Available online at www.sciencedirect.com
ScienceDirect
Psychiatry Research 157 (2008) 191 – 200
PSYCHIATRY RESEARCH
www.elsevier.com/locate/psychres

Malingering and uncooperativeness in psychiatric and psychological assessment: Prevalence and effects in a German sample of claimants

Andreas Stevens [a,*], Eva Friedel [a], Gisela Mehren [a], Thomas Merten [b]

[a] *Klinik für Psychiatrie und Psychotherapie der Universität Tübingen, Osianderstr. 22, D-72076 Tübingen, Germany*
[b] *Vivantes Netzwerk für Gesundheit, Klinikum im Friedrichshain, Klinik für Neurologie, Berlin, Germany*

Received 6 April 2006; received in revised form 22 December 2006; accepted 5 January 2007

Abstract

Effort has repeatedly been shown to have a pervasive effect on performance in psychological tests. The current study evaluates to what degree performance on various psychological tests is affected by lack of effort as compared with brain injury. Psychological and medical data from a sample of 233 patients referred from Workers' Compensation Boards or from claimants in personal injury litigation were retrospectively analyzed. Each patient underwent a battery of psychological tests and a medical examination. Measures of effort were derived from the Word Memory Test (WMT) and the Medical Symptom Validity Test (MSVT). Insuficient effort was shown by 44.6% of the patients. The frequency of patients failing the effort tests was independent of age, sex, referral source, and leading complaint. Effort accounted for up to 35% of the variance of performance in the domains of cognitive speed, memory and intelligence. After controlling for effort, there was no significant effect that could be attributed to substantial brain injury. The findings confirm that there is a general and strong effect of effort on psychological test results, which dwarfs the impact of substantial brain injury. Effort testing should become a standard procedure in psychological testing.

Keywords: Brain injury; Effort; Psychological testing; Malingering

1. Introduction

Psychological tests are routinely assumed to yield objective and standardized measures of an individual's mental abilities. However, it has long been recognized that test results may be completely invalidated if the patient is not cooperating. Standard psychological tests require good effort to yield valid results. The reason for this is that the reference values that are used to classify a given individual's performance as normal, suboptimal or superior are derived from normative samples composed of persons who perform to the best of their ability. In normative samples, effort is not formally assessed, but subjects participate voluntarily and are often compensated by payment. Thus, they have an interest in performing well and acting in compliance with test instructions. Moreover, they gain no advantage by showing mediocre effort. These assumptions apply neither to clinical settings nor to forensic cases, especially when compensation for some injury is at stake. Thus, in practice, test scores may fall well below published norms, not because of cognitive impairment, but due to lacking cooperation. Uncooperativeness may have many sources: mere lack of interest in taking the test, leading to careless or random responding, fatigue, general distrust of psychological tests, or malevolent

* Corresponding author. Tel.: +49 7071 2982322.
E-mail address: andreas.stevens@med.uni-tuebingen.de (A. Stevens).

0165-1781/$ - see front matter © 2007 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.psychres.2007.01.003

intentions such as fraud. Ideally, a test for cooperativeness would allow us to distinguish between a mere lack of interest and deliberate attempts to manipulate the test.

In a psychiatric examination, the assessment of cooperativeness seems similarly essential to establish psychiatric diagnoses. Most diagnoses rely on the patients' reports about their mental states, which the clinician is unable to validate by objective measurements. Feigning a mental disease is therefore a common problem in forensic as well as civil compensation cases.

The question of how uncooperative or malingering patients may be detected has been given ample treatment (Rogers, 1997; Hall and Poirier, 2001). According to Bender and Rogers (2004), there are several different strategies for detecting feigned impairments on psychological assessments: Floor Effect, Magnitude of Error, Performance Curve, Symptom Validity Testing (SVT), Response Time and Atypical Presentation. The Floor Effect refers to a malingerer's unawareness of those items that are too easy to be plausibly failed. The Magnitude of Error strategy is based on the assumption that feigners will exaggerate the amount by which they miss the correct answer. The Performance Curve strategy examines whether typical performance patterns are preserved, e.g. due to varying difficulty across the test items. Symptom Validity Testing examines whether the failure rate drops below chance levels. Response Time has also been tested as a detection strategy for feigned cognitive impairment.

The prevalence of malingering or uncooperativeness has been estimated at 47% in cases assessed for Workers' Compensation Boards (Youngjohn, 1991). The occurrence of malingered cognitive deficits by accident victims who suffered minor brain injury was estimated at 30% to 40% by Larrabee (2000), and similarly, by Binder (1993). Mittenberg et al. (2002) performed a prospective survey of clinicians, comprising 33,531 cases. There was a diagnosis of probable malingering in 30% of disability claimants, 19% of criminal cases, and 8% of medical cases. Several studies have reported on the magnitude of the effect which effort has on performance in psychological tests (Green et al., 2001; Green, 2004a,b). Large effects were found for memory and learning, psychomotor skills, executive functions and perceptual organization (Green et al., 2001). Gorissen et al. (2005), studying patients with schizophrenia, other psychiatric disorders, neurological illness and normal controls, reported large effects on the California Verbal Learning Test (Delis et al., 1987), Trail Making Test B (TMT-B) (Reitan, 1993) and the Stroop Test (Golden, 1975).

The present study describes a retrospective analysis of $n=233$ cases examined in 2004 and 2005. All of them had suffered an accident and were referred for assessment of medical and psychological impairments in the process of claiming compensation for injuries. All patients claimed to suffer from cognitive impairment, either as the direct result of brain injury or as the result of psychological trauma. All patients underwent a medical and psychological examination. Measures of effort were derived from the Word Memory Test (WMT, Green, 2003; German version as described by Brockhaus and Merten, 2004) and the Medical Symptom Validity Test (MSVT, Green, 2004a,b).

The following two hypotheses were tested: (a) Effort accounts for a considerable part (>20%) of the variance observed in a battery of psychological tests; (b) Certain domains of functioning are more susceptible to the influence of poor effort than others.

2. Methods

2.1. Subjects and instruments

Between March 2004 and June 2005, 233 adult patients underwent a neurological, psychiatric, and psychological examination (Table 1). Most of them were referrals from the German Workers' Compensation Board (49%) or plaintiffs in personal injury claim cases (20%). The category "other" (31%) includes claimants with personal accident insurance and civil servants eligible for accident compensation. For each patient, a detailed description of the accident including technical reports and the initial clinical assessments, including cranial CT scans or MRI findings, were

Table 1
Sample description

Whole sample		n	%	Age (years)	
		233	100	Mean	S.D.
Gender	Male	152	65.2	45.08	13.24
	Female	81	34.8	42.35	13.16
Education	11 years or less	170	72.9		
	More than 11 years	63	27.1		
Substantial brain damage	Absent	155	66.5		
	Present	78	33.4		
Referral source	Workers' Compensation Board	114	48.9		
	Personal injury claim	47	20.2		
	Other	72	31.0		
Complaints	Anxiety	40	17.2		
	Depression	62	26.6		
	Non-syndromatic	59	25.3		
	Medical	201	86.3		
	Whiplash	43	18.5		
	PTSD	63	27		

available. The examination was performed at least 1 year (median = 32 months) after the injury. The symptoms proffered were organized into six complaint categories: anxiety-like (including specific phobias), posttraumatic-stress-disorder-like (PTSD), depression-like, non-syndromal psychiatric complaints (including forgetfulness, irritability, fatigue), medical complaints (pain, dizziness, palsy, gastro-intestinal-complaints) and whiplash injury (complaints about neck pain and shoulder stiffness with cognitive complaints after a vehicle accident).

A detailed medical history was compiled, as well as a general medical, a neurological, and a psychiatric examination. Psychiatric diagnoses were made according to the DSM-IV-TR (American Psychiatric Association, 2000). Each patient had an EEG and, in cases of peripheral nerve injury, a neurophysiological assessment (electromyography, somatosensory evoked potentials). The results of the medical examination and the neurological diagnoses will not be given in detail, but are summarily reported as presence or absence of accident-related abnormal medical findings. Initial Glasgow Coma Scale ratings were unavailable in many cases; however, there were always neurological reports and cranial CT scans or MRI available. Based on these radiological findings, the cases were organized into two groups: the group with "substantial brain injury absent" includes all patients who either suffered no head injury or whose cranial imaging showed neither brain swelling nor hemorrhage. This category corresponds to absent, mild, or moderate brain injury in the standard nomenclature. The group with "substantial brain injury present" comprises all patients with radiological evidence of traumatic brain injury (hemorrhage, brain swelling, axonal damage). Each patient underwent a psychological test battery. A flexible approach was followed, depending on the questions the assessment was to answer. Because not every patient received the same test battery, the multivariate analyses are limited to subsets of cases. The assessment, including the medical examination, took 1 day. The psychological test battery was administered to each patient individually by a skilled psychologist. Each patient had given written informed consent to the anonymous use of his/her data.

The following psychological tests were used: to measure both effort and verbal memory, all patients were given either the computerized Word Memory Test (WMT; Green, 2003) or the computerized Medical Symptom Validity Test (MSVT) (Green, 2004a,b). The WMT is one of the most prominent SVTs available. It has repeatedly demonstrated its validity for diagnosing suboptimal performance (e.g., Gervais et al., 2004; Green et al., 1999; Green et al., 2001; for a review of the test: Green et al., 2002, and, more recently, Wynkoop and Denney, 2005) and was rated favorably in a comparison of different instruments (Hartman, 2002). On a computer screen, a list of 20 word pairs is presented twice to the patient. After that, the computer presents word pairs containing one of these words and one that had not been shown. The subject is required to select the word that was shown previously in the original list. Thus, a total of 40 test items are produced on the Immediate Recognition trial (IR). After a delay of 30 min, the same recognition testing is performed again, using different foil words in the Delayed Recognition trial (DR). The third effort variable is consistency between IR and DR. Some additional sub-tests were not included in the analyses so they need no further description here. A number of simulation studies have shown high sensitivity and specificity with correct classification rates of 99–100% in the studies of Tan et al. (2002), Brockhaus and Merten (2004) and Brockhaus and Peker (2003).

The MSVT is derived from the original WMT but comprises only 10 word pairs and, consequently, 20 test items each in the IR and DR trials. Moreover, the target word pairs are strongly associated with each other in an effort to facilitate the learning process.

Both tests are introduced to the patient as verbal memory tests, while, in fact, they are designed primarily to measure test motivation. The task is much easier than it appears and makes use of the Floor Effect; it cannot plausibly be failed unless there is bona-fide dementia or moderate to severe aphasia. According to the number of correct responses, both tests yield measures of effort as continuous variables, which are used to classify the patients on grounds of empirically derived cut-off values as "Fail" (showing insufficient effort) or "Pass" (showing sufficient effort).

The Structured Inventory of Malingered Symptomatology (SIMS) (Smith and Burger, 1997; German adaptation by Cima et al., 2003) is a paper–pencil self-report scale, which indicates symptom exaggeration. It contains 75 dichotomous items, describing a variety of symptoms, some of which are extremely unlikely to occur in real disorders but seem plausible to claimants with a tendency for over-reporting of symptoms. Patients are classified as "probably malingering" or "probably not malingering", according to empirically derived cut-off values. The following sub-tests of the Wechsler Intelligence Test for Adults (WAIS-R, Wechsler, 1981; German version by Tewes, 1991) were used: Object Assembly, Digit Span and Digit Symbol. Premorbid intelligence was estimated by a German multiple-choice Spot-the-Word test (Lehrl et al., 1995).

Table 2
General descriptors of subjects who fail and those who pass the effort tests

		Pass		Fail		χ^2
		n	%	*n*	%	
Whole sample		129	55.4	104	44.6	
Gender	Male	81	53.3	71	46.7	7.62
	Female	48	59.3	33	40.7	
Education	11 years or less	85	50.0	85	50.0	7.32 **
	More than 11 years	44	69.8	19	30.2	
Substantial brain damage	Absent	87	56.1	68	43.9	0.16
	Present	42	53.8	36	46.2	
Referral source	Compensation Board	58	50.9	56	49.1	1.82
	Personal injury claim	28	59.6	19	40.4	
	Other	43	59.7	29	40.3	
Complaints *	Anxiety	19	47.5	21	52.5	1.21
	Depression	35	56.5	27	43.5	0.40
	Non-syndromatic	28	47.5	31	52.5	2.00
	Medical	110	54.7	91	45.3	0.24
	Whiplash injury	22	51.2	21	48.8	0.00
	PTSD	35	55.6	28	44.4	0.37

* Multiple complaints allowed.
** $P<0.01$.

Verbal memory was assessed with the sub-tests Logic Memory Immediate and Delayed Recall as well as Paired Associates Learning from the Wechsler Memory Scale-Revised (WMS-R, Wechsler, 1987; German version by Härting et al., 1991). Visual memory was tested with the sub-test Visual Memory from the Visual and Verbal Memory Test (Schelling and Schächtele, 2001), a paper–pencil test. Attention was evaluated with the Attentional Network Test (ANT, Gauggel and Böcker, 2003). This is a computerized test measuring reaction time (ms) in a forced choice paradigm. It also evaluates responsiveness to cues, distracters, and alerting signals. Executive functions were assessed with the Trail Making Test B (Reitan, 1993), also a paper–pencil test.

Table 3
Structured Inventory of Malingered Symptomatology (SIMS) versus effort

		WMT pass		WMT fail		χ^2
		n	%	*n*	%	
SIMS	Pass	73	73.7	26	26.3	26.3 **
SIMS	Fail	23	33.8	45	66.2	

** $P<0.01$.

2.2. Statistical design

Contingency tabulation was used to test whether "Fail" and "Pass" patients were distributed randomly with respect to demographic descriptors. Hypothesis (a) was evaluated using three MANOVAs (a, b, c), one for each cognitive domain: "intelligence" (WAIS-R Vocabulary, Object Assembly, Digit Symbol, Digit Span and Visual Memory), "verbal memory" (WMS-R Immediate Recall, Delayed Recall) and "executive functions" (TMT-A, TMT-B, ANT Alertness, Orienting, Conflict, Reaction Time). The dependent variables consisted of performance on the individual tests (*z*-scores, derived from the normative tables provided for each test). The between-subject factors were effort ("Fail" and "Pass") and substantial brain injury ("none" and "yes"). For the ANT and TMT, there are no age-corrected norms. Therefore, raw data (ms) were used, including age and years of education as covariates. Homogeneity of error variance was examined using Levene's test, only failed by WAIS-R Digit Span data. All dependent variables passed tests for approximate normal distribution, with

Table 4
Effects of effort and of brain injury on test performance

	F	Eta^2	Power
(a) Intelligence sub-tests and visual memory			
(WAIS-R Vocabulary, Object Assembly, Digit Symbol, Digit Span and Visual Memory Test)			
n substantial brain injury = 44, *n* no brain injury = 74			
Substantial brain injury	$F(5,110)=1.33$	0.06	0.77
Effort	$F(5,110)=4.54$ **	0.17	
Brain injury × effort	$F(5,110)=0.54$	0.03	
(b) Verbal memory			
(WMS-R Pair Associates Immediate Recall, WMS-R Delayed Recall)			
n substantial brain injury = 36, *n* no brain injury = 49			
Substantial brain injury	$F(2,80)=0.26$	0.01	0.98
Effort	$F(2,80)=4.73$ *	0.11	
Brain injury × effort	$F(2,80)=1.58$	0.04	
(c) Executive function and attention			
(TMT-A, TMT-B, ANT Alertness, Orienting, Conflict, Reaction Time)			
n substantial brain injury = 18, *n* no brain injury = 21			
Substantial brain injury	$F(6,29)=1.70$	0.26	0.72
Effort	$F(6,29)=2.58$ *	0.35	
Brain injury × effort	$F(6,29)=1.05$	0.18	
Education years	$F(6,29)=1.29$	0.21	
Age	$F(6,29)=1.97$	0.29	

* $P<0.05$.
** $P<0.01$.

Table 5
Effects of effort on individual tests

Test	F(brain injury)	F(effort)	Cohen's d[a]
(a) Intelligence sub-tests and visual memory			
Vocabulary	$F(1,114)=0.03$	$F(1,114)=11.0$ **	0.62
Object Assembly	$F(1,114)=5.02$ *	$F(1,114)=7.62$ **	2.66
Digit Symbol	$F(1,114)=0.00$	$F(1,114)=15.98$ **	5.01
Digit Span	$F(1,114).=0.19$	$F(1,114)=8.48$ **	2.43
Visual Memory	$F(1,114)=0.21$	$F(1,114)=6.63$ *	3.94
(b) Verbal memory (pair associates)			
WMS-R Immediate Recall	$F(1,81)=0.00$	$F(1,81)=8.35$ **	3.51
WMS-R Delayed Recall	$F(1,81)=0.20$	$F(1,81)=8.99$ **	3.29
(c) Executive functions and attention			
TMT-A	$F(1,34)=1.49$	$F(1,34)=1.72$	−1.89
TMT-B	$F(1,34)=0.03$	$F(1,34)=0.26$	−0.73
ANT Alertness	$F(1,34)=0.83$	$F(1,34)=0.76$	1.25
ANT Orienting	$F(1,34)=1.17$	$F(1,34)=0.66$	−1.17
ANT Conflict	$F(1,34)=0.54$	$F(1,34)=2.04$	−2.03
ANT Reaction Time	$F(1,34)=1.33$	$F(1,34)=9.82$ **	−4.51

[a] Cohen's *d* calculated as [mean(pass) − mean(fail)]/pooled S.D.
* $P<0.05$.
** $P<0.01$.

the exception of ANT reaction time (this was expected since reaction times do not follow a normal distribution). Hypothesis (b) was evaluated by planned post-hoc comparisons of the sub-samples "Fail" and "Pass". Significance was assumed for $P<0.01$. The probability of type II errors was estimated by power analyses. The statistical software was SPSS for Windows (SPSS Inc., Chicago, Ill.).

3. Results

Of the 233 patients, 104 (44.8%) were classified as "Fail" according to the effort measures, and 129 (55.4%) as "Pass". Table 2 shows the distribution of "Fail" and "Pass" according to demographic descriptors. The likelihood to be classified as "Fail" was associated neither with gender, nor with presence or absence of substantial brain injury, referral source, or complaint. However, patients with more than 11 years of education were less likely to fail the effort tests ($\chi^2=7.32$, $df=1$, $P<0.01$). The classification results from the SIMS ("probably malingering" vs. "not malingering") were significantly associated with the WMT/MSVT classification ($\chi^2=26.28$, $df=1$, $P<0.01$), but not identical (Table 3): 24% of those passing the SIMS failed the WMT/MSVT; 36.9% of those failing the WMT/MSVT passed the SIMS. Cohen's kappa was 0.40 ($P<0.01$).

The MANOVAs (a, b, c) indicated a significant and strong effect for effort on test performance in each cognitive domain (Table 4). Effect sizes are indicated as partial eta^2, which is a conservative measure corrected for the effect of other factors in the model. The effect for substantial brain injury on test performance did not reach significance. The absence of a significant effect for brain injury is probably a substantial result, since the likelihood for a type II error was low ($0.02<P<0.28$).

Since the main effects were significant, post-hoc analyses were carried out (Table 5). Susceptibility for

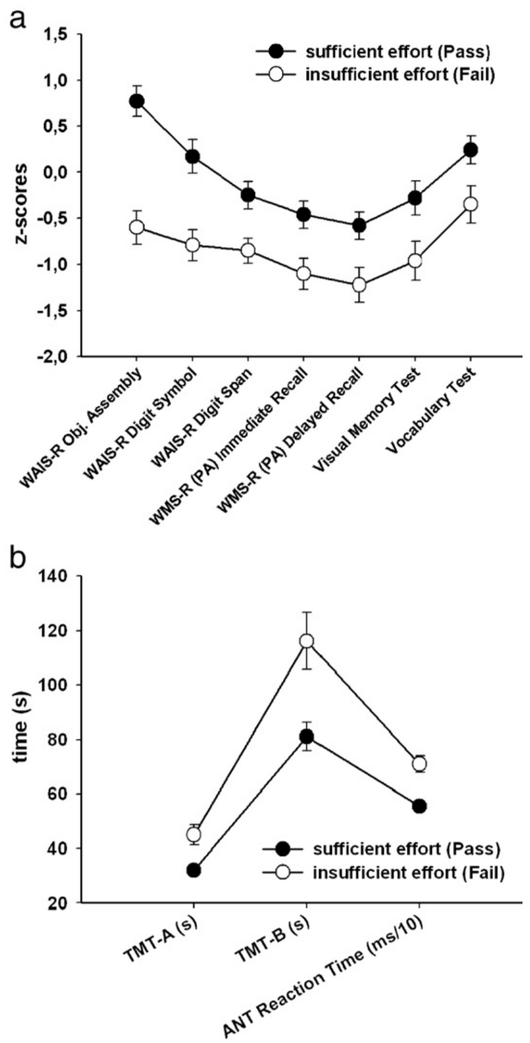

Fig. 1. The graph compares the performance of patients passing the effort test with the performance of those failing. The dots indicate average performance of the respective group, the bars SEM. Fig. 1a renders tests for which *z*-scores are available, Fig. 1b represents reaction times and time required for completing the Trail Making Test.

Table 6
Correlation of site of substantial brain injury with effort

Site		n (100%)	Pass		Fail	
			n	%	n	%
Frontal	Left	36	18	50	18	50
	Right	33	18	54	15	46
Temporal	Left	15	7	52	8	48
	Right	12	5	42	7	58
Parietal	Left	13	8	61	5	39
	Right	13	9	69	4	31
Occipital	Left	3	1	–	2	–
	Right	0	0	–	0	–

Note that there are multiple entries for patients with multiple brain lesions.

effort was neither evident for TMT-A and TMT-B nor for the ANT sub-tests Alertness, Conflict, and Orienting, while Reaction Time was markedly prolonged by lack of effort. Pronounced effects of effort on test performance were observed for all other tests. The post-hoc comparisons indicated that only WAIS-R Object Assembly performance was associated with the factor "substantial brain injury". However, the effect turned out to be contrary to what was expected: patients with substantial brain injury performed significantly better than those without (mean percentiles 32.00±1.37 as compared with 27.51±0.97, Cohen's $d=-3.84$). As expected, there was a significant effect of age on TMT-A completion time ($F=8.97$, $df=1$, $P<0.01$). However, as the combination of tests applied varied from subject to subject, the degree of overlap between the MANOVA subsets was low (50% for MANOVAs (a) and (b) and 54% for MANOVAs (a) and (c)). Fig. 1 shows the effects of effort by comparing test performances (as percentile) of the groups "Fail" and "Pass".

The site of brain injury might influence not only performance on certain cognitive tasks, but also the level of motivation. Contingency tabulation with the factor effort ("Pass" vs. "Fail") and the site of brain injury, however, gave no significant deviation from chance levels for any of the eight brain regions tested (Table 6) (all chi^2 $P \gg 0.05$). Since the sample size for some brain regions was very small, the analysis may be of insufficient power.

In order to estimate how effort relates to performance in individual tests, scatterplots were inspected (Fig. 2 shows two examples). The graphical representations suggested a unidirectional linear association of effort with test performance. "Unidirectional" means that with diminishing effort, there is a linear decline in test performance, while test performance does not predict effort. Spearman rank correlation coefficients indicated

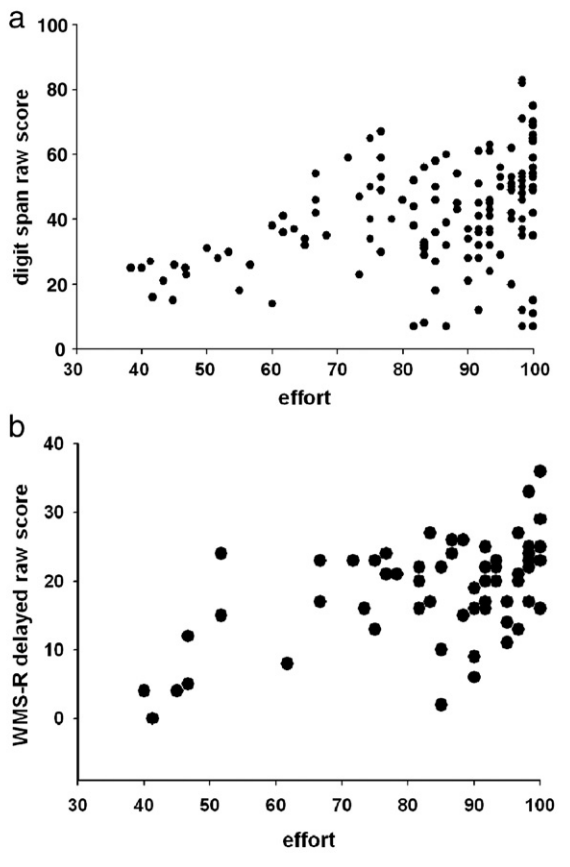

Fig. 2. The scatterplots suggest a unidirectional association between WMT/MACT effort scores and test performance. As examples, data for Digit Span (a) and WMS-R Delayed Recall (b) are shown. The dots represent individual patients. Note that reduced effort seems to cut out the upper range of performance.

Table 7
Bivariate correlations of effort with individual tests

Test	Spearman's rho	n
Vocabulary	0.28 *	115
WAIS-R sub-tests		
Object Assembly	0.24 **	138
Digit Symbol	0.39 **	134
Digit Span	0.42 **	129
Visual Memory	0.45 **	124
WMS-R Immediate Recall	0.40 **	63
WMS-R Delayed Recall	0.39 **	62
TMT-A	−0.49 **	54
TMT-B	−0.45 **	53
ANT sub-tests		
Alertness	0.13	54
Orienting	−0.06	54
Conflict	−0.25	54
Reaction Time	−0.69 **	54

* $P<0.05$.
** $P<0.01$.

significant associations between effort and performance on all psychological tests, except for the ANT sub-tests Alertness, Orienting and Conflict. The correlations ranged from rho = 0.28 for the vocabulary test to rho = −0.69 for the reaction time of the ANT (Table 7).

4. Discussion

The analysis of psychological test data from accident victims confirmed both initial hypotheses: (a) effort accounts for a considerable portion of the variance observed in a battery of psychological tests, and (b) effort has differential effects on performance in individual tests. A collateral result was that the effects of effort are larger than those of substantial brain injury.

In the present study, 44.8% of patients showed insufficient effort. This figure is well within the range reported by Youngjohn's (1991) study of claimants, and also comparable to the figures reported by Larrabee (2000), Binder (1993), and Mittenberg et al. (2002). In the present study, the prevalence of uncooperativeness was not associated with the leading complaint: the rate of uncooperativeness did not vary significantly across subgroups with whiplash injury, PTSD-like or non-syndromal complaints. However, Schmand et al. (1998) reported a particularly high prevalence of malingering (61%) in whiplash injury claimants. A high prevalence of malingering has also been observed in patients complaining of fibromyalgia and chronic fatigue (van der Werf et al., 2000; Gervais et al., 2001a,b). In the present study, uncooperativeness was associated neither with gender nor with age or referral source. Only education seemed to predict effort. The notion that persons with higher education seem less likely to malinger is not easy to account for. Possibly, better educated patients are less likely to exaggerate their symptoms, as they are also less likely to be diagnosed with somatoform disorder (Lieb et al., 2002). An alternative explanation would be that with well-educated patients, more subtle effort tests might be required.

There was moderate agreement between the SIMS and the WMT/MSVT classification of cooperativeness (Cohen's kappa = 0.40). To our knowledge, there is no other study commenting on agreement of SIMS ratings with other tests for malingering. The moderate agreement may partly be accounted for by methodological differences: the SIMS is a self-report inventory, while the WMT requires active behavior. Another difference is that the SIMS assesses a wide range of cognitive and affective complaints, while the WMT is presented as a memory test. Thus, patients complaining of psychological distress might not necessarily feign memory impairment.

In the present study, the proportion of variance explained by the measures of effort, derived from the WMT, was smaller (20–30%) than in previous studies (30–50%, Green et al., 2001). However, the estimate used here (partial eta^2) is rather conservative and excludes the influence attributed to other factors in the model. In agreement with Green et al. (2001), the effects of effort were much more distinct than those associated with presence or absence of substantial brain injury.

However, caution has to be advised: damage to some brain areas might compromise motivation, especially injury to frontal areas, which are thought to mediate executive functions. In fact, Gorissen et al. (2005) reported a high rate of non-compensation-seeking patients with DSM-IV-TR schizophrenia failing the WMT (72%). In their study, insufficient effort was correlated to negative symptoms. Considering that anhedonia and lack of interest are among the main symptoms of schizophrenia, the finding is not surprising. Patton et al. (2004) also used the WMT and found low effort in a sample of patients with cortical or subcortical dementia, apparently "false positives". In the present study, we found no evidence that injury to some brain regions (especially the frontal lobes) was associated with an increased probability to fail the effort tests. On the other hand, this lack of evidence must be viewed with skepticism, because the sample size was rather small for some brain regions. Similarly, Green et al. (1999) saw regular or superior effort in a sample with severe brain injury. These findings suggest that effort is usually not impaired by brain injury. The impact of depression-related symptoms on effort was not evaluated in the present study. In theory, depression might reduce motivation. However, Yanez et al. (2006) have reported that subjects with DSM-IV major depression show normal effort scores.

The psychological tests clearly differed in their susceptibility to lack of effort. WAIS-R Digit Symbol, Visual Memory, and WMS-R, as well as ANT Reaction Time, were strongly affected, while TMT-A and TMT-B, WAIS-R Digit Span and Object Assembly were more robust. However, even for the latter tests, Cohen's *d* for effort ranged from 0.6 to 2.5, indicating medium to large effects. Similar results were obtained by Green et al. (2001). The scatterplots suggest that those who failed the effort test "never did well" in any test, while those who passed it showed a normal range of test performance in other tests. Similarly, Constantinou et al. (2005) found that poor performance in a recognition

Table 8

Test	Substantial brain injury	Effort	Mean	S.D.
Descriptive statistics (mean z-scores and standard deviation) for MANOVA (a) intelligence sub-tests and visual memory				
Vocabulary	Absent	Pass	0.34	1.03
		Fail	−0.57	1.20
	Present	Pass	0.08	0.87
		Fail	−0.37	1.19
WAIS-R Obj. Assembly	Absent	Pass	0.05	1.04
		Fail	−0.55	1.04
	Present	Pass	0.40	0.85
		Fail	−0.04	0.92
WAIS-R Digit Symbol	Absent	Pass	0.28	1.17
		Fail	−0.63	1.08
	Present	Pass	0.25	1.28
		Fail	−0.60	0.96
WAIS-R Digit Span	Absent	Pass	−0.05	0.83
		Fail	−0.71	0.58
	Present	Pass	−0.15	1.10
		Fail	−0.47	0.91
Visual Memory	Absent	Pass	−0.22	1.18
		Fail	−0.99	1.30
	Present	Pass	−0.29	1.10
		Fail	−0.71	1.15
Descriptive statistics (mean z-scores and standard deviation) for MANOVA (b) verbal memory				
WMS-R Paired Associates Immediate Recall	Absent	Pass	−0.52	0.85
		Fail	−0.82	0.96
	Present	Pass	−0.20	1.00
		Fail	−1.18	1.08
WMS-R Paired Associates Delayed Recall	Absent	Pass	−0.71	1.02
		Fail	−1.15	0.93
	Present	Pass	−0.35	0.83
		Fail	−1.30	1.32
Descriptive statistics (time in s and ms) for MANOVA (c) executive functions and attention				
TMT-A (s)	Absent	Pass	41.2	19.0
		Fail	49.0	13.5
	Present	Pass	29.8	8.2
		Fail	40.8	15.2
TMT-B (s)	Absent	Pass	110.4	32.4
		Fail	105.6	39.3
	Present	Pass	75.2	30.4
		Fail	113.2	51.2
ANT Alertness (ms)	Absent	Pass	39.6	37.4
		Fail	7.0	61.7
	Present	Pass	61.4	49.4
		Fail	37.5	38.9
ANT Orienting (ms)	Absent	Pass	20.2	24.6
		Fail	72.3	27.7
	Present	Pass	21.7	30.1
		Fail	16.0	78.0
ANT Conflict (ms)	Absent	Pass	56.0	15.4
		Fail	92.6	119.1
	Present	Pass	79.8	37.2
		Fail	147.7	155.8
ANT RT (ms)	Absent	Pass	612.2	150.6
		Fail	808.6	243.0
ANT RT (ms)	Present	Pass	560.0	130.1
		Fail	710.7	159.2

memory test was associated with poor test performance in general.

An interesting but preliminary result was that only reaction time but neither the orienting nor the alerting and conflict responses of the ANT were modified by effort. The most likely interpretation seems to be that motor responses may be voluntarily delayed, but not so the networks engaged in the processing of distracters, cues and alerting signals. This hypothesis is supported by Collum et al. (1991), who showed that it is much more feasible to fake impairment of simple motor and cognitive performance than that of more complex tasks.

One might expect that the sub-sample of subjects who pass the effort tests and who did not suffer substantial brain damage will show normal testing results. Table 8 represents the testing results separately for the “Pass/Fail” and “brain injury absent/present” sub-samples. It turns out that even in the “Pass” sub-sample without radiological evidence for substantial brain injury, performance is subnormal for verbal and visual memory, reaction time and cognitive speed. There are several possible explanations. First, consider the possibility that someone with good cognitive abilities performs with reduced effort but is not exposed by the effort tests. The cut-off value for the “Fail/Pass” dichotomy is rather conservative. Thus, among those passing the effort tests, there will still be some that do not perform with good effort. Evidence for this is provided by the positive correlation of effort with test performance, showing that test performance still rises beyond the 82.5 percentile (the cut-off) of the WMT/MACT. Another possibility is that a subject may perform with good effort on the WMT but with less effort on other tests. A third alternative would assume that the “Fail/Pass” classification is correct but that there are real cognitive impairments even in those with non-substantial brain injury. Iverson (2005) has reviewed the literature about the long-term prospects of mild traumatic brain injury (MTBI). According to some textbooks, 10–20% of subjects with MTBI appear not to recover fully. However, he concludes, this may be an artifact created by substance abuse, litigation, depression or emotional and vocational stressors. MTBI effects seem to explain only a small share of the variance of cognitive function. Applied to our data, this would imply that there are some subjects without substantial brain injury but with real cognitive impairments. For example, we found that self-report of drug and alcohol abuse was unreliable in 30% of tests when cross-checked with hospital data files.

Some limitations of the present study need to be addressed: the most crucial issue pertains to the classification accuracy of the effort tests. For the tests employed in the current study, classification accuracy has been estimated above 95%, according to data from healthy controls and patients instructed to fake cognitive impairment (Green et al., 2002; Iverson et al., 2002; Tan et al., 2002; Brockhaus and Merten, 2004; Tydecks et al., 2006). Another important issue is that the data presented here stem from a retrospective analysis of data gathered for forensic purposes and are heterogeneous with respect to the psychological tests applied to each individual. For the same reason, the size of the analyzed subsets was quite heterogeneous and the degree of overlap was low. The severity of brain injury was classified on the basis of radiological data. This was necessary because Glasgow Coma Scale ratings and duration of consciousness were unavailable or found to be unreliable in 30% of cases, when cross-checked against the hospital files. However, the radiological classification might render the findings of the present study difficult to compare with reports using clinical data.

To summarize, effort testing should be an essential component of psychological testing. The reasons for this recommendation are twofold: first, the effects of uncooperativeness are strong and affect many psychological tests. Second, the effects of uncooperativeness (lack of effort) are much stronger than the effects of substantial brain injury.

References

American Psychiatric Association, 2000. APA Diagnostic and Statistical Manual of Mental Disorders, DSM-IV Text Revision. APA, Washington, DC.

Bender, S.D., Rogers, R., 2004. Detection of neurocognitive feigning: development of a multi-strategy assessment. Archives of Clinical Neuropsychology 19, 49–60.

Binder, L.M., 1993. Assessment of malingering after mild head trauma with the Portland Digit Recognition Test. Journal of Clinical and Experimental Neuropsychology 15, 170–182.

Brockhaus, R., Merten, T., 2004. Neuropsychologische Diagnostik suboptimalen Leistungsverhaltens mit dem Word Memory Test. Nervenarzt 75, 882–887.

Brockhaus, R., Peker, Ö., 2003. Testing effort in Turkish-speaking subjects: validation of a translation of the Word Memory Test (WMT). Journal of the International Neuropsychological Society 9, 520 [Abstract].

Cima, M., Hollnack, S., Kremer, K., Knauer, E., Schellbach-Matties, R., Klein, B., Merckelbach, H., 2003. Strukturierter Fragebogen Simulierter Symptome. Die deutsche Version des “Structured Inventory of Malingered Symptomatology: SIMS”. Nervenarzt 74, 977–986.

Collum, C.M., Heaton, R.K., Grant, I., 1991. Psychogenic factors influencing neuropsychological performance. In: Doerr, H.O., Carlin, A.S. (Eds.), Forensic Neuropsychology: Legal and Scientific Bases. The Guilford Press, New York, pp. 141–171.

Constantinou, M., Bauer, L., Ashendorf, L., Fisher, J.M., McCaffrey, R.J., 2005. Is poor performance on recognition memory effort measures indicative of generalized poor performance on neuropsychological tests? Archives of Clinical Neuropsychology 20, 191–198.

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

Gauggel, S., Böcker, M., 2003. Der Aufmerksamkeits-Netzwerk-Test. Neuropsychologisches Zentrum Kassel, Germany.

Gervais, R.O., Russell, A.S., Green, P., Allen, L.M., Ferrari, R., Pieschl, S.D., 2001a. Effort testing in patients with fibromyalgia and disability incentives. Journal of Rheumatology 28, 1892–1899.

Gervais, R.O., Green, P., Allen, L.M., Iverson, G.L., 2001b. Effects of coaching on symptom validity testing in chronic pain patients presenting for disability assessment. Journal of Forensic Neuropsychology 2, 1–20.

Gervais, R.O., Rohling, M.L., Green, P., Ford, W., 2004. A comparison of WMT, CARB, and TOMM failure rates in non-head injury disability claimants. Archives of Clinical Neuropsychology 19, 475–487.

Golden, C.J., 1975. The measurement of creativity by the Stroop Color and Word Test. Journal of Personality Assessment 39, 502–506.

Gorissen, M., Sanz, J.S., Schmand, B., 2005. Effort and cognition in schizophrenia patients. Schizophrenia Research 78, 199–208.

Green, P., 2003. Green’s Word Memory Test. User’s Manual. Green’s Publishing, Edmonton.

Green, P., 2004a. Green’s Medical Symptom Validity Test (MSVT) for Microsoft Windows. User’s Manual. Green’s Publishing, Edmonton, Canada.

Green, P., 2004b. The Pervasive Influence of Effort on Neuropsychological Tests. Green’s Publishing, Edmonton.

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

Green, P., Rohling, M.L., Lees-Haley, P.R., Allen, L.M., 2001. Effort has a greater effect on test scores than severe brain injury in compensation claimants. Brain Injury 15, 1045–1060.

Green, P., Lees-Haley, P.R., Allen, L.M., 2002. The Word Memory Test and the validity of neuropsychological test scores. Journal of Forensic Neuropsychology 2, 97–124.

Hall, H.V., Poirier, J.G., 2001. Detecting Malingering and Deception. Pacific Institute Series on Forensic Psychology, CRC Press.

Härting, C., Markowitsch, H.J., Neufeld, H., Calabrese, P., Deisinger, K. (Hrsg), 1991. Wechsler Gedächtnis Test-Revidierte Fassung. Deutsche Adaptation der revidierten Fassung der Wechsler-Memory-Scale,1. Auflage. Hogrefe Verlag.

Hartman, D.E., 2002. The unexamined lie is a lie worth fibbing. Neuropsychological malingering and the Word Memory Test. Archives of Clinical Neuropsychology 17, 709–714.

Iverson, G.L., 2005. Outcome from mild traumatic brain injury. Current Opinion in Psychiatry 18, 301–317.

Iverson, G.L., Lange, R.T., Green, P., Franzen, M.D., 2002. Detecting exaggeration and malingering with the Trail Making Test. The Clinical Neuropsychologist 16, 398–406.

Larrabee, G.J., 2000. Neuropsychology in personal injury litigation. Journal of Clinical and Experimental Neuropsychology 22, 702–707.

Lehrl, S., Triebig, G., Fischer, B., 1995. Multiple choice vocabulary test MWT as a valid and short test to estimate premorbid intelligence. Acta Neurologica Scandinavica 91, 335–345.

Lieb, R., Zimmermann, P., Friis, R.H., Höfler, M., Tholen, S., Wittchen, H.U., 2002. The natural course of DSM-IV somatoform disorder and syndromes among adolescents and young adults: a prospective-longitudinal community study. European Psychiatry 17, 321–331.

Mittenberg, W., Patton, C., Canyock, E.M., Condit, D.C., 2002. Base rates of malingering and symptom exaggeration. Journal of Clinical and Experimental Neuropsychology 24, 1094–1102.

Patton, C., Mittenberg, W., Loewenstein, D., Roberts, D., 2004. Validation of the Oral Word Memory Test in Dementia and Psychiatric Disorders. National Academy of Neuropsychology 24th Annual Conference, Nov 17–20, 2004.

Reitan, R.M., 1993. Trail Making Test. Reitan Neuropsychology Laboratories, Tucson, AZ.

Rogers, R., 1997. Introduction. In: Rogers, R. (Ed.), Clinical Assessment of Malingering and Deception, 2nd ed. The Guilford Press, New York, pp. 1–19.

Schelling, D., Schächtele, B., Hrsg. 2001. VVM. Visueller und Verbaler Merkfähigkeitstest. Hogrefe Verlag.

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.

Smith, G.P., Burger, G.K., 1997. Detection of malingering: validation of the Structured Interview of Malingered Symptomatology (SIMS). Journal of the American Academy of Psychiatry and the Law 25, 183–189.

Tan, J.E., Slick, D.J., Strauss, E., Hultsch, D.F., 2002. How'd they do it? Malingering strategies on symptom validity tests. The Clinical Neuropsychologist 16, 495–505.

Tewes, U., 1991. Hamburg-Wechsler-Intelligenztest für Erwachsene, Revision 1991. Huber, Bern.

Tydecks, S., Merten, T., Gubbay, J., 2006. The Word Memory Test and the One-in-Five-Test in an analogue study with Russian-speaking participants. International Journal of Forensic Psychology 1, 29–37.

van der Werf, S.P., Prins, J.B., Jongen, P.J., van der Meer, J.W., Bleijenberg, G., 2000. Abnormal neuropsychological findings are not necessarily a sign of cerebral impairment: a matched comparison between chronic fatigue syndrome and multiple sclerosis. Neuropsychiatry, Neuropsychology and Behavioral Neurology 13, 199–203.

Wechsler, D., 1981. Wechsler Adult Intelligence Scale-Revised. Psychological Corporation, San Antonio, TX.

Wechsler, D., 1987. Wechsler Memory Scale Revisited, The Psychological Corporation. Harcourt Brace Jovanovich Inc., San Antonio.

Wynkoop, T.F., Denney, R.L., 2005. Test review: Green's Word Memory Test (WMT) for Windows. Journal of Forensic Neuropsychology 4, 101–105.

Yanez, Y.T., Fremouw, W., Tennant, J., Strunk, J., Coker, K., 2006. Effects of severe depression on TOMM performance among disability-seeking outpatients. Archives of Clinical Neuropsychology 21, 161–165.

Youngjohn, J.R., 1991. Malingering of neuropsychological impairment: an assessment strategy. A Journal for the Expert Witness, The Trial Attorney, The Trial Judge 4, 29–32.