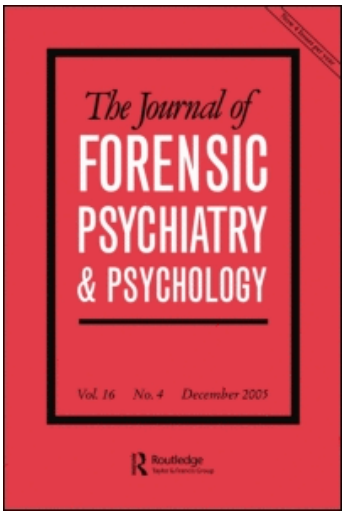

Journal of Forensic Psychiatry & Psychology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t714592861

Experimental simulation: type of malingering scenario makes a difference

Harald Merckelbach [a]; Tom Smeets [a]; Marko Jelicic [a]
[a] Faculty of Psychology, Maastricht University, Maastricht, The Netherlands

First Published:June2009

To cite this Article Merckelbach, Harald, Smeets, Tom and Jelicic, Marko(2009)'Experimental simulation: type of malingering scenario makes a difference',Journal of Forensic Psychiatry & Psychology,20:3,378 — 386

To link to this Article: DOI: 10.1080/14789940802456686

URL: http://dx.doi.org/10.1080/14789940802456686

The Journal of Forensic Psychiatry & Psychology
Vol. 20, No. 3, June 2009, 378–386

RESEARCH ARTICLE

Experimental simulation: type of malingering scenario makes a difference

Harald Merckelbach*, Tom Smeets and Marko Jelicic

Faculty of Psychology, Maastricht University, Maastricht, The Netherlands

(*Received 8 April 2008; final version received 29 August 2008*)

Using a simulation design, we tested in 486 undergraduates whether different case vignettes differentially affect the sensitivity of a self-report instrument designed to detect malingering (the Wildman Symptom Checklist). One group of participants ($n = 387$) was instructed to respond honestly, while three groups instructed to malinger (each $n = 33$) each received a different case vignette: in the manslaughter and the wrongful death vignettes malingering of symptoms was for the purpose of reducing criminal responsibility, while in the personal injury vignette malingering would lead to the acquisition of compensation money. The criminal case vignettes elicited malingering behaviour that was easier to detect than the civil case vignette, sensitivity rates being .82/.73 and .55, respectively. Thus, when testing the diagnostic qualities of such instruments it may be wise to rely on multiple case vignettes, including civil cases.

Keywords: malingering; simulated malingering; assessment; Wildman Symptom Checklist

Introduction

Research into the malingering of neurocognitive symptoms is blossoming (e.g., Boone, 2007; Delis & Wetter, 2007; Morel, 2008; Rogers & Payne, 2006). This is not surprising, as prevalence estimates indicate that the phenomenon is more widespread than once thought. For example, the seminal work by Mittenberg, Patton, Canyock, and Condit (2002) and more recently by Sullivan, Lange, and Dawes (2007) estimates that in personal injury cases referred for neuropsychological evaluation, the base rate of malingering is in the order of 13–29%, while in criminal cases the range is 17–19%.

Many studies have evaluated the diagnostic qualities of specialised instruments designed to detect malingering or – as it is sometimes called – the presentation of noncredible symptoms (Boone, 2007). Briefly, there are

*Corresponding author. Email: h.merckelbach@psychology.unimaas.nl

ISSN 1478-9949 print/ISSN 1478-9957 online

DOI: 10.1080/14789940802456686
http://www.informaworld.com

two classes of instrument: cognitive tasks designed to detect underperformance (e.g., the Amsterdam Short Term Memory Test or ASTM; Schmand et al., 1989), and self-report scales designed to detect overendorsement of symptoms (e.g., the Structured Inventory of Malingered Symptomatology or SIMS; Smith & Burger, 1997).

An important step towards establishing the diagnostic quality of these specialised tools is to conduct experimental simulation studies (Rogers, Harrell, & Liff, 1993). In such simulation designs, asymptomatic participants – usually undergraduate students – are asked to complete a malingering test either under instructions to respond honestly or under instructions to feign symptoms in a convincing manner. Several authors have pointed out that simulation designs have their limitations when it comes to generalising results to the clinical domain (Rogers & Cruise, 1998; Rogers et al., 1993). Therefore, when experimental simulation studies find that a new malingering test has promising diagnostic qualities, they should be followed up by so-called differential prevalence and known group studies (Rogers et al., 1993). Thus, experimental simulation studies are best viewed as a first, albeit important, step in the validation of malingering detection instruments.

An issue that has received little research attention in experimental studies is the precise nature of the scenario that is given to instructed malingerers (for an exception, see Rogers & Cruise, 1998). Studies differ widely in this respect. For example, some provided their experimental malingerers with general or vague instructions, such as: 'Assume that you are in a situation where it would benefit you greatly to appear mentally disturbed. Therefore, please respond so that you present yourself as someone with serious psychological problems' (Holden, Book, Edwards, Wasylkiw, & Starzyk, 2003, p. 1110). Other studies presented a civil case to their instructed malingerers. For example, Holmquist and Wanlass (2002, p. 147) asked their malingerers

> to pretend they were involved in litigation to determine how much financial compensation they would obtain from the party responsible for a motor-vehicle accident. They were told to imagine that they did not notice any difference in their mental or physical functioning as a result of the accident. However, the dissimulating subjects were to picture themselves as laborers presently working on a temporary assignment. They were to feel that they were justified in faking injuries and deserved all the money the courts would allow.

Yet other studies have presented a criminal scenario to their instructed malingerers. A case in point is the study by Smith and Burger (1997, p. 187), in which some participants were told 'to imagine themselves detained on a serious assault charge and to simulate a specific disorder to avoid a serious sentence'.

Many authors have noted that malingering is a rational strategy among people who face a difficult situation and who try to serve their own interests

in the best possible way (e.g., Rogers, 1997). One may wonder, therefore, whether the type of scenario with which instructed malingerers are provided would affect their performance. Consider the instructed malingerer who tries to identify with a suspect of manslaughter or wrongful death. In such criminal cases, the (instructed) malingerer may follow a strategy of reporting a broad range of symptoms so as to feign mental retardation or another serious condition that could support an argument of reduced criminal responsibility. Consider, on the other hand, the (instructed) malingerer who tries to fake personal injury symptoms (e.g., toxic encephalopathy) in a civil case. Here, the malingerer may want to appear as a reasonable and responsible person with a set of specific symptoms (e.g., neurological impairments). Therefore, the question arises: To what extent do different scenarios modulate the intensity with which symptoms are exaggerated and, in this way, affect the sensitivity of instruments designed to detect malingering?

The current study was a first attempt to look at this issue specifically. A large sample of undergraduates was given a malingering test encompassing several symptom dimensions. Most of them were asked to respond honestly to the test, while smaller groups were instructed to malinger on the basis of different scenarios.

Method

Participants

Our sample consisted of 486 first-year medical, health sciences, or psychology students (302 women) who participated in return for course credits. Their mean age was 20.1 years ($SD = 0.49$; range: 17–36). The study was approved by the Human Subjects Committee of the Faculty of Psychology.

Instruments

Participants completed the Wildman Symptom Checklist (WSC; Wildman & Wildman, 1999). This is a self-report scale designed to detect the overendorsement of symptoms and positive personal qualities. We selected this instrument because it is relatively unknown and there have been few attempts to validate it. In some respects, it is comparable to the more widely used SIMS (Smith & Burger, 1997). That is to say, both the WSC and the SIMS contain items that describe bizarre, atypical, or rare symptoms in several (e.g., psychological, neurological) domains. The respondent has to indicate whether or not each item applies to him/her. Unlike the SIMS, the WSC malingering items are embedded within a series of other items (e.g., 'I frequently experience headaches').

The WSC malingering items tap three different domains: 10 items list noncredible physical symptoms (e.g., 'The buzzing in my ears keeps

switching from the left to right'; 'I have allergies that I suffer from only at night'), 10 items describe noncredible mental symptoms (e.g., 'I have a terrible fear of street signs'; 'Someone is plotting to kill me and I know they will strike at midnight'), and 10 items describe 'fake good' tendencies (e.g., 'I never make a good story better'; 'I never find it difficult to talk to strangers'). Scores can be calculated for the physical, mental, and fake good subscales, separately (ranges: 0–10). A total malingering score can be obtained by summing across the subscales (range 0–30). Wildman and Wildman (1999) recommend that a total score of 4 or more is indicative of malingering.

In a pilot study, we administered both the WSC and the SIMS (Smith & Burger, 1997) to 149 undergraduates (95 women) and found that the WSC subscales correlated with each other, with r in the .52–.61 range. Also, the total malingering score on the WSC was significantly and positively correlated with endorsement of bizarre symptoms on the SIMS, although the correlation was modest ($r = .25$; $p < .01$).

Design and procedure

Participants were given one of four possible sets of instructions.

The control group ($n = 387$) was asked to respond honestly to the WSC items.

The manslaughter group ($n = 33$) was given a case vignette of 333 words describing a real Dutch case in which a young man trespassed on a medieval tower, thereby causing some of its stones to fall. One of the stones hit a young girl, who died instantly. The young man who climbed the tower was charged with manslaughter. Participants were asked to identify with the suspect and to malinger in a credible way a serious psychological condition so as to minimise criminal responsibility. Participants were asked to fill out the WSC with this in mind.

The wrongful death group ($n = 33$) was given a case vignette of 287 words about a real Dutch case in which a railroad worker failed to signal an incoming train, which then killed another railroad worker. The railroad worker was charged with wrongful death. Participants were instructed to imagine that they were the railroad worker who faced this charge, and to feign mental retardation so that their counsel could argue that the railroad company was responsible, having delegated safety measures to a person with insufficient cognitive capabilities. With these instructions in mind, participants in this group completed the WSC.

The personal injury group ($n = 33$) was given a 297-word summary of a real-life court case brought by a worker who had been a loyal employee for more than 25 years in a factory producing paint coatings. The worker had a conflict with his new manager, who wanted to fire him. The worker then refused to go back to work and instead filed a civil suit against the factory, claiming that he suffered from symptoms of chronic toxic encephalopathy.

Participants were asked to identify with the worker and to malinger in a plausible way complaints and symptoms that would support the claim while completing the WSC.

The instructed malingerer groups completed the WSC in small groups of 10–11 individuals. Several authors (e.g., Greve & Bianchini, 2004) have noted that when deriving diagnostic accuracy parameters from experimental simulation studies, the base rate of malingering is an important issue. Ideally, this base rate should reflect empirical estimates, which are in the 10–30% range. With this in mind, we wanted to recruit a relatively large control sample. Thus, for practical reasons, the control group was given the WSC during several mass testing sessions. Test sessions took about 25 minutes.

Results

Table 1 shows the mean scores of the four groups on the subscales of the WSC. It also shows the WSC total scores and the percentage of participants with total WSC scores exceeding the cutoff.

A one-way analysis of variance (ANOVA) indicated that the groups differed significantly in terms of their physical symptoms subscale scores ($F[3,482] = 60.5$, $p < .01$). Post-hoc Scheffé tests demonstrated that the control group had significantly lower subscale scores ($p < .01$) than the three malingering groups, which did not differ from each other (all $p > .08$). A different pattern was found for the mental symptoms subscale. Here, again, the overall ANOVA indicated that there were significant group differences ($F[3,482] = 66.4$, $p < .01$), but this time both the control group and the personal injury group had significantly lower scores than the two other malingering groups (both $p < .01$). As for the fake good scale, there were significant group differences ($F[3,482] = 26.5$, $p < .01$), with the control group scoring lower than the three malingering groups (all $p < .02$), which did not differ from each other (all $p > .15$). Most importantly, an ANOVA performed on the WSC total score also

Table 1. WSC scores of the control group ($n = 387$), the manslaughter group ($n = 33$), the wrongful death group ($n = 33$), and the personal injury group ($n = 33$). Standard deviations are given between parentheses.

	Physical symptoms (range 0–10)	Mental symptoms (range 0–10)	Fake good score (range 0–10)	Total score (range 0–30)	% ≥ 4
Control	0.74 (1.00)	0.34 (0.76)	1.08 (1.18)	2.17 (2.07)	16
Manslaughter	3.03 (2.04)	2.16 (2.08)	2.67 (2.20)	7.85 (5.49)	73
Wrongful death	2.46 (1.99)	2.39 (2.01)	2.67 (2.06)	7.52 (4.93)	82
Personal injury	2.24 (1.63)	0.72 (0.80)	1.88 (1.60)	4.85 (3.16)	55

revealed significant group differences, with both the control group and the personal injury group having lower overall scores than the two criminal case groups (all $p < .01$).

That participants instructed to malinger personal injury complaints in the context of a civil case are more difficult to differentiate from controls than participants instructed to malinger symptoms in the context of a criminal case is also evident when looking at the percentage of participants scoring above the cutoff. This percentage was significantly lower in the personal injury group than in the two other malingering groups ($\chi^2[2] = 6.03$, $p < .05$). Thus, with a specificity of 0.84 (i.e., 84% of control participants were correctly classified; see Table 1), the WSC was found to have a sensitivity between a low .55 and a high .82, depending on the precise instructions given to the malingerers.

Discussion

According to the widely accepted adaptational account, malingering is not a disorder, but a strategic response to adversarial circumstances (e.g., Rogers, 1997). A straightforward implication of this for simulation designs is that instructed malingerers can only engage in optimal strategic behaviour to the extent that they know the details of the adversarial circumstances. Therefore, it seems reasonable to expect malingering behaviour to depend on the way in which adversarial circumstances are portrayed to the instructed malingerers. Curiously enough, very few studies have looked at this in a systematic fashion. One is a study by Rogers and Cruise (1998) in which the researchers combined several types of scenario with positive or negative incentives. The authors found that incentives do make a difference: negative incentives (e.g., revealing the names of the 10 least successful malingerers) produced more focused forms of malingering.

The current study investigated whether malingering performance would be affected by the adversarial circumstances portrayed. We found that this is, indeed, the case. More specifically, a civil case scenario, in which the main character filed a suit complaining of toxic encephalopathy, elicited significantly lower endorsement rates on a test of malingering than did criminal case scenarios involving a suspect charged with manslaughter or wrongful death. In other words, the civil case vignette produced a less intense and therefore more difficult to detect form of malingering than the criminal case vignettes. The sensitivity of the malingering test dropped from .82/.73 for malingerers given the criminal case vignettes to .55 for malingerers given the civil case vignette.

One obvious explanation for this dramatic drop in sensitivity is that unlike the civil context, the criminal context induces the belief that only a dramatic presentation of symptoms and virtuous features will convince

others that one is not criminally responsible. This is in keeping with literature showing that the most bizarre and atypical symptoms can be found in criminal contexts where suspects face serious charges (Jaffe & Sharma, 1998). On the other hand, in the civil context, the malingerer does not want triers of fact to regard his/her presentation of symptoms and complaints as dubious because he/she is mentally confused.

One limitation of our study is that we did not manipulate legal context (e.g., civil versus criminal) and feigned symptoms (e.g., mental retardation versus 'painter's disease') in a completely orthogonal way. Future studies might look specifically at the sensitivity of instruments in detecting similar feigned symptoms across both contexts (e.g., simulated post-traumatic stress symptoms; Morel, 2008).

A second limitation of the current study is that we did not conduct exit interviews with our instructed malingers. This would have provided us with more direct information about how malingerers adapt their strategy to the details of the case. Clearly, this issue warrants further study.

Another limitation of our study is that we administered only a single test of malingering. The WSC is relatively unknown, and we selected it to illustrate our point that when validating new or unknown measures it is wise to include various instructed malingering conditions. Nevertheless, it would be informative to replicate this study with well-known malingering measures, including those that assess cognitive performance, like the ASTM (Schmand et al., 1989).

Thus, we consider our findings preliminary in the sense that they show that the features of a case vignette may affect the sensitivity of a malinger instrument. Future studies should look at the issue in a more systematic fashion, preferably using multiple tests of malingering.

Our finding that the sensitivity of a test of malingering can be easily manipulated has three implications. The first and most important is that clinicians should not blindly trust the sensitivity figures reported in test manuals, but critically evaluate the way in which these figures were obtained. Of course, even with further information about their sensitivities, malingering instruments should be employed with caution, particularly in legal settings. Their use requires expertise in the psychometrics of test batteries, and also clinical experience with diagnoses that superficially resemble malingering (e.g., somatoform disorders). The second and more methodological implication is that simulation designs should preferably rely on multiple detailed case vignettes, including one of a civil case, when instructing subjects to feign malingering. When testing and reporting the diagnostic quality of instruments to test for malingering, it is wise to include case vignettes that produce the most prudent lower-bound estimates of a test's sensitivity and specificity. A third implication is that malingering is, indeed, strategic behaviour that sometimes takes the form of positive symptomatology (i.e., bizarre symptoms) combined with

feigned impairments (i.e., pseudo-neurological dysfunctions) and at other times restricts itself to the latter type of responding (see also Kuperman, 2006).

More systematic research into the positive and negative manifestations of malingering is important – this could inform clinicians about which tests, in which contexts, can be safely included in a multi-strategy multi-method approach (e.g., Rogers & Payne, 2006) to detect malingering.

References

Boone, K.B. (2007). A reconsideration of the Slick et al. (1999) criteria for malingered neurocognitive dysfunction. In K.B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 29–49). New York: Guilford.

Delis, D.C., & Wetter, S.R. (2007). Cogniform disorder and cogniform condition: Proposed diagnoses for excessive cognitive symptoms. *Archives of Clinical Neuropsychology*, *22*, 589–604.

Greve, K.W., & Bianchini, K.J. (2004). Setting empirical cut-offs on psychometric indicators of negative response bias: A methodological commentary with recommendations. *Archives of Clinical Neuropsychology*, *19*, 533–541.

Holden, R.R., Book, A.S., Edwards, M.J., Wasylkiw, L., & Starzyk, K.B. (2003). Experimental faking in self-reported psychopathology: Unidimensional or multidimensional? *Personality and Individual Differences*, *35*, 1107–1117.

Holmquist, L.A., & Wanlass, R.L. (2002). A multidimensional approach towards malingering detection. *Archives of Clinical Neuropsychology*, *17*, 143–156.

Jaffe, M.E., & Sharma, K.K. (1998). Malingering uncommon psychiatric symptoms among defendants charged under California's 'three strikes and you're out' law. *Journal of Forensic Sciences*, *43*, 549–555.

Kuperman, V. (2006). Narratives of psychiatric malingering in works of fiction. *Journal of Medical Ethics*, *32*, 67–72.

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.

Morel, K.R. (2008). Development of a validity scale for combat-related posttraumatic stress disorder: Evidence from simulated malingerers and actual disability claimants. *Journal of Forensic Psychiatry and Psychology*, *19*, 52–63.

Rogers, R. (1997). Introduction. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 1–19). New York: Guilford.

Rogers, R., & Cruise, K.R. (1998). Assessment of malingering with simulation designs: Threats to external validity. *Law and Human Behavior*, *22*, 273–285.

Rogers, R., Harrell, E.H., & Liff, C.D. (1993). Feigning neuropsychological impairment: A critical review of methodological and clinical considerations. *Clinical Psychology Review*, *13*, 255–274.

Rogers, R., & Payne, J.W. (2006). Damages and rewards: Assessment of malingered disorders in compensation cases. *Behavioral Sciences and the Law*, *24*, 645–658.

Schmand, B., Lindeboom, J., Schagen, S., Heijt, R., Koene, T., & Hamburger, H.L. (1989). 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 Inventory of Malingered Symptomatology. *Journal of the American Academy of Psychiatry and the Law*, *25*, 183–189.

Sullivan, K., Lange, R.T., & Dawes, S. (2007). Methods of detecting malingering and estimated symptom exaggeration base rates in Australia. *Journal of Forensic Neuropsychology*, *4*, 49–70.

Wildman, R.W., & Wildman, R.W. (1999). The detection of malingering. *Psychological Reports*, *84*, 386–388.