WMT profile with normal effort scores but impaired memory scores. Their mean scores on the effort measures were below the cutoffs for incomplete effort and far lower than the scores of patients with severe brain injuries. A simple warning, therefore, has a minimal effect on the rate of failure on the WMT effort measures regardless of whether the subjects are chronic pain patients or sophisticated informed simulators.

The current findings are similar to those of Schmand and colleagues (1998), and Vickery et al. (1999), who reported high levels of cognitive exaggeration in chronic pain patients. The current findings support the suggestion of the latter authors that some chronic pain patients involved in disability-related psychological evaluations present with exaggerated cognitive difficulties, producing impaired but invalid scores on objective tests of cognitive abilities.

Exaggeration in one domain creates suspicion about self-reported symptoms. On the other hand, symptom exaggeration or incomplete effort during assessment might occur for a variety of reasons besides malingering. These include a cry for help, marked emotional distress, anger, hostility, factitious disorders, or many other reasons for non-compliance with instructions, all of which are worthy of further systematic investigation. Whereas the results of this study suggest the presence of cognitive exaggeration in some patients, the diagnosis of malingering is not necessarily implied. What is implied is that objective measures of ability are probably invalid in such cases. In people who fail SVTs, the clinical interpretation of invalid test results must rest upon a careful consideration of the entire clinical presentation and corroborating evidence. Whereas SVT failure is a statistical fact, the reason why a person fails SVTs is a clinical judgment.

The current findings support the clinical relevance of previous reports that coaching or “warning” simulators can raise scores on SVTs (Rose, Hall, Szalda-Petree, & Bach, 1998; Dunn et al., 1999) and on other measures of symptom exaggeration (Johnson & Lesniak-Karpiak, 1997). The demonstration of the effects of coaching on SVTs has important implications for work in forensic settings because it has been shown that coaching of litigating patients prior to assessment does occur (Youngjohn, 1995). The existence of differential effects of coaching on the WMT versus the CARB suggests that there are advantages to using SVTs that are relatively resistant to coaching. While the current findings indicate the need for further research to help us to understand biased responding in chronic pain patients, they also indicate the need for more research on the differential effects of coaching on SVTs that are already in widespread clinical use. As the tests become better known, there will be an increase in the potential for coaching and an increased need to employ counter-measures, such as the use of multiple SVTs or coaching-resistant SVTs. The use of SVTs with chronic pain patients holds promise for more accurate and objective assessments of reported symptoms and claimed cognitive impairment, contributing ultimately to a more equitable and effective allocation of rehabilitation and compensation resources.

References

Allen, L., & Green, P. (1999). Severe TBI Sample Performance on CARB and the WMT. (Supplement to the CARB '97 and Word Memory Test Manuals). Durham, NC: CogniSyst.

Allen, L., Conder, R. L., Green, P., & Cox, D. R. (1997). CARB' 97 Manual for the Computerized Assessment of Response Bias. Durham, NC: CogniSyst.

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

Binder, L. M. & Willis, S. C. (1991). Assessment of motivation after financially compensable minor head trauma. Psychological Assessment, 3, 175-181.

Branca, B., Giordani, B., Lutz, T., & Saper, J. R. (1996). Self-report of cognition and objective test performance in posttraumatic headache. Headache, 36, 300-306.

Conder, R., Allen, L., & Cox, D. (1992). Manual for the Computerized Assessment of Response Bias. Durham, NC: CogniSyst.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). California Verbal Learning Test Manual Adult Version, Version 1. San Antonio, TX: The Psychological Corporation.

Di Stefano, G. & Radnov, B. P. (1995). Course of attention and memory after common whiplash: A two-years prospective study with age, education and gender pair-matched patients. Acta Neurologica Scandinavica, 91, 346-352.

Dunn, T. M., Shear, P. K., Howe, S. R., & Ris, M. D. (1999). Effects of coaching and information regarding brain injury on measures designed to detect neuropsychological malingering. Poster session presented at the 27th annual meeting of the International Neuropsychological Society, Boston, MA.

Eccleston, C. (1995). Chronic pain and distraction: An experimental investigation into the role of sustained and shifting attention in the processing of chronic persistent pain. Behavior Research & Therapy, 33, 391-405.

Eccleston, C., Crombez, G., Aldrich, S., & Stannard, C. (1997). Attention and somatic awareness in chronic pain. Pain, 72, 209-215.

Frederick, R. I. (1997). Validity Indicator Profile Manual. Minnetonka, MN: NCS Assessments.

Frederick, R. I., & Crosby, R. D. (2000). Development and validation of the Validity Indicator Profile. Law and Human Behavior, 24, 59-82.

Frederick, R. I., & Foster, H. G. (1991). Multiple measures of malingering on a forced-choice test of cognitive ability. Psychological Assessment, 3, 596-602.

Gervais, R., Green, P., & Allen, L. (1999). Differential sensitivity to symptom exaggeration of verbal, visual, and numerical symptom validity tests [Abstract]. Archives of Clinical Neuropsychology, 14, 746-747.

Green, P., Allen, L., & Astner, K. (1996). The Word Memory Test: A User's Guide to the Oral and Computer-Administered Forms, US Version 1.1. Durham, NC: CogniSyst.

Green, P., & Allen, L., (1999). Performance of Neurological Patients on the Word Memory Test (WMT) and Computerized Assessment of Response Bias (CARB). (Supplement to the Word Memory Test and CARB '97 Manuals). Durham, NC: CogniSyst.

Green, P., & Allen, L., (1997). Memory Complaints Inventory. Durham, NC: CogniSyst.

Green. P., & Iverson, G. L. (In press). Validation of the computerized assessment of response bias in litigating patients with head injuries. The Clinical Neuropsychologist.

Green, P., Iverson, G. L, & Allen, L. (1999). Detecting malingering in head injury litigation with the Word Memory Test. Brain Injury, 13, 813-819.

Grigsby, J., Rosenberg, N. L., & Busenbark, D. (1995). Chronic pain is associated with deficits in information processing. Perceptual and Motor Skills, 81, 403-410.

Hart, R. P., Martelli, M. F., & Zaster, N. D. (2000). Chronic pain and neuropsychological functioning. Neuropsychology Review, 10, 131-149.

Hiscock, M. & Hiscock, C. K. (1989). Refining the forced-choice method for the detection of malingering. Journal of Clinical and Experimental Neuropsychology, 11, 967-974.

Iverson, G. L., & Binder, L. M. (2000). Detecting exaggeration and malingering in neuropsychological assessment. Journal of Head Trauma Rehabilitation, 15, 829-858.

Iverson, G., Green, P., & Gervais, R. (1999). Using the Word Memory Test to detect biased responding in head injury litigation. The Journal of Cognitive Rehabilitation (March/April), 2-6.

Iverson, G. L. & McCracken, L. M. (1997). Postconcussive symptoms in persons with chronic pain. Brain Injury, 11, 783-790.

Johnson, J. L. & Lesniak-Karpiak, K. (1997). The effect of warning on malingering on memory and motor tasks in college samples. Archives of Clinical Neuropsychology, 12, 231-238.

Karlsborg, M., Smed, A., Jespersen, H., Stephensen, S., Cortsen, M., Jennum, P., Herning, M., Korfitsen, E., & Werdelin, L. (1997). A prospective study of 39 patients with whiplash injury. Acta Neurologica Scandinavica, 95, 65-72.

Kay, N. R. M., & Morris-Jones, H. (1998). Pain clinic management of medico-legal litigants. Injury, 29, 305-308.

Kewman, D. G., Vaishampayan, N., Zald, D., & Han, B. (1991). Cognitive impairment in musculoskeletal pain patients. International Journal of Psychiatry in Medicine, 21, 253-262.

Kerns, R. D., Turk, D. C., & Rudy, T. E. (1985). The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). Pain, 23, 345-356.

Lezak, M. (1983). Neuropsychological Assessment (2nd ed.). New York: Oxford University Press.

Pankratz, L. & Binder, L. M. (1997). Malingering on intellectual and neuropsychological measures In R. Rogers (Ed.), Clinical assessment of malingering and deception (2nd ed., pp. 223-236). New York: Guilford.

Reynolds, C. R. (1998). Common sense, clinicians, and actuarialism in the detection of malingering during head injury litigation. In C. R. Reynolds (Ed.), Detection of malingering in head injury litigation (pp. 261-286). New York: Plenum Press.

Rohling, M. L., Binder, L. M., & Langhinrichsen-Rohling, J. (1995). A meta-analytic review of the association between financial compensation and the experience and treatment of chronic pain. Health Psychology, 14, 537-547.

Rose, F. E., Hall, S., Szalda-Petree, A. D., & Bach, P. J. (1998). A comparison of four tests of malingering and the effects of coaching. Archives of Clinical Neuropsychology, 13, 349-363.

Schmand, B., Lindeboom, J., Schagen, S., Heijt, R., Koene, T., & Hamburger, H. L. (1998). Cognitive complaints in patients after whiplash injury: The impact of malingering. Journal of Neurology, Neurosurgery, and Psychiatry, 64, 339-343.

Schnurr, R. & MacDonald, M. (1995). Memory complaints in chronic pain. The Clinical Journal of Pain, 11, 101-111.

Slick, D., Hopp, G., Strauss, E., Hunter, M., & Pinch, D. (1994). Detecting dissimulation: Profiles of simulated malingerers, traumatic brain-injury patients, and normal controls on a revised version of Hiscock and Hiscock's forced choice memory test. Journal of Clinical and Experimental Neuropsychology, 16, 472-481.

Smed, A. (1997). Cognitive function and distress after common whiplash injury. Acta Neurologica Scandinavica, 95, 73-80.

Sweet, J. J. (1999). Malingering: differential diagnosis. In J. J. Sweet, Forensic neuropsychology fundamentals and practice. Lisse: Swets and Zeitlinger.

Taylor, A., Cox, C., & Mailis, A. (1996). Persistent neuropsychological deficits following whiplash: Evidence for chronic mild traumatic brain injury? Archives of Physical Medicine and Rehabilitation, 77, 529-535.

Tombaugh, T. N. (1996). Test of Memory Malingering. Toronto: Multi-Health Systems.

Vickery, C. D., Ranseen, J. D., Cooley, A. T., & Berry, D. T. R. (1999). Use of malingering tests in the forensic evaluation of pain patients [Abstract]. Archives of Clinical Neuropsychology, 14, 102.

Youngjohn, J. R. (1995). Confirmed attorney coaching prior to neuropsychological evaluation. Assessment, 2, 279-283.

Table 1

Memory Complaints Inventory Scores for Chronic Pain and N-STBI Groups

	Pain (n=59)		N-STBI (n=51)		t	p
	Mean	SD	Mean	SD		
GMP*	34.8	25.1	27.0	22.0	-1.69	.094
NIP	38.7	25.0	31.8	23.8	-1.48	.141
VSMP	29.5	25.4	23.7	20.8	-1.29	.200
VMP	46.4	29.1	42.4	29.2	-0.73	.465
PIM	63.7	26.5	28.1	30.7	-6.52	.000
MIW	40.1	30.7	34.3	30.2	-0.99	.324
IRM	20.5	18.1	16.7	14.7	-1.19	.235
ACB	22.1	22.5	19.0	21.0	-0.75	.456
AAB	10.4	15.4	8.1	12.5	-0.85	.398
Mean MCI	34.0	21.7	25.7	19.8	-2.09	.039

* Note. The MCI scales in the table are General Memory Problems (GMP), Numeric Information Problems (NIP), Visuospatial Memory Problems (VSP), Verbal Memory Problems (VMP), Pain Interferes with Memory (PIM), Memory Interferes with Work (MIW), Impairment of Remote Memory (IRM), Amnesia for Complex Behavior (ACB), and Amnesia for Antisocial Behavior (AAB), and mean MCI score.

Table 2

CARB and WMT Effort Measure Scores and Failure Rates

	Pain Patient Groups				
	N-STBI	Uncoached	Test Coached		
			CARB	WMT	p
N	97	57	32	29	
CARB failure rate	2%	41%	6%	38%	.002*
WMT failure rate	1%	42%	34%	41%	.772**
CARB mean % correct	97.8	85.4	97.9	84.8	.000***
(SD)	(3.5)	(17.0)	(4.6)	(21.0)	
WMTIR mean % correct	95.8	83.3	82.1	80.6	.783***
(SD)	(4.5)	(16.6)	(18.7)	(18.4)	
WMTDR mean % correct	95.9	82.5	84.0	80.2	.789***
(SD)	(4.5)	(17.4)	(16.6)	(20.0)	
WMT-CONS mean % correct (SD)	93.1	79.0	79.5	76.5	.868***
	(6.1)	(16.7)	(16.0)	(18.9)	

Note. * Significance of difference between CARB-Coached and Uncoached group. ** Significance of difference between WMT-Coached and Uncoached group. *** Kruskal-Wallis ANOVA comparing Uncoached, CARB-Coached, and WMT-Coached groups.

Table 3

CARB and WMT Scores for 118 DSM-IV Pain Disorder Patients and 97 N-STBI Patients

	N-STBI		All Pain			
	M	SD	M	SD	U	z
CARB	97.8	(3.6)	88.7	(16.8)	3776.0	-4.20 **
WMT-IR	95.8	(4.5)	82.3	(17.5)	2620.5	-6.89 **
WMT-DR	95.9	(4.5)	82.3	(17.8)	2928.5	-6.09 **
WMT-CONS	93.1	(6.1)	78.5	(17.0)	2791.5	-6.35 **
WMT-MC	85.5	(14.5)	65.0	(24.7)	2796.0	-6.01 **
WMT-PA	77.9	(18.7)	59.1	(24.6)	2912.5	-5.42 **
WMT-DFR	44.6	(14.9)	36.0	(17.2)	3436.0	-3.90 **
WMT-LDFR	44.4	(16.5)	36.4	(17.9)	2720.5	-3.40 *

Note. Mean scores are expressed as percent correct. * $p = .001$ ** $p < .0005$