Detecting coached feigning using the Test of Memory Malingering (TOMM) and the Structured Inventory of Malingered Symptomatology (SIMS)

Marko Jelicic[1], Erik Ceunen[2], Maarten J.V. Peters[1] & Harald Merckelbach[1]

[1]Forensic Psychology Section, Maastricht University, The Netherlands

[2]Department of Psychology, University of Leuven, Belgium

Correspondence: Marko Jelicic, Forensic Psychology Section, Department of Clinical Psychological Science, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Email: m.jelicic@maastrichtuniversity.nl

Abstract

The aim of the present study was to compare the accuracy of the Test of Memory Malingering (TOMM) with that of the Structured Inventory of the Malingered Symptomatology (SIMS) in detecting feigning of cognitive dysfunction in participants who received different types of coaching. Ninety undergraduate students were administered the TOMM and the SIMS and asked to respond honestly (controls; $n = 30$), or instructed to feign cognitive dysfunction due to head injury ($n = 2x30$). Before both instruments were administered, symptom-coached feigners ($n = 30$) were provided with some information about brain injury, while feigners who received a mix of symptom-coaching and test-coaching ($n = 30$) were given the same information plus advice on how to defeat symptom validity tests. When used separately, the subtests of the TOMM correctly classified 87-97 percent of the symptom-coached feigners and 80 percent of the feigners who received symptom-coaching plus test-coaching. When these subtests were used in combination, the TOMM correctly identified 97 percent of the symptom-coached feigners and 87 percent of the symptom/test-coached feigners. The SIMS correctly classified 93 percent of the symptom-coached feigners and 86 percent of the symptom/test-coached feigners. None of the honest responders were misclassified as feigners by either the TOMM or the SIMS. Our findings suggest that both the TOMM and the SIMS are relatively resistant to the effects of coaching.

Key words: feigning, brain injury, coaching, TOMM, SIMS

Introduction

Evidence suggests that a nontrivial percentage of neuropsychological patients seeking financial compensation for cognitive dysfunction perform below their actual cognitive capabilities and/or engage in overreporting of symptoms (Mittenberg, Patton, Canyock, & Condit, 2002; Sharland & Gfeller, 2007). Different strategies have been proposed to investigate symptom validity in medicolegal cases. Some authors opine that a negative response basis can be detected by looking at inconsistent test performance on standard neuropsychological tests (Hartlage, 1998). For example, attention and concentration mediate memory performance. If a patient performs well on memory tests, but poorly on tests of attention and concentration, this would be a possible indication of underperformance on some of the tests. Research, however, has shown that clinicians have great difficulty in detecting negative response bias. In a number of experiments, experienced clinical neuropsychologists were given neuropsychological test protocols from several individuals (Faust, Hart, & Guilmette, 1988; Heaton, Smith, Lehman, & Vogt, 1978). Some protocols were from healthy participants who were instructed to feign cognitive dysfunction, while others were from patients with genuine brain injury. The clinicians were asked to judge which protocols were indicative of feigning. Although the neuropsychologists were fully informed about the purpose of the experiments, they often were unable to distinguish instructed feigners from genuine patients.

The finding that clinicians have difficulty in detecting negative response bias has inspired researchers to develop special symptom validity tests. These tests are usually based on the notion that people engaged in exaggeration and feigning have limited knowledge of true neurological and psychopathological symptoms (Larrabee, 2007). Broadly speaking, there are two types of symptom validity tests (Iverson, 2006). The first category consists of cognitive tests designed to measure underperformance or poor effort. These tests usually involve an easy memory task requiring only passive recognition. Feigners do not know that most genuine brain-injured patients do quite well on simple recognition tasks. Because they want to convince the examiner that they suffer from brain damage, feigners will perform more poorly than true patients. One of the most widely used and best studied performance tests designed to detect cognitive underperformance is the Test of Memory Malingering (TOMM; Tombaugh, 1996). The second category of symptom validity tests are self-report instruments that measure overendorsement of symptoms. These tests present patients with items that refer to atypical and bizarre symptoms that – in the eyes of laypeople – seem to be related to a particular syndrome or condition (such as brain damage). Because feigners want to appear suffering from a syndrome or condition, they will endorse substantially more of these atypical and bizarre items than genuine patients. A widely used self-report tool developed to measure overendorsement of symptoms is the Structured Inventory of the Malingered Symptomatology (SIMS; Smith & Burger, 1997).

There is reason to believe that attorneys sometimes educate their clients about how to feign symptoms of brain injury prior to a neuropsychological evaluation (Youngjohn, 1995; see also Rosen, 1995). Such "attorney coaching" may undermine the accuracy of symptom validity tests like the TOMM and the SIMS. Therefore, Youngjohn, Lees-Haley, and Binder (1999) argued that providing feigners with information about true patients and warning them not to exaggerate symptoms could lead to sophisticated feigning behavior. Meanwhile, systematic research on the effect of coaching on the accuracy of symptom validity tests is largely lacking (Blaskewitz & Merten, 2007; Gorny & Merten, 2005; Suhr & Gunstad, 2007). What is known, though, is that when coaching is brief and only involves a superficial introduction to the to-be-feigned symptoms, its undermining effects on the ability of symptom validity tests to detect feigning are limited (e.g., Rose, Hall, Szalda-Petree & Bach, 1998). On the other hand, one can also find examples in the literature where more extensive forms of coaching did lead to considerable drops in the efficacy of symptom validity tests to detect feigned disorders (e.g., Gunstad & Suhr, 2001). According to Powell, Gfeller, Hendricks, & Sharland (2004) there are two types of coaching. Firstly, potential feigners can be educated about the symptoms of the to-be-feigned disorder. This type of coaching has been labelled “symptom-coaching”. Secondly, they can also be provided with information on how to defeat symptom validity tests. This latter type of coaching has been termed “test-coaching”.

The present study was initiated to study the effects of different kinds of coaching on the accuracy of two widely used symptom validity tests, the TOMM and the SIMS, used to investigate negative response bias. Although there are hints in the literature that both the TOMM (see Powell et al., 2004) and the SIMS (see Jelicic, Hessels, & Merckelbach, 2006; Jelicic, Merckelbach, Candel, & Geraerts, 2007) are relatively immune to a moderate degree of coaching (symptom-coaching), we hypothesized that more extensive coaching (symptom-coaching plus test-coaching) would undermine the diagnostic accuracy of both instruments.

Methods

Participants

There were 90 participants (20 men) who volunteered to take part in a simulation experiment. They were undergraduate students, alumni, or employees of Maastricht University. Some of the students received course credit, while others participated in the experiment in return for a small financial compensation (approximately 8 US$). The experiment was approved by the standing ethical committee of the Faculty of Psychology and Neuroscience of Maastricht University. Mean age was 22 years (*SD* = 4.7; range 17-49 years). Participants with a history of neurological or psychiatric disorders were excluded from the study. None of the participants had any knowledge of psychological or psychiatric methods. They were randomly assigned to one of the three groups.

Materials

Participants were administered the TOMM and the SIMS. The TOMM (Tombaugh, 1996) contains two learning trials where the participant is presented with 50 line drawings of common objects, followed by a forced choice recognition task. Feedback is given after each item. A Retention Trial, given 20 minutes after the last learning trial, consists of the forced choice task only. A score of less than 45 on the second learning trial (Trial 2) or the Retention Trial is considered indicative of underperformance. The SIMS (Smith & Burger, 1997) is a self-report measure designed to screen for feigning in five domains: low intelligence, affective disorders, neurological impairment, psychosis, and amnesia. It consists of 75 yes-no items that refer to bizarre and atypical symptoms and complaints. Total SIMS scores range from 0 to 75. Scores exceeding the cut-off of 16 are regarded to be indicative of overreporting. The Dutch version of the SIMS (Merckelbach & Smith, 2003) was used in the present study.

Procedure

Because the experimenter had to remain “blind” for the assignment of the participants to the three conditions, participants were given a blank envelope containing the instructions. To ensure random distribution of the three types of instructions, the envelopes were shuffled before the data collection started (i.e., before the first participant took part in the study). Participants were requested not to disclose the instructions to the experimenter before the end of the test session. In the control group (n = 30), participants were asked to do their best on the tests and to fill out the questionnaire honestly. In the symptom-coached feigning group (n = 30), participants were given a scenario about an accident in which they sustained head injury, and they were asked to act as though they had deficits resulting from that accident. These participants were specifically instructed to feign or exaggerate symptoms of brain injury in a believable way. They were also provided with information about cognitive dysfunction after head injury, and were warned that at least one of the tests they would receive was designed to catch them feigning. The scenario and instructions for the symptom-coached group were taken from Suhr and Gunstad (2000). In the feigning group who received a mix of symptom-coaching and test-coaching (n = 30), participants received the same scenario and the same information about cognitive dysfunction after head injury. Like in the group that underwent symptom-coaching, they received a list of symptoms and were told that at least one of the tests they would receive was designed to catch them feigning. However, they also received the following instruction: *Note that, instruments developed to identify faking of cognitive problems look just like normal tests and questionnaires. However, on these instruments patients with brain damage and healthy people who are not engaged in faking cognitive problems obtain similar scores. Therefore, try to score normally on instruments that could be designed to catch you faking!* The participants in the three groups first completed the first two trials of the TOMM. They then had to fill out the SIMS. Finally, they were given the Retention Trial of the TOMM.

Results

Unfortunately, four participants (one in the control group, one in the symptom-coached feigning group, and two in the feigning group that underwent symptom-coaching plus test-coaching) had skipped a number of questions on the SIMS. They were therefore excluded from the analyses of the SIMS scores.

Mean total TOMM and SIMS scores of the three groups are presented in Table 1. One-way Analysis of Variance (ANOVA) showed significant differences between the three groups with regard to scores on Trial 2 of the TOMM [F (2, 87) = 48.8; p < .05]. Post-hoc analyses using Student-Newman-Keuls (SNK) tests demonstrated that participants in the control group had higher scores on this subtest of the TOMM than those in the two feigning groups (both p's < .05), while there was no reliable difference in Trial 2 scores between the two feigning groups. An ANOVA also showed significant differences between the three groups with respect to scores on the Retention Trial of the TOMM [F (2, 87) = 52.6; p < .05]. Post-hoc comparisons using SNK tests showed that participants in the control group again had higher scores on this subtest of the TOMM than those in the two feigning groups (both p's < .05). In addition, the two feigning groups did not differ in their scores on the Retention Trial of the TOMM. An ANOVA also indicated significant differences in SIMS scores between the three groups [F (2, 83) = 93.7; p < .05]. Post-hoc analyses using SNK tests showed that participants in the control group had lower SIMS scores than those in the two feigning groups (both p's < .05). There was no reliable difference between the two feigning groups with regard to the SIMS scores.

Table 1 about here

Table 2 shows the number (and percentage) of participants in each group that were correctly classified as honest responders or feigners on the basis of the TOMM or SIMS cut-offs. All the control participants were correctly classified as honest responders (specificity). Moreover, 87 percent of the symptom-coached participants and 80 percent of the symptom/test-coached participants were correctly identified as feigners by Trial 2 of the TOMM, while 97 percent of the symptom-coached participants and 80 percent of the symptom/test-coached participants were correctly classified by the Retention Trial of the TOMM (sensitivity).The difference in detection rate between feigners who received symptom-coaching only and those who underwent symptom-coaching and test-coaching was significant for the Retention Trial [χ^2 (1) = 4.04, $p < .05$], but not for Trial 2. Note that there were participants who had a score above the cut-off on Trail 2 of the TOMM and below the cutting point on the Retention trial, and vice versa. Only one participant in the symptom-coached feigning group and four participants in the symptom/test-coached feigning group had scores above the cut-off on both subtests of the TOMM. When participants with at least one score below the cutting point are labelled as feigners, 97 percent of the symptom-coached feigners and 87 percent of the symptom/test-coached feigners were correctly identified. As for the SIMS, 93 percent of the symptom-coached participants and 86 percent of the symptom/test-coached participants were correctly classified as feigners. The difference in detection rates between feigners who received symptom-coaching only and those who underwent symptom-coaching plus test-coaching failed to reach significance.

Table 2 about here

Discussion

The results of our study can be summarized as follows. Neither the TOMM nor the SIMS yielded any false positives. That is, on both tests there were no control participants who were misclassified. When using only Trial 2 or the Retention Trial of the TOMM, 87-97 percent of the symptom-coached participants and 80 percent of the symptom/test-coached participants were correctly classified as feigners. When both subtests of the TOMM were used in combination, 97 percent of the symptom-coached participants and 87 percent of the symptom/test-coached participants were correctly identified as feigners. The SIMS correctly classified 93 percent of the symptom-coached participants and 86 percent of the symptom/test-coached participants as feigners. Although the accuracy of the TOMM and SIMS appears to be somewhat reduced by a mix of symptom-coaching and test-coaching, even feigners who underwent this type of coaching were most of the times defeated by the symptom validity tests. Thus, it seems that both the TOMM and the SIMS are relatively resistant to the effects of coaching.

Some authors (e.g., Iverson, 2006) propose that multiple symptom validity tests should be used in forensic neuropsychological assessments. This notion is supported by the results of the present study. In our study, there were only two instructed feigners (both from the group that received a mix of symptom-coaching and test-coaching) who were able to score normally on all three symptom validity measures.

The present findings are well in line with previous studies. Powell et al. (2004) studied the effect of symptom-coaching and test-coaching on the ability of the TOMM to detect feigned cognitive dysfunction. Participants were asked to feign and received either symptom-coaching or test-coaching. Powell and colleagues found that 93 percent of the symptom-coached and 96 percent of the test-coached participants were correctly classified as feigners. Jelicic et al. (2007) examined the effect of symptom-coaching on the efficacy of the SIMS to detect feigning of cognitive dysfunction and found that 90 percent of the coached participants were correctly identified as feigners.

A few potential limitations of the current study warrant comment. Most of our participants were women. Although there is no a priori reason to assume that coaching leads to more sophisticated feigning behavior in men relative to women, the overrepresentation of female participants may have influenced our results. A more serious limitation pertains to the use of instructed feigners. In our study, as well as in many other studies, normal participants were instructed to feign cognitive dysfunction. Some authors argue that real life feigners have a greater motivation to feign in a sophisticated way than research participants who are instructed to feign cognitive dysfunction (Rogers, 2008). Although one could argue that our findings may have limited external validity, there is evidence that students can mimic neuropsychological test scores of brain injured patients in quite a believable way (Haines & Norris, 2001). Also, if symptom validity tests such as the TOMM and the SIMS appear to be relatively insensitive to coached feigning attempts of college students, their robustness might even be better in a forensic or clinical setting who often have lower levels of education.

There are hints in the literature that bona fide patients, especially those with severe cognitive dysfunction, do not always pass symptom validity tests (Teichner & Wagner, 2004). Because all our controls were normal participants who were asked to respond honestly, the absence of false positives in our study may be somewhat unrealistic. To get a more realistic idea of specificity rates in medicolegal and clinical settings, future research on the efficacy of the TOMM and the SIMS should use an additional control group consisting of patients with documented brain injury.

Finally, one could argue that our test-coaching was somewhat superficial. We educated participants undergoing this type of coaching about the basic idea behind symptom validity tests. They were not provided with detailed information about the TOMM and the SIMS. In future studies, it would be interesting to investigate the robustness of the TOMM and SIMS using more thorough forms of test-coaching.

Notwithstanding the above limitations, it seems that both the TOMM and the SIMS are relatively robust to a combination of symptom-coaching and test-coaching. Thus, both instruments may be used in the assessment of cognitive dysfunction in medicolegal cases.

References

Blaskewitz, N., & Merten, T. (2007). Diagnostik der Beschwerdenvalidät: Diagnostik der Simulationsverdacht. Ein Update 2002-2005 [Diagnostics of malingering: An update]. *Fortschritte der Neurologie und Psychiatrie, 75,* 140-154.

Faust, D., Hart, K.J., & Guilmette, T.J. (1988). Pediatric malingering: The capacity of children to fake believable deficits on neuropsychological testing. *Journal of Consulting and Clinical Psychology, 56,* 578-582.

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.

Gunstad, J., & Suhr, A. (2001). Courting the clinician: Efficacy of full and abbreviated forms of the Portland Digit Recognition Test: Vulnerability to coaching. *The Clinical Neuropsychologist, 15,* 171-182.

Haines, M.E., & Norris, M.P. (2001). Comparing student and patient simulated malingerers' performance on standard neuropsychological measures to detect feigned cognitive deficits. *Clinical Neuropsychologist, 15,* 171-182.

Hartlage, L.C. (1998). Clinical detection of malingering. In C.R. Reynolds (Ed.), *Detection of malingering during head injury litigation: Critical issues in neuropsychology* (pp. 239-260). New York: Plenum.

Heaton, R.K., Smith, H.H., Lehman, R.A.W., & Vogt, A.T. (1978). Prospects for faking believable deficits on neuropsychological testing. *Journal of Consulting and Clinical Psychology, 46,* 892-900.

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

Jelicic, M., Hessels, A., & Merckelbach, H. (2006). Detection of feigned psychosis with the Structured Inventory of Malingered Symptomatology (SIMS): A study of coached and uncoached simulators. *Journal of Psychopathology and Behavioral Assessment, 28,* 19-22.

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. (2007). *Assessment of malingered neuropsychological deficits.* New York: Oxford University Press.

Merckelbach, H., & Smith, G.P. (2003). Diagnostic accuracy of the Structured Inventory of Malingered Symptomatology (SIMS) in detecting instructed malingering. *Archives of Clinical Neuropsychology, 18,* 145-152.

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.

Powell, M.R., Gfeller, J.D., Hendricks, B.L., & Sharland, M. (2004). Detecting symptom- and test-coached simulators with the Test of Memory Malingering. *Archives of Clinical Neuropsychology, 19,* 693-702.

Rogers, R. (2008). *Clinical assessment of malingering and deception.* New York: Oxford University Press.

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

Rosen, G.M. (1995). The Aleutian Enterprise sinking and posttraumatic stress disorder: Misdiagnosis in clinical and forensic settings. *Professional Psychology: Research and Practice, 26,* 82-87.

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.

Smith, G.P., & Burger, G.K. (1997). Detection of malingering: Validation of the Structured Inventory of Malingered Symptomatology. *Journal of the American Academy of Psychiatry and the Law, 25,* 183-189.

Suhr, J.A., & Gunstad, J. (2000). The effects of coaching on the sensitivity and specificity of malingering measures. *Archives of Clinical Neuropsychology, 15,* 415-424.

Suhr, J.A., & Gunstad, J. (2007). Coaching and malingering: A review. In G.J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 287-311). New York: Oxford University Press.

Teigner, G., & Wagner, M.T. (2004). The Test of Memory Malingering (TOMM): Normative data from cognitively intact, cognitively impaired, and elderly patients with dementia. *Archives of Clinical Neuropsychology, 19,* 455-464.

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

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

Youngjohn, J.R., Lees-Haley, P.R., & Binder, L.M. (1999). Comment: Warning malingerers produces more sophisticated malingering. *Archives of Clinical Neuropsychology, 14,* 511-515.

Table 1.

Mean Scores (SD) on the Test of Memory Malingering (TOMM) and Structured Inventory of Malingered Symptomatology (SIMS) of Participants in the Three Groups

	Control Participants (TOMM *n* = 30; SIMS *n* = 29)	Symptom-Coached Malingerers (TOMM *n* = 30; SIMS *n* = 29)	Symptom/Test-Coached Malingerers (TOMM *n* = 30; SIMS *n* = 28)
TOMM Trial 2	50.0 (0.2)	34.8 (8.0)	38.0 (7.4)
TOMM Retention Trial	50.0 (0)	32.1 (7.9)	36.3 (9.3)
SIMS	4.6 (2.7)	25.8 (6.8)	27.2 (9.8)

Table 2.

Number (and Percentage) of Participants in the Three Groups Correctly Classified by the Test of Memory Malingering (TOMM) and the Structured Inventory of the Malingered Symptomatology (SIMS) in the Three Groups

	Control Participants (TOMM *n* = 30; SIMS *n* = 29)	Symptom-Coached Malingerers (TOMM *n* = 30; SIMS *n* = 29)	Symptom/Test-Coached Malingerers (TOMM *n* = 30; SIMS *n* = 28)
TOMM Trial 2	30 (100%)	26 (87%)	24 (80%)
TOMM Retention Trial	30 (100%)	29 (97%)	24 (80%)
TOMM Trial 2 and Retention Trial combined	30 (100%)	29 (97%)	26 (87%)
SIMS	29 (100%)	27 (93%)	24 (86%)
The two TOMM trials and SIMS combined	29 (100%)	29 (100%)	26 (93%)