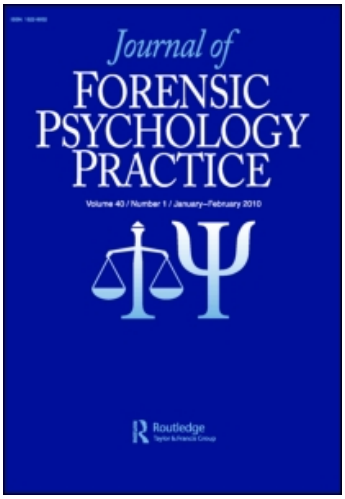

Journal of Forensic Psychology Practice
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t792304004

Detection of Feigned Crime-Related Amnesia: A Multi-Method Approach

Peter Giger[a]; Thomas Merten[b]; Harald Merckelbach[c]; Margit Oswald[a]
[a] Department of Psychology, University of Berne, Berne, Switzerland [b] Department of Neurology, Vivantes Klinikum im Friedrichshain, Berlin, Germany [c] Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands

Online publication date: 05 November 2010

To cite this Article Giger, Peter , Merten, Thomas , Merckelbach, Harald and Oswald, Margit(2010) 'Detection of Feigned Crime-Related Amnesia: A Multi-Method Approach', Journal of Forensic Psychology Practice, 10: 5, 440 — 463
To link to this Article: DOI: 10.1080/15228932.2010.489875
URL: http://dx.doi.org/10.1080/15228932.2010.489875

Journal of Forensic Psychology Practice, 10:440–463, 2010

ISSN: 1522-8932 print/1522-9092 online
DOI: 10.1080/15228932.2010.489875

Routledge
Taylor & Francis Group

Detection of Feigned Crime-Related Amnesia: A Multi-Method Approach

PETER GIGER, MSc
Department of Psychology, University of Berne, Berne, Switzerland

THOMAS MERTEN, PhD
Department of Neurology, Vivantes Klinikum im Friedrichshain, Berlin, Germany

HARALD MERCKELBACH, PhD
Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands

MARGIT OSWALD, PhD
Department of Psychology, University of Berne, Berne, Switzerland

Claims of crime-related amnesia appear to be common. Using a mock crime approach, the diagnostic power of seven symptom validity instruments was investigated. Sixty participants were assigned to three conditions: responding honestly; feigning crime-related amnesia; feigning amnesia with a warning not to exaggerate. High sensitivity and specificity were obtained for the Structured Inventory of Malingered Symptomatology, the Amsterdam Short-Term Memory Test, and the Morel Emotional Numbing Test. Only three warned malingerers went undetected. The results demonstrate that validated instruments exist to support forensic decision making about crime-related amnesia. Yet, warning may undermine their effectiveness, even when using a multi-method approach.

KEYWORDS malingering, crime-related amnesia, symptom validity testing, coaching, symptom overgeneralization

Address correspondence to Peter Giger, Wankdorffeldstrasse 101, CH-3014 Bern, Switzerland. E-mail: p.giger@gmx.net

INTRODUCTION

Offenders relatively often claim to experience difficulties remembering their crimes. Several studies have concluded that the estimated prevalence rate of this crime-related amnesia varies depending on the type of crime, with the highest estimates being found for violent offences. Thus, as a rule of thumb, one may assume that 25% to 50% of murder and manslaughter suspects claim crime-related amnesia (Pujol & Kopelman, 2003). Experts differ in their opinion as to how many of these claims are authentic (e.g., due to crime-related stress causing genuine posttraumatic amnesia for the crime) and how many of them are feigned (i.e., malingered). On the basis of his clinical experience, Centor (1982) argued that he did not encounter a single case of genuine dissociative amnesia over the course of his lengthy forensic career. Likewise, in their study of more than 300 forensic inmates, Cima, Nijman, Merckelbach, Kremer, and Hollnack (2004) found no support for a trauma-dissociation account of crime-related amnesia. These authors argued that claims of amnesia must be treated with skepticism. The fact remains, however, that cases of genuine amnesia for criminal acts have been described in the literature. Often, these cases involve neurological conditions such as intoxication, seizure, or sleep disorder (Bourget & Whitehurst, 2007).

Given the legal repercussions that authentic crime-related amnesia may have in terms of competence to stand trial and criminal responsibility (e.g., Smith & Resnick, 2007), the detection of feigned crime-related amnesia is of central importance to forensic decision making (Bourget & Whitehurst, 2007; Merckelbach & Christianson, 2007). How amnesia is feigned depends on laypersons' beliefs about different kinds of amnesia, beliefs that are markedly formed by naive Hollywood examples (Baxendale, 2004).

Base rate estimates of negative response bias in populations of criminal offenders have recently been summarized by Denney (2007). When instruments that were primarily developed for civil forensic applications are used in criminal populations, elevated rates of response distortion can be found (Ardolf, Denney, & Houston, 2007).

Various psychological instruments have been developed to detect feigned amnesia. The most prominent of these symptom validity tests are based on a forced-choice format with a known probability of the possible answers (e.g., "was the victim wearing glasses: yes or no?"), which permits the detection of below-chance performance. Among forensic neuropsychologists who have been active in designing symptom validity tests, below-chance responding is widely regarded as the most secure criterion for the diagnosis of malingered neurocognitive symptoms (Slick, Sherman, & Iverson, 1999). In the words of Iverson (2003, p. 169): "The patient scored below chance on a . . . forced-choice procedure, indicating that she knew the correct answer and deliberately chose the incorrect answer. This performance invalidates the entire set of neuropsychological test results."

Forced-choice testing has been adapted to address criminal events for which forensic evaluees claimed no memory (Frederick, Carter, & Powel, 1995). Denney (1996) illustrated the usefulness of this approach with three case reports, whereas Jelicic, Merckelbach, and van Bergen (2004) employed it in experimental malingering research. Apart from forced-choice tasks specifically tailored to criminal cases (e.g., Denney, 1996), standardized forced-choice tests have been developed to evaluate claims of memory impairment. These tests have a general format and may also be used outside the criminal forensic arena. Often, these tests focus not so much on below-chance performance but on implausible performance (i.e., poor performance below the level of that of patients suffering from severe neurological diseases). In this category, one widely used instrument is the Amsterdam Short-Term Memory Test (ASTM; Schagen, Schmand, de Sterke, & Lindeboom, 1997). The ASTM is available in English, German, and Dutch, and it looks like a memory test but, in fact, measures poor effort. Other standardized symptom validity tests focusing on memory performance that are in widespread use today are the Word Memory Test (WMT; Green, 2003) and the Test of Memory Malingering (TOMM; Tombaugh, 1996), with both of them having a strong body of research.

Whereas tests such as the ASTM address claims of memory impairment, other forced-choice instruments tap more specific forms of psychopathology. A case in point is an instrument developed by Morel (1998) to detect feigned posttraumatic stress disorder (PTSD). This task, dubbed the Morel Emotional Numbing Test (MENT), resorts to the symptom of emotional numbing in PTSD patients and requires examinees to match descriptive words with pictures of emotional facial expressions. Given that each trial of words and pictures has a correct and an incorrect response, the MENT follows a forced-choice set-up. Morel and Shepherd (2008) summarized data illustrating its usefulness. The MENT is relevant to the issue of crime-related amnesia because suspects claiming such amnesia often say that they have been traumatized themselves by the crime. Two test modifications have been described in the literature. Messer and Fremouw (2007) developed a revision (MENT-R) using different pictures of facial affect expressions, and Geraerts, Jelicic, and Merckelbach (2006) developed a European test version, again with different photographs.

Besides forced-choice testing intending to tap *poor effort*, a number of other approaches have been used in the detection of feigning (e.g., of amnesia), with most of them trying to detect *exaggeration* of symptoms (e.g., Iverson, 2006). The best studied examples of these approaches are the fake bad scales derived from the MMPI-2 (e.g., Gervais, Ben-Porath, Wygant, & Green, 2007; Lees-Haley, English, & Glenn, 1991). A psychological symptom validity test that intends to tap exaggeration of symptoms is the Structured Inventory of Malingered Symptomatology (SIMS; Widows & Smith, 2005). Experimental and field studies show that the SIMS is a promising instrument

to detect feigning of neurocognitive deficits, among those amnesia (e.g., Alwes, Clark, Berry, & Granacher, 2008; Edens, Poythress, & Watkins-Clay, 2007; Jelicic, Merckelbach, Candel, & Geraerts, 2007; Vitacco, Rogers, Gabel, & Munizza, 2007).

There is a widespread belief in the literature that, compared to forced-choice testing, approaches that intend to detect exaggeration of symptoms often yield lower degrees of diagnostic accuracy (Slick et al., 1999; but see Jelicic et al., 2007). However, ideally, forensic evaluations of crime-related amnesia claims include both measures designed to detect poor effort (e.g., ASTM, MENT) and tests designed to tap exaggeration of symptoms (e.g., MMPI-2, SIMS). Such a multi-method approach would enable the forensic expert to gather converging evidence for the feigning of amnesia. Curiously enough, the multi-method approach has not been studied in the context of crime-related amnesia.

A strong argument for using a multi-method approach can be seen in the fact that different strategies may be employed by individuals who feign mental disorders. This well-known fact is the underlying cause for relatively low intercorrelations between various symptom validity measures found in a number of empirical studies. Another factor responsible for low intercorrelations between some validity measures is the differential sensitivity and specificity of the tests. Ruocco et al. (2008) have recently reported a low overlap between measures of malingering in the cognitive and psychiatric domain in a sample of compensation-seeking patients. Nelson, Sweet, Berry, Bryant, and Granacher (2007) have conducted a factor-analytic study and shown that different tests for response bias loaded on different factors that were interpreted as representing different feigning strategies. In this vein, the position paper of the National Academy of Neuropsychology (Bush et al., 2005) stated:

> In the case of sophisticated examinees, an approach that involves multiple methods at multiple points in time is typically required in order to obtain a sufficient understanding of the validity of the examinee's symptoms and performance. Use of multiple SVTs generally provides nonredundant information regarding examinee credibility. (p. 422)

Moreover, the use of multiple symptom validity tests increases diagnostic accuracy, as has recently been demonstrated by Larrabee (2007). Also, with feigned crime-related amnesia, it is largely unknown which variety of feigning strategies is followed by the offenders. Though the administration of multiple symptom validity tests in such cases seems almost self-evident, there is one reservation that psychologists often voice: Multiple tests may result in bona fide amnestics being classified as malingerers (i.e., false-positive outcomes). In contrast to this perceived reservation, with a thorough assessment of symptom validity done by a qualified examiner, convergent lines of evidence can be obtained and, consequently, the danger of false-positive results may be reduced.

In the current study, we wanted to examine how well feigned amnesia can go undetected when relying on multiple tests. However, we were also interested in the other side of the coin: To what extent do multiple tests produce false-positive outcomes? Thus, in an experimental simulation study, we instructed participants to commit a violent mock crime, after which they were told to feign amnesia for the crime. Next, multiple symptom validity tests were administered to the participants to examine whether a combination of tests is effective and safe in detecting crime-related amnesia. We also explored to what extent the warning not to overplay (which may be received from counsel or other sources) may undermine detection rates.

METHOD

Participants

The study included male participants only because severe violent crimes with subsequent claims of crime-related amnesia are mostly confined to male offenders. Thus, 63 German-speaking Swiss men participated in the study. They were recruited partly from the community ($n = 24$) and partly from a military base of the Swiss army ($n = 39$). Participation was strictly voluntary, and results were exclusively used for the purpose of this study. All participants gave written informed consent. Two individuals were excluded from the analyses because they did not conform with the instructions as became clear from the post-experimental check (see below). One person refused to participate for ethical reasons.

The mean age of the remaining 60 participants was 27.4 years ($SD = 5.3$; range: 18–42 years). Participants' educational background varied from elementary education to university and approximately reflected the current realities in Swiss society. Assignment of participants to conditions (see below) was strictly randomized. Post hoc comparisons showed no differences in age and education between the conditions.

Procedure

Participants were instructed to commit a mock crime that was described in the form of a scenario. The scenario consisted of written instructions that participants should follow step by step, reading one command at a time and following the instruction immediately before proceeding to the next command. More specifically, they had to enter a room with the intention to steal secret information that they could sell profitably to a third party. When they had entered the room, they unexpectedly found another male person sitting at a desk in the room (in fact, a dummy was placed at the desk; cf. Figures 1 and 2). As his presence interfered with their plans, they

FIGURE 1 Mock Crime Scene.

FIGURE 2 Victim.

had to get rid of the person. Participants were instructed to hit the person with great force. They found a locked cabinet in the room and its key in the pockets of the crime victim. However, instead of the information they had expected to find in the cabinet, it contained only a banknote (equivalent to about $200).

In the aftermath of the crime, so the scenario went, the participant was arrested by the police because witnesses had seen him at the scene of the

crime. He learned about the victim's death after the attack. In the subsequent legal process, a judge ordered a forensic examination of the perpetrator.

At this point, participants were assigned to either of three conditions. The first group (subsequently referred to as *honest*; $n = 20$) was instructed to respond honestly to all questions and to perform as well as they could during psychological testing. The second group (referred to as *naïve malingerers*; $n = 20$) was told to feign crime-related amnesia in an attempt to avoid criminal responsibility. The third group (referred to as *warned malingerers*; $n = 20$) should also feign crime-related amnesia, but they received an explicit warning not to overplay so as to make a credible impression during the forensic examination. The forensic examiner, they were told, might even use special tests to find out whether their report was believable.

Following recommendations of Nies and Sweet (1994), we administered a pre-assessment check to test role understanding. Thus, participants had to respond to a set of multiple-choice questions about the scenario. All of their responses had to be correct. If not, instructions were repeated until they were fully understood. For six participants, repetition was necessary.

Next, participants underwent a psychological examination as detailed below. As an incentive, participants in the malingering groups could win an equivalent of $100 for a convincing demonstration of amnesia (i.e., without being identified as malingerer by the tests). After the completion of the study, one of the honestly responding participants was chosen by chance and received the same sum of money. All persons were evaluated by the same examiner (first author).

As a post-experimental manipulation check to evaluate role commitment, a separate set of questions was given to the participants. As mentioned earlier, two experimental malingerers indicated that they had not followed the feigned amnesia scenario, and their records were excluded from all analyses.

The total time for the experiment was about 2 hours; all participants were tested individually.

Instruments

Based on considerations discussed in Iverson (2006), we decided to include in our multi-method battery tests that are able to detect suboptimal performance and measures developed to assess symptom exaggeration. Also, we decided to include only measures that are well-researched.

(1) Following recommendations of Frederick et al. (1995) and Denney (1996), a forced-choice Symptom Validity Test (FC-SVT) was constructed specifically addressing details of the crime scene and the environment in which the crime took place. Potential items were evaluated with a Doob and Kirshenbaum (1973) procedure so that from an original pool

of 50 forced-choice items, only 19 were accepted for the final version, the inclusion criterion being that incorrect answer options had to be as plausible to naïve pilot participants ($N = 30$) as correct answer options. These 19 pertinent items were mixed with another set of 19 items for which no correct answer existed (bogus items; Verschuere, Meijer, & Crombez, 2008). Examples of pertinent items are "What color was the victim's jacket? A. Fawn; B. Black." and "Where did you find the key for the cabinet? A. In the pocket of the jacket; B. In the trouser pocket." Typical bogus items were "What was the shape of the ash tray? A. Round; B. Square" (there was no ash tray in the room) and "How many boxes were located under the table? A. Two; B. Three" (in fact, there were no boxes in the room at all). Participants were instructed to guess if they did not know the correct answer. The number of correct answers to the pertinent items was counted. Below-chance responding with $p < .05$ was used as a criterion for feigned amnesia.

(2) The ASTM (Schagen et al., 1997) is a standardized forced-choice task in which each trial starts with five examples of a common category (such as vehicles, colors, wild animals) to be read aloud by the examinee. Next, a mathematical filler problem is presented, after which again five words from the same category are presented. Three words are old (i.e., were also present on the first list), and two words are new. The task consists of recognizing the old words. In total, 30 trials are presented. Thus, the maximum correct score is 90. To detect implausible performance, we employed the empirically established cutoff proposed by Schagen et al.[1]

(3) The modified version of the MENT (Morel, 1998) was based on a European adaptation of this test (Geraerts et al., 2009). The MENT is a test specifically designed for detecting feigned posttraumatic symptomatology. Basically, the test alludes to the PTSD symptom of emotional numbing. Patients with false claims of PTSD may display implausible difficulties in the perception of emotions. The test instructions were different from the original instructions in that participants were told that involvement in violent crimes may produce PTSD symptomatology. Though this addition sets limits to the generalizability of the results to the original test format, it was felt that participants should be alerted to the link between PTSD and violence, or the test might not work in the context of this application. In the current study, we used the empirically derived cutoff proposed by Morel (1998).

(4) The SIMS (Widows & Smith, 2005; German version: Cima et al., 2003) is a 75-item questionnaire developed to assess the endorsement of unlikely, bizarre, or rare symptoms in the following domains (which also constitute the subscales): low intelligence, affective disorders, neurological

[1] For reasons of test security, we do not specify the precise cutoffs. They can be obtained from the first author.

impairment, psychosis, and amnestic disorders. In the current study, three items of the German SIMS version were slightly reformulated so as to adapt them to Swiss German. These changes were related to the items no. 14 (number of Switzerland's administrative districts or cantons), no. 63 (name of a well-known Swiss political leader), and no. 67 (change to Swiss currency). A total SIMS score can be obtained by summing the number of bizarre symptom endorsements. For the total SIMS score, we used the total cutoff that Cima et al. (2003) proposed for the German version.

(5) Due to practical considerations, there was a strict limit to the total duration of the current experiment. Therefore, we could not administer the full MMPI-2. Instead, participants completed three MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kammer, 1989) validity scales, notably the original Infrequency scale (*F* scale), the Fake Bad Scale (FBS; Lees-Haley et al., 1991), and the Response Bias Scale (RBS; Gervais et al., 2007). The items of these three scales were pooled and given in the order of the original item numbers. Because these scales are not commonly used in German-speaking countries, the data thus gathered should be considered as preliminary in nature. We employed for the RBS the cutoffs proposed by Gervais et al. (2007), for the FBS the cutoff proposed by Lees-Haley (1992), and for the *F* scale the conservative cutoff of *T* scores exceeding 90 (Butcher, 2004). The latter, in fact, is the cutoff that is also recommended for the *F* scale of the German version of the MMPI-2 (Engel, 2000) whereas scores beyond $T = 70$ may be interpreted as "questionable profile validity" or "possible malingering." Having in mind that other authors have proposed different cutoffs for a number of MMPI-2 validity scales (such as $T > 80$ for the *F* scale; Rubenzer, 2006), it was outside the scope of this study and the data available for analysis to investigate the question of finding an optimal cutoff.

(6) In addition, we administered self-report scales to measure trait[2] and state dissociation[3] to add non-malingering instruments to the test battery and for the purpose of further analyses. Because these tests have been dealt with in a separate paper (Giger, Merten, Merckelbach, & Oswald, 2010), they are not presented in detail here.

To prevent order effects, test presentation was randomized (i.e., the sequence of tests was varied for each participant, according to chance allotment), with two restrictions: (a) the crime-related SVT was always given right at the beginning, to avoid interference from normal forgetting; and (b) the MMPI-2 fake bad scales were never given directly before or after the SIMS.

[2] Dissociative Experiences Scale (DES; Bernstein & Putnam, 1986; extended German version: Spitzer, Stieglitz, & Freyberger, 2005).

[3] Peritraumatic Dissociative Experiences Questionnaire (PDEQ; Marmar, Weiss, & Metzler, 1997; German version: Maercker, 1998).

Data Analyses

Because for most variables homogeneity of variance could not be assumed, non-parametric analyses were conducted (Erceg-Hurn & Mirosevich, 2008). Differences between the three groups were evaluated using separate Kruskal-Wallis tests (*H* tests), which can be seen as non-parametric equivalents of traditional analyses of variance. Post hoc comparisons were also conducted with non-parametric tests (separate Mann-Whitney *U* tests with variable-wise Bonferroni corrections).

RESULTS

The present article concentrates on the efficacy of symptom validity assessment. The analysis and discussion of dissociation, its assessment, and its proneness to malingering have been dealt with elsewhere (Giger et al., 2010).

Table 1 shows the mean scores, standard deviations, and ranges of scores of the three groups on the instruments. As can be seen, for all tests, significant group differences were evident. Furthermore, post hoc comparisons showed that relative to the honest group, both groups of malingerers performed more poorly on the tasks (i.e., FC-SVT, ASTM, and MENT) and exaggerated more on the questionnaires (i.e., SIMS, RBS, FBS, and F). Moreover, with the exception of the modified MENT, warning had a significant effect in that poor performance and exaggeration were reduced upon warning.

Using cutoffs found in the literature, we calculated diagnostic accuracy parameters for the symptom validity tests. They are shown in Table 2. High specificity rates were found for all tests. For five of seven tests, the classification of honest participants was perfect. Two participants of the honest group scored in the chance range of the crime-related forced-choice SVT (i.e., just below the threshold), and one scored somewhat above the cutoff of the SIMS, leading to their false (i.e., false-positive) classification as amnesiacs or malingerers respectively.

A different picture appeared for sensitivity. Whereas some tests (ASTM, SIMS, MENT) yielded respectable results in identifying naïve malingerers (80%–95%), this was not the case for the crime-related FC-SVT and the MMPI-2 validity scales. Moreover, sensitivity estimates for the latter instruments dropped considerably upon warning. The overall hit rate of the instruments varied from a low 45% for the MMPI-2 *F* scale to a high 88% for the ASTM.

Separate cross-table analyses were conducted to evaluate group differences in accuracy parameters. First, accuracy results of the honest group were contrasted with all malingerers. Second, naïve malingerers were compared to warned malingerers. The corresponding Chi-square statistics are

TABLE 1 Descriptive Parameters for Symptom Validity Tests

	FC-SVT	ASTM	MENT	SIMS	RBS	FBS	F
Honest Group ($n = 20$)							
Range (%)	68–100	94–100	0–12	1–23	4–50	14–42	2–25
M (*SD*)	16.6 (1.7)	88.4 (1.6)	2.6 (1.6)	5.8 (3.7)	6.4 (3.2)	12.2 (3.3)	5.4 (3.0)
Naïve Malingerers ($n = 20$)							
Range (%)	5–58	33–98	0–95	11–66	21–82	16–86	12–72
M (*SD*)	5.8 (2.8)	53.2 (19.7)	29.6 (18.6)	37.9 (13.8)	16.2 (4.0)	24.4 (7.1)	23.1 (9.5)
Warned Malingerers ($n = 20$)							
Range (%)	5–79	33–100	2–87	5–69	14–75	23–63	5–77
M (*SD*)	9.3 (3.5)	71.0 (17.9)	20.3 (14.3)	22.3 (12.6)	12.5 (5.4)	18.1 (5.1)	15.3 (11.1)
H Test ($df = 2$)							
$\chi 2$	42.5*	31.4*	25.7*	39.0*	27.9*	30.2*	31.7*
Post-hoc Comparisons *U* Test (*Z*)							
Honest vs. Naïve	5.43*	5.25*	4.47*	5.35*	5.01*	4.88*	5.23*
Honest vs. Warned	5.19*	3.76*	4.01*	4.58*	3.43*	3.53*	3.70*
Naïve vs. Warned	3.15*	2.52*	1.48	3.25*	2.17*	3.13*	2.59*

FC-SVT = forced-choice crime-related Symptom Validity Test; ASTM = Amsterdam Short-Term Memory Test; MENT = Morel Emotional Numbing Test; SIMS = Structured Inventory for Malingered Symptomatology; RBS = MMPI-2 Response Bias Scale; FBS = MMPI-2 Fake Bad Scale; *F* = MMPI-2 Infrequency Scale.

*$p < .05$.

TABLE 2 Diagnostic Accuracy Parameters of Symptom Validity Tests across Groups

		FC-SVT	ASTM	MENT	SIMS	RBS	FBS	F
Honest Group ($n = 20$)								
Correct classification as honest	(n)	18	20	20	19	20	20	20
False classification	(n)	2	0	0	1	0	0	0
Specificity	(%)	90	100	100	95	100	100	100
Naïve Malingerers ($n = 20$)								
Correct classification as malingering	(n)	9	19	16	19	8	10	5
False classification	(n)	11	1	4	1	12	10	15
Sensitivity	(%)	45	95	80	95	40	50	25
Warned Malingerers ($n = 20$)								
Correct classification as malingering	(n)	2	14	13	13	5	2	2
False classification	(n)	18	6	7	7	15	18	18
Sensitivity	(%)	10	70	65	65	25	10	10
Correct overall classification rate	(%)	48	88	82	85	55	53	45
Comparison honest group vs. all malingerers	X^2 ($df = 1$)	2.41	36.67*	28.07*	30.30*	8.30*	7.50*	3.96
Comparison naïve vs. warned malingerers	X^2 ($df = 1$)	6.13*	4.33*	1.13	5.63*	1.03	7.62*	1.56

Note: See Table 1 for listing of tests.
*$p < .05$.

TABLE 3 Number of Failures to Pass Symptom Validity Tests

	Number of test failures							
	0	1	2	3	4	5	6	7
Honest Participants	19*	1	—	—	—	—	—	—
Naïve Malingerers	—	1	2	5	1	7	1	3
Warned Malingerers	3	1	5	7	1	3	—	—

*Two of the 19 participants scored in the chance range of the crime-related FC-SVT but not below chance, so they were not classified as malingering.

also shown in Table 2. Due to its low sensitivity, the crime-related SVT failed to separate honest participants from malingerers. The same was true for the *F* scale. Although ASTM and SIMS were good at identifying malingering, they were also sensitive to warning. Only the modified MENT was both able to reliably identify malingerers and appeared to be robust against warning. However, in terms of sensitivity, the MENT performed below the ASTM both for the naïve (80% and 95%, respectively) and the warned group (65% and 70%, respectively).

Table 3 shows how many symptom validity tests participants failed to pass. One honest participant failed to pass one test (a near miss on the SIMS), and no honest participant failed to pass two or more tests. In contrast, not a single naïve malingerer passed all symptom validity scales. One of them failed to pass a single measure (which was the ASTM). Two of them failed on two tests (which were in both cases the ASTM and the SIMS). For warned malingerers, three participants were able to go undetected throughout the entire test battery. One member of the warned group failed in only one test (which, again, was the ASTM).

For the SIMS, the results were broken down into the five subscales and results presented in Tables 4 and 5. As can be seen from Table 5, the SIMS subscale Amnestic Disorders almost perfectly classified honest participants and malingers, with only two undetected warned participants. Moreover, a general trend to over-generalizing symptom report was observed. Consequently, all other SIMS subscales were relatively effective in identifying malingering. Their classification accuracy was well above that of the MMPI-2 fake bad scales. Warning had a slight effect on malingerers' ability to avoid detection, but this effect was non-significant in all cases.

DISCUSSION

The present study investigated the accuracy of different types of symptom validity tests in the context of forensic assessment of crime-related amnesia. To this purpose, a number of well-established tests, namely the ASTM, the

TABLE 4 Descriptive Parameters for the Subscales of the Structured Inventory of Malingered Symptomatology

	AM	P	LI	NI	AF
Honest Group ($n = 20$)					
Range (%)	0–13	0–20	0–27	0–13	0–60
M (*SD*)	0.8 (0.8)	0.8 (1.1)	1.2 (1.0)	0.9 (0.8)	2.2 (2.2)
Naïve Malingerers ($n = 20$)					
Range (%)	27–100	0–100	7–100	0–93	13–87
M (*SD*)	12.2 (3.5)	5.4 (4.4)	5.3 (4.2)	7.8 (4.0)	7.3 (3.3)
Warned Malingerers ($n = 20$)					
Range (%)	0–100	0–80	0–73	0–67	0–73
M (*SD*)	8.3 (4.2)	2.9 (3.6)	2.9 (3.1)	3.5 (2.5)	4.8 (3.1)
H Test ($df = 2$)					
$\chi 2$	40.0*	17.7*	14.3*	33.3*	21.8*
Post hoc Comparisons *U* Test (*Z*)					
Honest vs. Naïve	5.46*	3.80*	3.70*	4.90*	4.50*
Honest vs. Warned	4.83*	1.92	1.98	4.16*	2.79*
Naïve vs. Warned	2.96*	1.94	1.94	3.30*	2.15

AM = Amnestic Disorders; P = Psychosis; LI = Low Intelligence; NI = Neurological Impairment; AF = Affective Disorders.
*$p < .05$.

TABLE 5 Analysis of the Structured Inventory for Malingered Symptomatology on the Subscale Level

		AM	P	LI	NI	AF
Honest Group ($n = 20$)						
Correct classification as honest	(n)	20	15	19	20	19
False classification	(n)	0	5	1	0	1
Specificity	(%)	100	75	95	100	95
Naïve Malingerers ($n = 20$)						
Correct classification as malingering	(n)	20	15	12	17	13
False classification	(n)	0	5	8	3	7
Sensitivity	(%)	100	75	60	85	65
Warned Malingerers ($n = 20$)						
Correct classification as malingering	(n)	18	10	8	13	8
False classification	(n)	2	10	12	7	12
Sensitivity	(%)	90	50	40	65	40
Correct overall classification rate	(%)	97	67	65	83	67
Comparison honest group vs. all malingerers	X^2 ($df = 1$)	51.82*	7.50*	11.87*	30.00*	12.96*
Comparison naïve vs. warned malingerers	X^2 ($df = 1$)	2.11	2.67	1.60	2.13	2.51

See Table 4 for listing of subscales.

*$p < .05$.

MENT, and the SIMS, were compared with the alternative-choice method first proposed by Frederick et al. (1995) and Denney (1996). To the authors' knowledge, it is the first published study in which a forced-choice crime-related SVT was embedded in a multi-method approach.

The main results of our study can be summarized as follows. To begin with, our findings demonstrate that a multi-method approach along the lines sketched by Iverson (2006)—that is, one that includes both tests to detect underperformance and tests to tap symptom exaggeration—is successful in identifying feigned amnesia, while keeping the risk of false-positives (i.e., honest perpetrators who were wrongly classified as malingerers) at an acceptable minimum. All 20 naïve malingerers were detected by at least one of the tests, up to a maximum of all seven tests' being positive for three malingerers. A closer look at the different instruments made clear that three of them obtained high sensitivity estimates (hit rates of $\geq$ 80%): the ASTM, the modified MENT, and the SIMS. However, the SIMS also produced one false-positive case, demonstrating that notwithstanding its excellent sensitivity, one should never use this instrument in isolation. Indeed, there was not a single honest participant who failed to pass two or more symptom validity tests. Thus, to the extent that a forensic expert bases his or her clinical inference of malingering on multiple tests, the risk of false-positives is considerably reduced rather than inflated.

Second, though the crime-related forced-choice SVT is popular even up to the point that it has been introduced into popular TV drama (e.g., Cracker—Mad Woman in the Attic, directed by Jimmy McGovern, Great Britain, 1993), we found that it has a relatively low sensitivity (i.e., 45%). This rate is in keeping with what other simulation studies found (Jelicic et al., 2004; Verschuere et al., 2008). In the current study, this low sensitivity might have to do with the fact that our FC-SVT relied on only 19 pertinent items, so that the threshold for below-chance responding was as low as five correct answers. This low item number was also the basis for the fact that two participants of the honest group responded in the chance range of the test. Bearing in mind that the knowledge about the correct responses had to be acquired through incidental learning during the experiment, chance-range responding per se is not indicative for malingering or insufficient test effort.

At this point, it has to be borne in mind that with cognitive effort testing, below-chance performance is just the most striking form of distorted response behavior (cf. Slick et al., 1999), but the same decision rules cannot be employed for individually constructed symptom validity testing in cases of claimed amnesia. There, performance in the chance range would be compatible with true amnesia and not indicative of malingering as is the case with most cognitive SVTs.

Note that from the original pool of 50 forced-choice items, only 19 were found to be suitable to be included in the final test. In forensic practice, it

will be difficult to develop a set of psychometrically sound items that is considerably larger than the one used in the current study. Most important, we found that the FC-SVT is not very robust against warning, a result that was also reported by Verschuere et al. (2008). Concerning the fact that the FC-SVT is the only diagnostic opportunity to directly measure the authenticity of crime-related amnesia to date, it has to be noted that a number *as high as* 45% of naïve malingerers scored only five or fewer points in the test. Thus, having the mentioned limitations in mind, we consider the specifically designed crime-related forced-choice SVT to have a specific additional power to a multi-method evaluation of crime-related amnesia, but items have to be constructed with great circumspection and care.

Third, the study included, on a more tentative basis, three MMPI-2 malingering scales. They were developed on very different backgrounds: the *F* scale for the detection of unusual or atypical symptom endorsement, the FBS for identifying exaggerated symptom report in litigants after personal injury, and the RBS for malingered neurocognitive symptoms. To the authors' knowledge, no study so far has analysed the potential of these scales to identify feigned crime-related amnesia. Though rather disappointing sensitivity estimates were found for the MMPI-2 fake bad scales, the data have to be interpreted cautiously because the items were given in isolation, that is, outside the context of the complete questionnaire. With this limitation in mind, it also appears noteworthy that participants of the honest group passed all three MMPI-2 fake bad scales at a 100% rate.

One may argue that only the RBS is directly relevant to feigned memory deficit. However, even for this scale, the sensitivity was only 40% for the naïve and 25% for the warned group. Moreover, different cutoffs have been proposed for a number of validity scales, in particular for the FBS (Butcher, Gass, Cumella, Kally, & Williams, 2008; Larrabee, 2003; Lees-Haley & Fox, 2004). If the cutoff proposed by Butcher et al. (2008), which is the one currently used by Pearson Assessments, would have been used for the classification, FBS sensitivity rates would drop considerably from 50% to 30% for the naïve and from 10% to 0% for the warned malingerers. One of the major challenges to forensic research and practice in the German-speaking countries appears to be that there are, altogether, very little data on the application of the MMPI-2 validity scales and the different cutoffs proposed for the original American version. Though recent results appear to be promising (Diederich, Kathmann, & Merten, 2009), the authors feel that the MMPI-2 validity scales deserve more attention in empirical research.

Fourth, warning decreased the sensitivity of the symptom validity tests to varying degrees. At the level of raw scores, considerable score reductions were observed for the warned group. Only for the modified MENT, this reduction did not reach the level of statistical significance, but even here the trend was visible. Even with this relative resistance against warning, the MENT did not outperform the ASTM, neither for the naïve nor

for the warned group. For some tests, sensitivity estimates were considerably reduced for the warned group, such that hit rates as low as 10% were obtained for the crime-related FC-SVT, the FBS, and the *F* scale. These results demonstrate the ramifications of a real-world phenomenon. Patients and suspects are increasingly able to find information about assessment techniques, using modern media (Bauer & McCaffrey, 2006). Moreover, in both civil and forensic contexts, coaching of test persons is a problem of increased importance (e.g., Wetter & Corrigan, 1995; Youngjohn, 1995). These developments undermine the diagnostic accuracy of instruments and assessment procedures. Not surprisingly, studies on the effect of coaching have become an important branch of malingering research (e.g., Gorny & Merten, 2005; Merten, Diederich, & Stevens, 2008). In this context, warning is usually considered to be an elementary form of coaching. The problem of warning is even more pertinent as informed consent of patients in forensic assessments may contain an explicit warning not to exaggerate and to be honest, and that specific assessment techniques are used to check cooperation. Though there is no agreement yet on whether such a procedure is mandatory (Sharland & Gfeller, 2007), some authors maintain that an explicit warning may be advisable (Bush et al., 2005). The apparent consequence of all this for malingering detection is that an appropriate symptom validity test should be able to maintain high sensitivity rates when faced with warned evaluees (e.g., Hartman, 2002). In our study, only the ASTM, MENT, and SIMS seemed to meet that criterion, suggesting that a multi-method approach should at the very least contain these tests or their equivalents.

Finally, the analyses of the SIMS subscales showed some noticeable points. The participants were specifically instructed to fake crime-related amnesia, and consequently, almost all of them (100% of the naïve and 90% of the warned malingerers) scored above cutoff in the subscale Amnestic Disorders. However, experimental malingerers did not limit their exaggerated symptom report to the domain of memory but overgeneralized so much so that all other SIMS subscales were successful in detecting some of the participants. This was particularly true for the scale Neurological Impairment, for which 85% of naïve and 65% of warned malingerers were correctly identified. It has to be noted that subscale analyses may be restricted due to limited reliability of some of the scales, but the results presented here appear to be of interest specifically with regard to the phenomenon of symptom overgeneralization.

Several limitations of the current study deserve some comment. First of all, the test battery was composed almost exclusively of tests for malingering detection. In a more realistic scenario, assessment strategies would combine personality measures, symptom report questionnaires, validity scales, and possibly cognitive performance measures in a way that makes it harder to discover to what category a given test belongs. This is, in a real-world design, fundamental for collecting valid data in forensic assessment. A difficult

question is that of which tests to choose from the repertory of existing instruments. This question was particularly pertinent for the cognitive symptom validity tests, with a number of potent instruments available. Some effort tests combine validity measures with measures of memory functioning (e.g., the WMT; Green, 2003). The TOMM (Tombaugh, 1996) may be a good choice for application in a population with a high percentage of foreign language users. Digit memory tests such as the Victoria Symptom Validity Test (VSVT; Slick, Hopp, Strauss, & Spellacy, 1996) have shown promising capability to detect malingering in civil forensic evaluations. In the context of this European study, the authors opted for using one of the few original European symptom validity tests, which appears to be as sound an instrument as the above mentioned tests (cf. Henry, 2009). The WMT, the TOMM, the VSVT, or a number of other cognitive symptom validity tests could have been legitimately employed instead of the ASTM in the framework of this study. However, in a real-world forensic evaluation, the expert has to check carefully the properties of the tests available for a special question and try and opt for most suitable of all instruments for that question.

With regard to the MENT, which is the only symptom validity measure to date specifically designed to tap malingered PTSD, the test instructions were slightly changed to induce a connection between PTSD and crime-related amnesia. This, obviously, sets limits to the generalizability of the results to the original test format.

Furthermore, it has to be kept in mind that the examination of the participants was always conducted by the same person (the first author) who was not blind to the condition to which participants were assigned. This can be seen as a potentially limiting factor, but it was not possible to perform the study under different terms.

Another limitation of the current and of any other analog study is the problem of external validity (Dearth et al., 2005). Healthy experimental participants and the conditions under which they are tested will differ in many ways from evaluees and test conditions found in real-world forensic assessments. Also, results obtained with healthy Swiss volunteers can, of course, not be generalized to a criminal population of Swiss or any other origin. One aspect of the experimental procedure whereby such limitations become obvious is the time lapse between (mock) offence and the forensic evaluation. Having said that, internal validity is usually higher in analog studies because they allow for thorough experimental control (Vickery et al., 2004). Rogers (1997) suggested that both analog studies and known-groups designs should be conducted before a particular test or procedure is implemented in forensic evaluations. Thus, it would be worthwhile to replicate our multimethod approach in suspects who maintain or gave up their amnesia claims. Furthermore, it would be interesting to test real amnestics to specify how many of these patients would probably have been classified as malingerers (false-positives).

Even though our sample size was relatively small, the study adds to the findings of previous studies in that it shows that a number of promising methods exist to support forensic decision making. Thus, it appears to be relatively difficult to feign convincingly amnesia without being identified by sophisticated assessment techniques in a multi-method approach. So far, few authors have addressed the methods of crime-related SVTs, so their real potential and limitations are underrepresented in the relevant literature. More analog studies but, even more important, more well-planned naturalistic studies are needed to determine which tests and diagnostic approaches are most appropriate for application in the criminal forensic arena when potential feigning of crime-related amnesia is an issue.

REFERENCES

Alwes, Y. R., Clark, J. A., Berry, D. T. R., & Granacher, R. P. (2008). Screening for feigning in a civil forensic setting. *Journal of Clinical and Experimental Neuropsychology*, *30*, 133–140.

Ardolf, B. R., Denney, R. L., & Houston, C. M. (2007). Base rates of negative response bias and malingered neurocognitive dysfunction among criminal defendants referred for neuropsychological evaluation. *The Clinical Neuropsychologist*, *21*, 899–916.

Bauer, L., & McCaffrey, R. J. (2006). Coverage of the Test of Memory Malingering, Victoria Symptom Validity Test, and Word Memory Test on the internet: Is test security threatened? *Archives of Clinical Neuropsychology*, *21*, 121–126.

Baxendale, S. (2004). Memories aren't made of this: Amnesia at the movies. *British Medical Journal*, *329*, 18–25.

Bernstein, E. M., & Putnam, F. W. (1986). Development, reliability, and validity of a dissociation scale. *Journal of Nervous and Mental Disease*, *174*, 727–735.

Bourget, D., & Whitehurst, L. (2007). Amnesia and crime. *Journal of the American Academy of Psychiatry and the Law*, *35*, 469–480.

Bush, S. S., Ruff, R. M., Tröster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., et al. (2005). Symptom validity assessment: Practice issues and medical necessity. *Archives of Clinical Neuropsychology*, *20*, 419–426.

Butcher, J. N. (2004). *A beginner's guide to the MMPI-2* (2nd ed.). Washington, DC: American Psychological Association.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis: University of Minnesota Press.

Butcher, J. N., Gass, C. S., Cumella, E., Kally, Z., & Williams, C. L. (2008). Potential for bias in MMPI-2 assessments using the Fake Bad Scale. *Psychological Injury and the Law*, *1*, 191–209.

Centor, A. (1982). Criminals and amnesia: Comment on Bower. *American Psychologist*, *37*, 240.

Cima, M., Hollnack, S., Kremer, K., Knauer, E., Schellbach-Matties, R., Klein, B., et al. (2003). Strukturierter Fragebogen Simulierter Symptome. Die deutsche Version des "Structured Inventory of Malingered Symptomatology: SIMS" [The German version of the Structured Inventory of Malingered Symptomatology: SIMS]. *Nervenarzt*, *74*, 977–986.

Cima, M., Nijman, H., Merckelbach, H., Kremer, K, & Hollnack, S. (2004). Claims of crime-related amnesia in forensic patients. *International Journal of Law and Psychiatry*, *27*, 215–221.

Dearth, C. S., Berry, D. T. R., Vickery, C. D., Vagnini, V. L., Baser, R. E., Orey, S. A., et al. (2005). Detection of feigned head injury symptoms on the MMPI-2 in head injured patients and community controls. *Archives of Clinical Neuropsychology*, *20*, 95–110.

Denney, R. L. (1996). Symptom validity testing of remote memory in a criminal forensic setting. *Archives of Clinical Neuropsychology*, *11*, 589–603.

Denney, R. L. (2007). Assessment of malingering in criminal forensic neuropsychological settings. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 428–452). New York: Guilford.

Diederich, C., Kathmann, N., & Merten, T. (2009, May). *An extension of the MMPI-2 Meyers Validity Index for the detection of malingered depression and PTSD*. Poster presented at the First European Symposium on Symptom Validity Assessment, Wuerzburg, Germany.

Doob, A. N., & Kirshenbaum, H. M. (1973). Bias in police line-ups partial remembering. *Journal of Police Science and Administration*, *1*, 287–293.

Edens, J. F., Poytress, N. G., & Watkins-Clay, M. M. (2007). Detecting malingering in psychiatric unit and general population prison inmates: A comparison of the PAI, SIMS, and SIRS. *Journal of Personality Assessment*, *88*, 33–42.

Engel, R. R. (Ed.). (2000). *MMPI-2. Minnesota Multiphasic Personality Inventory-2: Manual*. Bern, Switzerland: Huber.

Erceg-Hurn, D. M., & Mirosevich, V. M. (2008). Modern robust statistical methods: An easy way to maximize the accuracy and power of your research. *American Psychologist*, *63*, 591–601.

Frederick, R. I., Carter, M., & Powel, J. (1995). Adapting symptom validity testing to evaluate suspicious complaints of amnesia in medicolegal evaluations. *Bulletin of the American Academy of Psychiatry and the Law*, *23*, 231–237.

Geraerts, E., & Brennen, T. (2006). Investigating cognitive abnormalities in posttraumatic stress disorder. In M. J. Roy (Ed.), *Novel approaches to the diagnosis and treatment of posttraumatic stress disorder* (pp. 31–40). Amsterdam: IOS Press.

Geraerts, E., Jelicic, M., & Merckelbach, H. (2006). Symptom overreporting and recovered memories of childhood sexual abuse. *Law and Human Behavior*, *30*, 621–630.

Geraerts, E., Kozaric-Kovacic, D., Merckelbach, H., Peraica, T., Jelicic, M., & Candel, I. (2009). Detecting deception of war-related posttraumatic stress disorder. *Journal of Forensic Psychiatry and Psychology*, *20*, 278–285.

Gervais, R. O., Ben-Porath, Y. S., Wygant, D. B., & Green, P. (2007). Development and validation of a Response Bias Scale (RBS) for the MMPI-2. *Assessment*, *14*, 196–208.

Giger, P., Merten, T., Merckelbach, H., & Oswald, M. (2010). Willentliche Testverfälschung bei Verfahren zur Erfassung von Dissoziation [Response distortions in scales for measuring dissociation] *Praxis der Rechtpsychologie, 20*, 131–147.

Gorny, I., & Merten, T. (2005). Symptom information—warning—coaching: How do they affect successful feigning in neuropsychological assessment? *Journal of Forensic Neuropsychology*, *4*, 71–97.

Green, P. (2003). *Green's Word Memory Test: User's manual*. Edmonton, Canada: Green's Publishing.

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.

Henry, M. (2009). Beschwerdenvalidierungstests in der zivil- und sozialrechtlichen Begutachtung: Verfahrensüberblick [Symptom validity tests in civil forensic assessment: A review of instruments]. In T. Merten & H. Dettenborn (Eds.), *Diagnostik der Beschwerdenvalidität* (pp. 118–161). Berlin, Germany: Deutscher Psychologen Verlag.

Iverson, G. L. (2003). Detecting malingering in civil forensic evaluations. In A. M. Horton & L. C. Hartlage (Eds.), *Handbook of forensic neuropsychology* (pp. 137–177). New York: Springer.

Iverson, G. L. (2006). Ethical issues associated with the assessment of exaggeration, poor effort, and malingering. *Applied Neuropsychology*, *13*, 77–90.

Jelicic, M., Merckelbach, H., & van Bergen, S. (2004). Symptom validity testing of feigned amnesia for a mock crime. *Archives of Clinical Neuropsychology*, *19*, 525–531.

Jelicic, M., Merckelbach, H., Candel, I., & Geraerts, E. (2007). Detection of feigned cognitive dysfunction using special malinger tests: A simulation study in naïve and coached malingerers. *International Journal of Neuroscience*, *117*, 1185–1192.

Larrabee, G. J. (2003). Exaggerated MMPI-2 symptom report in personal injury litigants with malingered neurocognitive deficits. *Archives of Clinical Neuropsychology*, *18*, 673–686.

Larrabee, G. J. (2007). Aggregation across multiple indicators improves the detection of malingering: Relationship to likelihood ratios. *The Clinical Neuropsychologist*, *22*, 666–679.

Lees-Haley, P. R. (1992). Efficacy of MMPI-2 validity scales and MCMI-II modifier scales for detecting spurious PTSD claims: F, F-K, Fake Bad Scale, Ego Strength, Subtle-Obvious subscales, DIS, and DEB. *Journal of Clinical Psychology*, *48*, 681–688.

Lees-Haley, P. R., English, L. T., & Glenn, W. J. (1991). A Fake Bad Scale on the MMPI-2 for personal injury claimants. *Psychological Reports*, *68*, 208–210.

Lees-Haley, P. R., & Fox, D. D. (2004). Commentary on Butcher, Arbisi, Atlis, and McNulty (2003) on the Fake Bad Scale. *Archives of Clinical Neuropsychology*, *19*, 333–336.

Maercker, A. (1998). *Die deutsche Version des Peritraumatic Dissociation Experience Questionnaire (PDEQ)* [German version of the PDEQ]. Unpublished manuscript. Dresden: Technical University.

Marmar, C. R., Weiss, D. S., & Metzler, T. J. (1997). The Peritraumatic Dissociative Experience Questionnaire. In J. P. Wilson & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD* (pp. 197–217). New York: Guilford.

Merckelbach, H., & Christianson, S. A. (2007). Amnesia for homicides as a form of malingering. In S. A. Christianson (Ed.), *Offenders' memories of violent crimes* (pp. 165–190). Chichester, England: Wiley.

Merten, T., Diederich, C., & Stevens, A. (2008). Vorgetäuschte Beschwerden nach Distorsionstrauma der Halswirbelsäule: eine experimentelle Simulationsstudie [Feigned symptoms after whiplash injury: An experimental analogue study]. *Aktuelle Neurologie*, *35*, 8–15.

Messer, J. M., & Fremouw, W. J. (2007). Detecting malingered posttraumatic stress disorder using the Morel Emotional Numbing Test-Revised (MENT-R) and the Miller Forensic Assessment of Symptoms Test (M-FAST). *Journal of Forensic Psychology Practice*, *7*, 33–57.

Morel, K. R. (1998). Development and preliminary validation of a forced-choice test of response bias for posttraumatic stress disorder. *Journal of Personality Assessment*, *70*, 299–314.

Morel, K. R., & Shepherd, B. E. (2008) Meta-analysis of the Morel Emotional Numbing Test for PTSD: Comment on Singh, Avasthi, and Grover. *German Journal of Psychiatry*, *11*, 128–131.

Nelson, N. W., Sweet, J. J., Berry, D. T., Bryant, F. B., & Granacher, R. P. (2007). Response validity in forensic neuropsychology: Exploratory factor analytic evidence of distinct cognitive and psychological constructs. *Journal of the International Neuropsychological Society*, *13*, 440–449.

Nies, K. J., & Sweet, J. J. (1994). Neuropsychological assessment of malingering: A critical review of past and present strategies. *Archives of Clinical Neuropsychology*, *9*, 501–552.

Pujol, M., & Kopelman, D. (2003). Psychogenic amnesia. *Practical Neurology*, *3*, 292–299.

Rogers, R. (1997). Researching dissimulation. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 398–426). New York: Guilford Press.

Rubenzer, S. (2006). Malingering of psychiatric problems, brain damage, chronic pain, and controversial syndromes in a personal injury context. *FDCC Quarterly*, *56*, 499–523.

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*, 547–564.

Schagen, S., Schmand, B., de Sterke, S., & Lindeboom, J. (1997). Amsterdam Short-Term Memory test: A new procedure for the detection of feigned memory deficits. *Journal of Clinical and Experimental Neuropsychology*, *19*, 43–51.

Sharland, M. J., & Gfeller, J. D. (2007). A survey of neuropsychologists' beliefs and practices with respect to the assessment of effort. *Archives of Clinical Neuropsychology*, *22*, 213–223.

Slick, D. J., Sherman E. M., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: proposed standards for clinical practice and research. *The Clinical Neuropsychologist*, *13*, 545–561.

Slick, D. J., Hopp, G., Strauss, E., & Spellacy, F. J. (1996). Victoria Symptom Validity Test. Efficiency for detecting feigned memory impairment and relationship to neuropsychological tests and MMPI-2 validity scales. *Journal of Clinical and Experimental Neuropsychology*, *18*, 911–922.

Smith, D., & Resnick, Ph. (2007). Amnesia and competence to stand trial. *Journal of the American Academy of Psychiatry and the Law*, *35*, 541–543.

Spitzer, C., Stieglitz, R.-D., & Freyberger, H. J. (2005). *Fragebogen zu Dissoziativen Symptomen: Testmanual zur Kurz- und Langform (FDS-20 und FDS)* [Dissociative Experiences Scale: Manual of the short version and the extended version]. Bern, Switzerland: Huber.

Tombaugh, T. N. (1996). *Test of Memory Malingering (TOMM)*. North Tonawanda, NY: Multi-Health Systems.

Verschuere, B., Meijer, E., & Crombez, G. (2008). Symptom validity testing for the detection of simulated amnesia: Not robust to coaching. *Psychology, Crime & Law*, *14*, 523–528.

Vickery, C. D., Berry, D. T. R., Dearth, C. S., Vagnini, V. L., Baser, R. E., Cragar, D. E., et al. (2004). Head injury and the ability to feign neuropsychological deficits. *Archives of Clinical Neuropsychology*, *19*, 37–48.

Vitacco, M. J., Rogers, R., Gabel, J., & Munizza, J. (2007). An evaluation of malingering screens with competency to stand trial patients: A known-group comparison. *Law and Human Behavior*, *31*, 249–260.

Wetter, M. W., & Corrigan, S. K. (1995). Providing information to clients about psychological tests: A survey of attorneys' and law students' attitudes. *Professional Psychology: Research and Practice*, *26*, 474–477.

Widows, M. R., & Smith, G. P. (2005). *SIMS— Structured Inventory of Malingered Symptomatology: Professional manual*. Lutz, FL: Psychological Assessment Resources.

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