



# Symptom overreporting obscures the dose–response relationship between trauma severity and symptoms

Harald Merckelbach<sup>a,\*</sup>, Willie Langeland<sup>b</sup>, Gerard de Vries<sup>c</sup>, Nel Draijer<sup>b</sup>

<sup>a</sup> Faculty of Psychology and Neuroscience, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands

<sup>b</sup> Department of Psychiatry, and EMGO+ Institute, Free University Medical Center, Amsterdam, The Netherlands

<sup>c</sup> Scientific Council For Government Policy, The Hague, The Netherlands

## ARTICLE INFO

### Article history:

Received 6 September 2013

Received in revised form

27 January 2014

Accepted 14 March 2014

### Keywords:

Symptom overreporting

Trauma

Self-reports

Brief Symptom Inventory

## ABSTRACT

We investigated whether symptom overreporting affects the dose–response relationship between self-reported abuse severity and psychiatric symptoms in two samples. The first sample ( $N=599$ ) consisted of adults who had previously reported to a public commission that they had been witnesses to or victims of childhood sexual abuse by Roman Catholic Church representatives. The second sample ( $N=1756$ ) consisted of general population respondents who indicated that they had been victims of non-familial childhood sexual abuse. Using a web-based data collection procedure, both samples completed the Brief Symptom Inventory (BSI-18), items addressing abuse severity, and items flagging symptom overreporting. Adjusting for overreporting reduced the proportion of participants with clinically raised BSI-18 scores from 60% to 47% in sample 1 and from 26% to 22% in sample 2. Also, in both samples, normal range reporting participants exhibited the typical dose–response relationship between trauma severity and BSI-18 scores, whereas this pattern was largely non-significant in overreporting participants. Our findings show that symptom overreporting has a psychometric impact that may obscure relationships between clinically relevant variables and should therefore preferably be monitored in surveys.

© 2014 Elsevier Ireland Ltd. All rights reserved.

## 1. Introduction

Psychiatry research often has to rely on scales measuring participants' subjective reports about their symptoms. Such self-reports may be vulnerable to distorted responses. There are two types of distortions (Meade and Craig, 2012). The first type of distortion is independent of the questionnaire content and consists of “yea-saying”, “nay-saying” and/or inattentive responding due to, for example, participants' lack of interest when filling out a lengthy test. The second is content dependent and consists of underreporting or overreporting symptoms. Underreporting may be motivated by participants' reluctance to endorse symptoms because of the stigma that surrounds such symptoms, whereas overreporting may be motivated by the prospect of advantages, such as sympathy, attention, and incentives (Berry et al., 2008).

The current study focuses on overreporting of symptoms in survey respondents. In an influential paper, McGrath et al. (2010) argued that the importance response biases such as overreporting is often overestimated and that their distorting effect on research

outcomes is limited. Still, neuropsychological studies show that symptom overreporting is typically pronounced in people who anticipate incentives (Iverson, 2006; Stevens et al., 2008). These incentives may involve legal benefits or monetary compensation, yet striving for recognition of one's status as a patient or victim can also be considered as a form of what is often termed “secondary gain” (Shapiro et al., 2013).

One way to adjust for symptom overreporting is to include items that allude to absurd symptoms. The idea here is that people who endorse such bogus items might overstate their mental health problems (Lanyon, 2003; Cooper et al., 2011). Once this group has been identified, researchers can adjust for the contribution of symptom overreporting. An example of this approach is provided by Wiggins et al. (2012). These authors employed validity indicators of the Minnesota Multiphasic Personality-2 Restructured Form (MMPI-2 RF) to identify symptom overreporting in a sample of disability litigants. The meaningful correlations that are typically found between clinical scales disappeared in the overreporting subgroup. Accordingly, Wiggins et al. (2012; p. 170) concluded that “response bias weakens our ability to describe clinical functioning and predict various clinical constructs.”

In the current study, we tested whether adjusting for symptom overreporting would affect the typical dose–response relationship

\* Corresponding author. Tel.: +31 43 3881945; fax: +31 43 3884196.

E-mail address: [h.merckelbach@maastrichtuniversity.nl](mailto:h.merckelbach@maastrichtuniversity.nl) (H. Merckelbach).

between self-reported trauma severity and mental health problems (McNally and Robinaugh, 2011). We also wanted to know whether adjusting for symptom overreporting would suppress prevalence estimates of the number of individuals with clinically raised symptom levels. We investigated these issues in two separate samples that differed in secondary gain expectations. We predicted that they would also differ in the prevalence of symptom overreporting.

## 2. Methods

### 2.1. Participants

We collected survey data from the archival sources of a public commission chaired by Dutch Council of State member Wim Deetman (Deetman et al., 2011). With help of the Dutch survey agency TNS NIPO, his commission gathered survey data in 2010 and 2011. The commission was installed by the Dutch Roman Catholic Church (DRCC), which also financed the investigations of the commission. However, the commission operated independently of the DRCC and the DRCC did not have any influence on the final report of the commission, or on the content of the current article (<http://www.onderzoekrkr.nl/home.html>).

The current paper focuses on two samples. The first sample consisted of adults who themselves had contacted the commission by telephone, letter or e-mail, reporting that as a child or adolescent, they had witnessed or experienced abuse by Roman Catholic Church representatives. These reports were made against the background of an intense public debate about sex abuse crimes in the churches. The survey agency made attempts to approach all these persons ( $N=883$ ) by e-mail with the request to fill out a set of web-based questionnaires. In 47 cases (5%), the survey agency was unable to locate an e-mail address, while in 42 cases (5%) the e-mail address proved to be dysfunctional. In 30 cases (3%), respondents explicitly reacted negatively, saying that they refused to cooperate. In 16 cases (2%), respondents said they were unable to complete the items before the deadline because they were on vacation or were ill. All in all, 68% (599) respondents agreed to participate. The large majority of them (505; 84%) reported that as a child, they had been victimized by perpetrator(s) related to the church. As it was widely known that the commission was an investigative commission and was not involved in financial compensation procedures, it is unlikely that respondents in sample 1 anticipated monetary incentives, although many of them wanted to be recognized as abuse victims.

The second sample was selected from one of the largest online community panels in the Netherlands comprising 145,785 Dutch citizens from 60,412 households. Recruitment involved two steps. To make the age distribution as similar as possible to that of sample 1, the survey agency included in a first step only those members of the panel aged 40 years and older ( $N=65,536$ ) and asked them to fill out a series of items, among which a general item about non-familial sexual abuse. In the second step, a stratified sample ( $N=2812$ ) of those who responded positively or negatively to this question, were invited by e-mail to participate in a follow-up web-based survey that consisted of the same questionnaires and items administered to sample 1 (see below). Stratification involved oversampling and was based on age, religious background, non-familial abuse, and educational career. Of the stratified sample, 91% was willing to participate in the second step. Below, we focus on the subsample that previously responded affirmatively to the sexual abuse item ( $N=1756$ ). The sampling methods are described in detail elsewhere (Deetman et al., 2011).

Table 1 summarizes demographic information for the two samples. The samples differed in several respects. Firstly, the sample recruited from reports made to the commission (sample 1) comprised more men than the sample

recruited through the general population panel (sample 2), proportions being 81% versus 38% ( $\chi^2(1)=344.5, p<0.01$ ). Secondly, sample 1 was somewhat older than sample 2 ( $t(2353)=6.82, p<0.01$ ), with the 95% confidence interval (CI) of the difference ranging from 2.2 to 3.9 years. Thirdly, sample 1 had more fully employed respondents than sample 2 ( $\chi^2(1)=73.9, p<0.01$ ). Fourthly, participants in sample 1 more often said that they had informed the police about the abuse than respondents in sample 2 ( $\chi^2(1)=90.0, p<0.01$ ). Fifthly, sample 1 respondents more often reported that they had sought psychological help for the emotional sequelae of sexual abuse than sample 2 respondents ( $\chi^2(1)=834.3, p<0.01$ ).

### 2.2. Measures

Both samples completed the Dutch version of the 18-item Brief Symptom Inventory (BSI-18; Cronbach's  $\alpha$  sample 1=0.95; Cronbach's  $\alpha$  sample 2=0.89), a widely used instrument specifically designed to screen for psychiatric distress (Derogatis, 2000; De Beurs, 2011). Its items refer to anxiety, depression, and somatization symptoms and respondents indicate on 5-point scales (anchors: 0=not at all; 4=always) to what extent they experienced these symptoms in the past week. We calculated a total score by summing across items (range: 0–72). Following Dutch normative data (De Beurs, 2011), we employed a cut-off of 11 to estimate the prevalence of individuals with clinically significant levels of distress. We also calculated BSI subscale scores for anxiety (Cronbach's  $\alpha$ 's: 0.81; 0.92), depression (Cronbach's  $\alpha$ 's: 0.82, 0.92), and somatization (Cronbach's  $\alpha$ 's: 0.69; 0.85).

Abuse severity was assessed with five questions about the characteristics of the abusive event. Respondents were asked whether there had been one or more perpetrators (scored as 1 and 2), whether they had been victims of penetrative abuse or other types of abuse such as unwanted sexual touching or sexual assault (scored as 1 and 0), how often the abuse had taken place (with scores ranging from 0=don't know to 5=very often), whether they had been threatened by the perpetrator(s) (0=no, 1=yes), and whether or not the abuse had spanned a longer period of time (ranging from 0=don't know to 3=longer than a year). Scores on these items were summed to generate an abuse severity composite (range: 1–12), with higher scores reflecting more severe abuse reports.

Respondents were also given four items from the Wildman Symptom Checklist that addresses non-credible symptoms (Wildman and Wildman, 1999). In a previous study (Merckelbach et al., 2008), the selected items were found to discriminate optimally (at least a difference in endorsement rate of 30%) between honest responders and participants instructed to exaggerate symptoms. These items were: "I have headaches that are so severe my feet hurt"; "The buzzing in my ears keeps switching from the left to the right"; "I notice that the color of objects around me keeps shifting"; and "I find myself frequently blacking out when I sit down." The non-credible symptom items were rated on a five-point scale (anchors: 0=not at all, 4=extremely) that closely resembled that of the BSI-18. To obtain a symptom overreporting index, the scores were summed across the four items (range: 0–16). We used a threshold score of 4 to identify respondents with an overreporting response style. Scores above this cut-off reflect affirmative answers of some degree to at least two non-credible items, a pattern that is unlikely to be the result of mere incidental report errors.

### 2.3. Procedure

Respondents were given 20 days to complete the online questionnaires. They were not paid for this. One week before the deadline, a reminder was sent to those who had not filled out the questionnaires. The non-credible symptoms were interspersed among the BSI-18 items (see for a similar strategy: Cooper et al., 2011). All respondents first completed the symptoms items and then the items about the abuse characteristics. They were told that their scores would be processed anonymously. Completion of the questionnaires took about 30 min. Ethics approval was obtained from the standing human subjects committee of the Faculty of Psychology and Neuroscience, Maastricht University, The Netherlands.

### 2.4. Data analyses

Using  $t$ -tests and  $\chi^2$  tests (in the case of categorical data) and associated 95% CI's for differences between means and for Odds Ratio's (OR's), we compared the two samples with regard to BSI-18 scores, abuse severity, and symptom overreporting. We also contrasted normal range (i.e., non-overreporting) participants of the two samples. For both samples, we calculated Pearson product-moment correlations between the abuse severity index and BSI-18 indices in overreporting and normal range respondents, separately. Differences in correlational strength were evaluated with Fisher's  $Z$ .

## 3. Results

Table 2 summarizes psychometric scores of samples 1 and 2. Sample 1 had higher scores on BSI-18 indices, abuse severity, and

**Table 1**  
Demographics of samples.

	Sample 1 $N=599$	Sample 2 $N=1756$
Men/women*	486/113	654/1102
Age (S.D., range)*	60.1 (7.6; 40–87)	57.0 (10.2; 40–85)
Education <sup>a</sup>		
University degree (%)	105 (18)	295 (17)
Some college degree (%)	469 (78)	1405 (80)
< 6 years of education (%)	18 (3)	56 (3)
Employed (%)*	227 (38)	357 (20)
Informed police (%)*	40 (7)	7 (0.4)
Sought help (%)*	330 (55)	68 (4)

\*  $p < 0.05$ .

<sup>a</sup> Seven missing values in sample 1.

symptom overreporting than sample 2 (all  $t$ 's (2353) > 10.3, all  $p$ 's < 0.01). Accordingly, the associated CI's for group differences never included zero. Furthermore, the proportion of respondents in sample 1 with clinically raised BSI-18 total scores was higher than that in sample 2 ( $\chi^2$  (1) = 228.2,  $p$  < 0.01, OR = 4.2). Also, the proportion of participants with an overreporting style was higher in sample 1 than sample 2 ( $\chi^2$  (1) = 62.9,  $p$  < 0.01, OR = 3.7).

In sample 1, 76 respondents (13%; 95% CI: 10–16%) displayed an overreporting style (defined in terms of scores > 4 on the symptom overreporting index). All of them also had BSI-18 Total scores above the cut-off of 11 employed to identify raised levels of distress. Taking only the normal range respondents (scores ≤ 4 on the symptom overreporting index) into account reduced the rate of cases with raised BSI-18 levels from 358 (60%) to 282 (54% of the normal range subsample, 47% of the full sample). In sample 2, 66 respondents (4%; 95% CI: 3–5%) exhibited an overreporting style. Of these, 60 respondents also had a BSI-Total score exceeding the cut-off of 11. Looking only at the combination of high BSI-18 Total scores and normal range responding reduced the estimated rate of clinically raised levels of distress from 453 (26%) to 393 (23% of normal range subsample, 22% of full sample) respondents.

When overreporting participants were removed from both samples, sample 1 still had higher scores on BSI-18 indices, abuse severity, and symptom overreporting than sample 2 (all  $t$ 's (2211) > 6.1, all  $p$ 's < 0.01). Sample 1 also continued to have more individuals with clinically raised BSI-18 levels ( $\chi^2$  (1) = 177.2,  $p$  < 0.01, OR = 3.9). However, as can be seen in the right panel of Table 2, except for abuse severity, confidence intervals were consistently smaller compared with confidence intervals relying on the full sample (left panel).

In both (full) samples, symptom overreporting (as a continuous variable) correlated significantly with BSI-18 parameters. Correlations were in the 0.59–0.71 range in sample 1 and in the 0.45–0.58 range in sample 2 (all  $p$ 's < 0.01). The correlational strength between symptom overreporting and trauma severity was much lower (sample 1:  $r$  = 0.19,  $p$  < 0.01; sample 2:  $r$  = 0.16,  $p$  < 0.01), but still significant due to the large samples sizes.

Table 3 gives Pearson product-moment correlations between abuse severity and BSI-18 scales for normal responding and overreporting subgroups within each sample. The correlations between abuse severity and symptomatology were significant and followed a dose–response relationship in normal range respondents, but were largely non-significant in those with an overreporting style. Differences between the correlations in the two groups were significant in sample 1, but fell short of significance in sample 2.

#### 4. Discussion

The main results of the current study can be summarized as follows. Firstly, using the criterion of affirmative responses to at least two of four non-credible symptoms, we found a symptom

overreporting style in 13% of sample 1, the sample in which secondary gain expectations (i.e., societal recognition of victim status) were arguably strongest. In the community sample (sample 2) such expectations will have played no role. In line with this, an overreporting style was less prevalent (4%) in that sample.

Secondly, we found that overreporting has a profound psychometric impact. Thus, differences between the two samples, though still significant, generally became smaller when overreporting participants were excluded from the samples.

Thirdly, the typical dose–response relationship between trauma severity and psychopathology was largely non-significant in the overreporting groups of both sample. That is, in overreporting participants, less severe forms of self-reported trauma were not accompanied by lower levels of psychopathology. The difference in dose–response relationships between normal range and overreporting participants was more pronounced in sample 1 than in sample 2 because sample 1 comprised both more participants who had experienced the most severe forms of abuse and more participants who engaged in overreporting. Our finding that overreporting obscures dose–response relationships is reminiscent of Wiggins et al. (2012) who observed that apriori plausible associations between clinical variables are present in normal range respondents, but disappear in those with an overreporting style.

The first-order correlations between trauma severity and psychopathology that we observed in the normal range respondents in both sample (i.e., 0.13–0.24) replicates the magnitudes commonly reported in the literature. For example, in their community sample of women with a history of childhood sexual abuse ( $N$  = 102), McNally and Robinaugh (2011) found a correlation of 0.21 ( $p$  = 0.02) between self-reported abuse severity and PTSD symptoms, a pattern that is consistent with dose-dependent mental health effects of trauma intensity.

**Table 3**

Pearson product-moment correlations with abuse severity for non-overreporting and overreporting respondents in both samples.

	Non-overreporting $n$ = 523	Overreporting $n$ = 76	Fisher's $Z$
Sample 1			
BSI total	0.23*	−0.05	2.27*
BSI anxiety	0.24*	−0.04	2.28*
BSI depression	0.19*	−0.04	1.86*
BSI somatization	0.17*	−0.05	1.77*
Sample 2			
BSI total	0.17*	0.22	0.41
BSI anxiety	0.15*	0.21	0.48
BSI depression	0.14*	0.30*	1.31
BSI somatization	0.13*	0.00	1.02

\*  $p$  < 0.05 (one-tailed).

**Table 2**

Mean scores of full samples and non-overreporting participants on BSI-18 indices, abuse severity, and symptom overreporting (SO). 95% Confidence Intervals (CI) for differences between sample means and for OR's are also shown.

	Sample 1 $N$ = 599	Sample 2 $N$ = 1756	95% CI	Sample 1 non-over reporting $n$ = 523	Sample 2 non-over reporting $n$ = 1690	95% CI
BSI total	19.2 (16.0)	8.7 (9.6)	9.4–11.6	15.8 (13.2)	8.0 (8.5)	6.8–8.8
BSI > 11 (%) <sup>a</sup>	358 (60)	453 (26)	3.5–5.2	282 (54)	393 (23)	3.1–3.9
BSI anxiety	6.8 (6.1)	2.8 (3.7)	3.6–4.4	5.6 (5.2)	2.5 (3.3)	2.7–3.5
BSI depression	7.6 (6.5)	3.2 (4.1)	3.9–4.9	6.4 (5.7)	2.9 (3.8)	3.1–3.9
BSI Somatization	4.9 (5.0)	2.8 (3.2)	1.8–2.5	3.8 (4.0)	2.5 (2.9)	1.0–1.6
Abuse severity	6.0 (3.3)	3.8 (2.6)	1.9–2.5	5.8 (3.2)	3.7 (2.5)	1.8–2.4
SO index	1.6 (2.6)	0.7 (1.5)	0.7–1.1	0.8 (1.2)	0.5 (0.9)	0.2–0.4
SO index > 4 (%) <sup>a</sup>	76 (13)	66 (4)	2.6–5.3	–	–	–

<sup>a</sup> 95 CI's in these rows pertain to OR's.



Several studies noted that a subgroup of respondents with a self-reported history of trauma exhibits a tendency to fail on symptom validity measures (Rogers et al., 2009; Williamson et al., 2012). Together with the current results, this implies that a measure of symptom overreporting is crucial when researchers employ self-reports to investigate the link between trauma severity and mental health problems (Cooper et al., 2011). It is precisely because they lack validity scales that instruments such as the BSI-18 are vulnerable to symptom exaggeration. This might be, as our data show, particularly problematic when respondents anticipate secondary gains. Self-report measures of psychopathology with embedded validity indices addressing symptom overendorsement are available (Mogge et al., 2008). Researchers might be well advised to employ these instruments or to embed response validity indices in self-report instruments that lack such validity scales (Cooper et al., 2011).

Our results do *not* imply that respondents who consistently endorse bogus items such as “I have headaches that are so severe my feet hurt” are completely fabricating their current health problems and/or traumatic history. These respondents might have a victimization history and they might magnify genuine symptoms (Rogers et al., 2009). Our point is that when a person repeatedly endorses bogus symptoms, researchers can no longer take his or her self-reported mental health problems for granted (Iverson, 2006; Stevens et al., 2008; Freeman et al., 2008). In studies that intend to yield precise estimates, symptom overreporting should not be ignored, particularly not because it might obscure correlations between background characteristics (e.g., abuse severity), clinical variables (e.g., symptoms, treatment outcome), and even biological parameters (Rienstra et al., 2013). Measuring symptom overreporting allows researchers to present statistics in a way that adjust for this type of response bias, thereby contributing to the precision of the data sets (Rosen, 2003; Freeman et al., 2008; Rohling et al., 2011). As the current study shows, this is especially important when relying on self-report data of respondents who may have a motive to obtain societal recognition.

Our findings should be interpreted in the context of the following limitations. Firstly, non-response was higher in sample 1 than in sample 2 (32% versus 9%). To a large extent, this differential non-response relates to the fact that sample 2 was drawn from an existing online panel, while sample 1 was drawn from a group in which people with no internet connection were overrepresented. Thus, our web-based method might have introduced selective non-response in the sense that less educated people with no reliable access to internet or no experience in how to use it, may have been overrepresented in non-responders. However, in general, web-based data collection methods yield high quality data (Shapiro et al., 2013). Furthermore, we have good reasons to assume that self-reported trauma severity was not higher in non-responders. As a matter of fact, the reverse appears to be true: the proportion of those who reported penetrative abuse was 15% in the group that had contacted the commission, while it was 23% and 8% in sample 1 and 2, respectively.

Secondly, the survey character of our study made it necessary to rely on brief indicators. Had we included lengthy measures, we might have induced reactive forms of response distortion reflecting respondents' annoyance at the large number of items (Meade and Craig, 2012). Thirdly, we focused on symptom overreporting. We did not address underreporting, i.e., the tendency to minimize symptoms and adverse life events due to, for example, reluctance to discuss sensitive topics (Goodman et al., 2003).

A fourth limitation is that we only included respondents who reported being the victim of non-familial abuse. It remains to be seen whether our findings can be generalized to respondents who report being victims of familial sexual abuse. Perhaps, symptom overreporting and symptom underreporting play differential roles

in both categories of victims, with symptom underreporting due to feelings of shame being more pronounced in familial abuse groups and symptom overreporting being more pronounced in non-familial abuse groups that seek official recognition. This issue warrants systematic research.

Rogler et al. (2001) reviewed major psychiatric journals and concluded that the overwhelming majority of self-report studies that had appeared in these journals did not address the problem of distorted symptom responding. Our study illustrates that this is unfortunate because symptom overreporting is a non-trivial phenomenon, particularly in samples where there might be secondary gain expectations (Stevens et al., 2008; Cooper et al., 2011). Although it is evident that childhood sexual abuse has mental health implications (Chen et al., 2010), our results caution against the overly optimistic reliance on self-reports and demonstrate that survey research in this domain may benefit from symptom overreporting indicators.

## Acknowledgments

This work was supported by a grant (No 16564) from The Netherlands Organization for Health Research and Development (ZonMw, The Hague).

## References

- Berry, D.T.R., Sollman, M.J., Schipper, L.J., Clark, J.A., Shandera, A.L., 2008. Assessment of feigned psychological symptoms. In: Butcher, J. (Ed.), *Oxford Handbook of Personality Assessment*. Oxford University Press, New York, pp. 613–642.
- Chen, L.P., Murad, M.H., Paras, M.L., Colbenson, K.M., Sattler, A.L., Goranson, E.N., Elamin, M.B., Seime, R.J., Shinozaki, G., Prokop, L.J., Zirakzadeh, A., 2010. Sexual abuse and lifetime diagnosis of psychiatric disorders: systematic review and meta-analysis. *Mayo Clinic Proceedings* 85, 618–629.
- Cooper, D.B., Nelson, L., Armistead-Jehle, P., Bowles, A.O., 2011. Utility of the Mild Brain Injury Atypical Symptom Scale as a screening measure for symptom overreporting in Operation Enduring Freedom/Operation Iraqi Freedom service members with post-concussive complaints. *Archives of Clinical Neuropsychology* 26, 718–727.
- De Beurs, E., 2011. Brief Symptom Inventory (BSI) en BSI 18. Handleiding 2011. PITS, Leiden, The Netherlands.
- Deetman, W., Draijer, N., Kalbfleisch, P., Merckelbach, H., Monteiro, M., de Vries, G., 2011. Seksueel misbruik van minderjarigen in de Rooms-Katholieke Kerk. Balans, The Hague, The Netherlands.
- Derogatis, L.R., 2000. BSI-18: Administration, Scoring, and Procedures Manual. National Computer Systems, Minneapolis, MN.
- Freeman, T., Powell, M., Kimbrell, T., 2008. Measuring symptom exaggeration in veterans with chronic posttraumatic stress disorder. *Psychiatry Research* 158, 374–380.
- Goodman, G.S., Ghetti, S., Quas, J.A., Edelstein, R.S., Alexander, K.W., Redlich, A.D., Cordon, I.M., Jones, D.P.H., 2003. A prospective study of memory for child sexual abuse: new findings relevant to the repressed-memory controversy. *Psychological Science* 14, 113–118.
- Iverson, G.L., 2006. Ethical issues associated with the assessment of exaggeration, poor effort, and malingering. *Applied Neuropsychology* 13, 77–90.
- Lanyon, R.L., 2003. Assessing the misinterpretation of health problems. *Journal of Personality Assessment* 81, 1–10.
- Meade, A.W., Craig, S.B., 2012. Identifying careless responses in survey data. *Psychological Methods* 17, 437–455.
- McGrath, R.E., Mitchell, M., Kim, B.H., Hough, L., 2010. Evidence for response bias as a source of error variance in applied assessment. *Psychological Bulletin* 136, 450–470.
- McNally, R.J., Robinaugh, D.J., 2011. Risk factors and posttraumatic stress disorders: are they especially predictive following exposure to less severe stressors? *Depression and Anxiety* 28, 1091–1096.
- Merckelbach, H., Smeets, T., Jelicic, M., 2008. Onwaarschijnlijke symptomen simuleren: De Wildman Symptom Checklist. *Neuropsychologie* 12, 53–57.
- Mogge, N.L., Steinberg, J.S., Fremouw, W., Messer, J., 2008. The Assessment of Depression Inventory (ADI): an appraisal of validity in an outpatient sample. *Depression and Anxiety* 25, 64–68.
- Rienstra, A., Groot, P.F.C., Spaan, P.E.J., Majoie, C.B.L.M., Nederveen, A.J., Walstra, G.J.M., De Jonghe, J.F.M., van Gol, W.A., Olabarriaga, S.D., Korkhov, V.V., Schmand, B., 2013. Symptom validity testing in memory clinics: hippocampal-memory associations and relevance for diagnosing mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology* 35, 59–70.

- Rogers, R., Payne, J.W., Correa, A.A., Gillard, N.D., Ross, C.A., 2009. A study of the SIRS with severely traumatized patients. *Journal of Personality Assessment* 91, 429–438.
- Rogler, L.H., Mroczek, D.K., Fellows, M.M.A., Loftus, S.T., 2001. The neglect of response bias in mental health research. *Journal of Nervous and Mental Disease* 189, 182–187.
- Rohling, M.L., Larrabee, G.J., Greiffenstein, M.F., Ben-Porath, Y.S., Lees-Haley, P., Green, P., 2011. A misleading review of response bias: comment on McGrath, Mitchell, Kim, and Hough (2010). *Psychological Bulletin* 137, 708–712.
- Rosen, G.M., 2003. Litigation and reported rates of posttraumatic stress disorder. *Personality and Individual Differences* 36, 1291–1294.
- Shapiro, D.N., Chandler, J., Mueller, P.A., 2013. Using Mechanical Turk to study clinical populations. *Clinical Psychological Science* 1, 213–220.
- Stevens, A., Friedel, E., Mehren, G., Merten, T., 2008. Malingering and uncooperativeness in psychiatric and psychological assessment: prevalence and effects in a German sample of claimants. *Psychiatry Research* 157, 191–200.
- Wiggins, C.W., Wygant, D.B., Hoelzle, J.B., Gervais, R.O., 2012. The more you say the less it means: overreporting and attenuated criterion validity in a forensic disability sample. *Psychological Injury and Law* 5, 162–173.
- Wildman, R.W., Wildman, R.W., 1999. The detection of malingering. *Psychological Reports* 84, 386–388.
- Williamson, D.J., Holsman, M., Chaytor, N., Miller, J.W., Drane, D.L., 2012. Abuse, not financial incentive, predicts non-credible cognitive performance in patients with psychogenic non-epileptic seizures. *Clinical Neuropsychologist* 26, 588–598.