

Decreasing Invalid Symptom Reporting: A Comment on Horner, Turner, VanKirk, and Denning (2017)

Isabella J.M. Niesten*, Alfons van Impelen, Harald Merckelbach

Clinical Psychological Science, Maastricht University, Maastricht, the Netherlands

*Corresponding author at: Clinical Psychological Science, Maastricht University, PO Box 616, 6200 MD, Maastricht, the Netherlands.
Tel.: +31-43-388-17-33.

E-mail address: elly.niesten@maastrichtuniversity.nl

Editorial Decision 2 January 2018; Accepted 10 January 2018

Recently, Horner, Turner, VanKirk, and Denning (2017) posited that invalid performance stems from a rational cost–benefit analysis and may be discouraged by warning individuals about both cost and benefits of the behavior (i.e., deterrence theory; Becker, 1968). They gave patients ($n = 121$; intervention group) a handout that listed consequences of valid and invalid responding, stressed the importance of effort, and required their signature. The handout also raised awareness of Performance Validity Tests (PVTs) in the test battery and warned that if invalid responding were to be detected, a report would be sent to the patient’s treatment team. Controls ($n = 122$) received a handout with general information about neuropsychological testing. Interestingly, a lower proportion of PVT-failure was observed *only* for those in the intervention group with a self-reported interest in disability benefits (i.e., 44% vs. 65% of controls). This led the authors to conclude that “even if it [the intervention] were effective only in a minority of cases, it would still provide substantial benefit with essentially no cost” (Horner et al., 2017; p. 9). We argue that several limitations call for a reassessment of this conclusion.

First, the researchers only report percentages of PVT-failure. This may invite readers to overvalue the data due to “collective statistical illiteracy” (Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2007): people—including (health) professionals—are not equipped to interpret numbers and more complex statistics in particular. Reporting raw frequencies instead—or in addition—improves transparency and interpretation. Table 1 depicts the absolute numbers of patients in Horner et al.’s study who failed or passed in each condition per self-reported disability-seeking status. The intervention handout led to a gain of six passes among patients who reported interest in benefits when compared to those who reported such interest but received the control handout. Among those not reporting interest in benefits there were five more failures in the intervention as opposed to the control condition. Effect sizes were not reported by the authors, but our calculation for the interaction between intervention and disability-seeking status ($N = 243$, $X^2 = 3.89$) yields a modest effect, Cohen’s $d = .255$.

Second, the findings rely heavily on self-report. Does the group that reported seeking benefits reflect the actual number of patients interested in benefits? The authors report that 182 patients were known to receive benefits, whereas 70 self-reported they were currently seeking them. While receipt of benefits does not necessarily imply *current* interest, these divergent numbers may suggest that some patients underreported the role benefits play in their symptom reporting. Relatedly, the impact of the intervention was determined by the Medical Symptom Validity Test (MSVT; Green, 2005) and clinicians’ judgment based on the results of this particular PVT—combined with sources not provided in the paper. Both approaches revealed a significant effect among patients reporting interest in benefits. However, if clinicians’ determination of invalid reporting is largely based on the MSVT—and why would it not be?—their judgment does not remediate the shortcoming that the battery included only one PVT; to the extent that clinical judgment is dependent on the MSVT reporting it as an additional route to scrutinize the data might lead the less cautious reader to interpret the finding as additional support for a robust effect of the intervention.

Third, studies on warning-based interventions and coaching suggest that warnings may help patients evade detection. Consequently, authors have advised against their use to ensure test-security (Suhr & Gunstad, 2000; Youngjohn et al., 1999). Horner et al. report data on the free recall trial of the MSVT to demonstrate that lower MSVT failure rates reflect honest responding: On this trial, those reporting interest in benefits who received the intervention performed similar to controls

Table 1. Absolute numbers of patients who passed and failed the MSVT

	Intervention (<i>n</i> = 121)		No intervention (<i>n</i> = 122)		Total <i>N</i>
	Fail	Pass	Fail	Pass	
Not seeking disability benefits	25 (29%)	60 (71%)	20 (23%)	68 (77%)	173
Seeking disability benefits	16 (44%)	20 (56%)	22 (65%)	12 (35%)	70

Notes: The authors did not report which subtests of the MSVT were used. While the original sample consisted of 251 participants, data about (self-reported) disability-seeking status was missing for eight patients, resulting in a sample of *N* = 243. Numbers in bold represent the percentages reported by the authors to illustrate the significant effect among patients reporting interest in disability benefits.

Table 2. Mean (*SD*) scores on the free recall trial of the MSVT for patients with a self-reported interest in disability benefits (*n* = 70) per condition and MSVT performance

	MSVT-free recall
Intervention – Fail	41.4 (18.0)
Intervention – Pass	61.5 (17.6)
Control – Pass	64.2 (15.2)

Notes: Scores on the MSVT-free recall for patients who received the control handout and failed the MSVT were not reported by the authors.

passing the MSVT, and they outperformed those who received the intervention but failed the MSVT (see Table 2). In our view, these data do not rule out the possibility that warnings helped some individuals evade detection. Given that patients may not know the placement of the MSVT's validity indices (i.e., assuming the test is not that transparent), their response would logically entail more careful performance across its subtests and resemble the symptom profiles of controls passing the MSVT. In fact, the pattern of scores observed for patients who received the intervention and reported interest in benefits seems in line with what patients who intentionally over-report their symptoms wish to achieve: a portrayal of themselves as someone with genuine pathology. Clearly, the inclusion of baseline performance would have allowed a more nuanced analysis of the intervention's effect (see for example, Suchy, Chelune, Franchow, & Thorgusen, 2012). One main issue, nevertheless, remains: the effect is difficult to study with known-groups designs because the ground truth about symptom exaggeration is not available.

In sum, although Horner et al.'s data may appear encouraging their conclusion should be tempered: the only effect found was modest in magnitude and ambiguities surrounding the data complicate their interpretation. It remains unclear how warning-based interventions affect patients' symptom reporting and there is reason to believe that they may have perilous effects due to sophisticated feigning but possibly also moral licensing: approaches that advocate honesty may—paradoxically—foster an overly positive self-evaluation that permits excusing one's own dishonest behavior (Niesten, Muller, Merckelbach, Dandachi-FitzGerald, & Jelicic, 2017). The decisional steps that patients go through when encountering situations in which invalid responding may be beneficial are poorly understood. Similarly, relatively little is known about patient-characteristics that contribute to symptom distortion. Data on these issues could aid the development of interventions that discourage invalid responding but minimize undesirable side-effects (i.e., more sophisticated feigning, moral licensing). Indeed, it may be timely for research to focus from warning-based interventions to studying the complexities of patients' decision-making.

Conflict of interest

We declare that we have no conflicts of interest.

References

- Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 78, 169–217.
- Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest*, 8, 53–96.
- Green, P. (2005). *Medical Symptom Validity Test manual*. Edmonton, Alberta, Canada: Green's Publishing.
- Horner, M. D., Turner, T. H., VanKirk, K. K., & Denning, J. H. (2017). An intervention to decrease the occurrence of invalid data on neuropsychological evaluation. *Archives of Clinical Neuropsychology*, 32, 228–237. <http://doi.org/10.1093/arclin/acw094>.
- Niesten, I. J. M., Muller, W., Merckelbach, H., Dandachi-FitzGerald, B., & Jelicic, M. (2017). Moral reminders do not reduce symptom over-reporting tendencies. *Psychological Injury & Law*, 10, 368–384.

- Suchy, Y., Chelune, G., Franchow, E. I., & Thorgusen, S. R. (2012). Confronting patients about insufficient effort: The impact on subsequent symptom validity and memory performance. *The Clinical Neuropsychologist*, *26*, 1296–1311.
- 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.
- 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. [http://doi.org/10.1016/S0887-6177\(98\)00049-3](http://doi.org/10.1016/S0887-6177(98)00049-3).