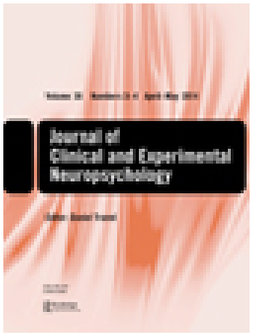

Routledge
Taylor & Francis Group

Journal of Clinical and Experimental Neuropsychology

ISSN: 1380-3395 (Print) 1744-411X (Online) Journal homepage: https://www.tandfonline.com/loi/ncen20

Evaluating the Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) over-reporting scales in a military neuropsychology clinic

Paul B. Ingram, Brittney L. Golden & Patrick J. Armistead-Jehle

To cite this article: Paul B. Ingram, Brittney L. Golden & Patrick J. Armistead-Jehle (2020): Evaluating the Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) over-reporting scales in a military neuropsychology clinic, Journal of Clinical and Experimental Neuropsychology, DOI: 10.1080/13803395.2019.1708271

To link to this article: https://doi.org/10.1080/13803395.2019.1708271

Published online: 03 Jan 2020.

JOURNAL OF CLINICAL AND EXPERIMENTAL NEUROPSYCHOLOGY
https://doi.org/10.1080/13803395.2019.1708271

Evaluating the Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) over-reporting scales in a military neuropsychology clinic

Paul B. Ingram [a,b], Brittney L. Golden[a] and Patrick J. Armistead-Jehle[c]

[a]Department of Psychological Sciences, Texas Tech University, Lubbock, TX, USA; [b]Dwight D. Eisenhower VAMC, Eastern Kansas Veteran Healthcare System, Leavenworth, KS, USA; [c]Munson Army Health Center, Concussion Clinic, Fort Leavenworth, KS, USA

ABSTRACT

Introduction: This study examines the utility of the Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) validity scales to detect invalid responding within a sample of active duty United States Army soldiers referred for neuropsychological evaluations.

Method: This study examines the relationship between performance validity testing and performance on the MMPI-2-RF over-reporting scales. Specifically, mean differences between those who passed (n = 152; 75.6%) or failed (n = 49; 24.4%) performance validity testing were compared. Receiver operator characteristic analyzes were also conducted to expand available information on the MMPI-2-RF over-reporting sensitivity and specificity in an Army sample.

Results: This study has two distinct findings. First, effect size differences between those passing and failing performance validity testing are classified as small to medium in magnitude (ranging from d = . 30/g = .32 on F-r to d = .66/g = .73 on RBS). Second, over-reporting scales have higher specificity and poorer sensitivity. Likewise, performance of the over-reporting scales suggests that those who exceeding recommended cut scores are likely to have failed extra-test performance validity measures.

Conclusion: These findings suggest that many who fail external performance measures may be undetected on the MMPI-2-RF over-reporting scales and that those exceeding recommended cut scores are likely to have failed extra-test performance validity testing. Implications for research on, and practice with, the MMPI-2-RF in military populations are discussed.

ARTICLE HISTORY
Received 28 May 2019
Accepted 13 December 2019

KEYWORDS
MMPI-2-RF; feigning; military; validity testing; over-reporting

The Minnesota Multiphasic Personality Inventory–2–Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008/2011) is amongst the most widely used instruments assessing psychopathology (Wright et al., 2017). In fact, in recent surveys of Veteran Affairs (VA) providers it was used more frequently than any other personality inventory during neuropsychological evaluations (Russo, 2018) and, more generally, it is among the most utilized symptom validity test (Young, Roper, & Arentsen, 2016). This wide-ranging use is likely because of the MMPI-2-RF's abbreviated length and noted improvements in psychometric strengths (e.g., Simms, Casillas, Clark, Watson, & Doebbeling, 2005). In addition to measuring under-reporting and acquiescent responding, the MMPI-2-RF has five validity scales assessing over-reporting. Three over-reporting scales were revised versions of scales developed on the MMPI-2 (e.g., Infrequent Responses [F-r], Infrequent Psychopathology Responses [Fp-r], and Symptom Validity [FBS-r]; Ben-Porath, 2012). However, Response Bias (RBS; Gervais, Ben-Porath, Wygant, & Green, 2007) and Infrequent Somatic Complaints (Fs; Wygant, Ben-Porath, & Arbisi, 2004) were introduced to strengthen an under-assessed area of over-reporting (infrequent somatic complaints in Fs) and to provide an alternative approach to identifying feigners (excessive failure of external performance validity tests [PVT] in RBS). Effective detection of over-reporting is important as it otherwise leads to weaker substantive scale relationships and reduces the predictive utility of validated assessments (Burchett & Ben-Porath, 2010; Wershba, Locke, & Lanyon, 2015; Wiggins, Wygant, Hoelzle, & Gervais, 2012).

Recently, two meta-analyzes have synthesized the literature on these five over-reporting scales. Both studies found substantially large effect sizes differentiating honest respondents and those identified as exaggerating or feigning their symptoms (Ingram & Ternes, 2016; Sharf, Rogers, Williams, & Henry, 2017). In addition to general support for the efficacy of the MMPI-2-RF over-reporting scales, Ingram and Ternes identified numerous moderators influencing the effectiveness of

CONTACT Paul B. Ingram pbingram@gmail.com Department of Psychological Sciences, Texas Tech University, Lubbock, TX, USA

The views, opinions, and/or findings contained in this article are those of the authors and should not be construed as an official Department of the Army, Department of Veteran Affairs, Department of Defense, or U.S. Government position, policy or decision unless so designated by other official documentation.

those scales, including some related to military service (e.g., veteran status, posttraumatic stress disorder [PTSD]).

Such moderation across the MMPI-2-RF's over-reporting scales is not surprising given that military service is a distinctive identity component for those who have served (Orazem et al., 2017) and that those with military service consistently demonstrate differences in their clinical presentation and evaluative needs. For instance, despite similarities in many treatments and principles of practice, traumatic brain injury and PTSD pose distinctive challenges for mental health providers conducting evaluations and providing treatment to those with military service (Armistead-Jehle, Soble, Cooper, & Belanger, 2017; Coll & Weiss, 2016; Dursa, Reinhard, Barth, & Schneiderman, 2014). Indeed, the prevalence of disorders amongst those with military service histories are distinct from the general population, resulting in unique health epidemics (e.g., Kilpatrick et al., 2013). Ingram and Ternes (2016) also noted a relative paucity of research assessing military and veteran samples relative to other populations. Given the patterns of moderation they found on the MMPI-2-RF validity scales for domains related to military service, a more limited literature is problematic.

Several studies have evaluated the MMPI-2-RF validity scales in samples with psychiatric and neuropsychological problems common among military personnel; however, most of these studies have utilized veteran participants rather than active duty or other current service members. Research has included simulation studies with veterans experimentally feigning common psychopathologies (e.g., Goodwin, Sellbom, & Arbisi, 2013), as well as descriptive findings from clinically derived samples. Findings from clinically derived samples are particularly important because simulation studies over-estimate what will be observed during clinical evaluation (e.g., Ingram & Ternes, 2016). For instance, Jurick et al. (2018) found that among Operation Enduring Freedom/Operation Iraqi Freedom (OEF/OIF) veterans, those with mild traumatic brain injury (mTBI) were likely to demonstrate symptom exaggeration on one or more MMPI-2-RF over-reporting scales, with the frequency of invalidity ranging from 50–87%, depending upon the cut score used. This high rate of profile invalidity is consistent with other research on mTBI in veterans (Nelson et al., 2011) and may reflect a broader pattern of response style for veterans (e.g., Ingram, Tarescavage, Ben-Porath, & Oehlert, 2019a, 2019b, 2019c), rather than the unique influence of mTBI on performance. This is likely, in part, due to the evaluative context in which testing is done within the VA relative to disability evaluations (Armistead-Jehle, 2010; Russo, 2017) although other factors, including evaluative presentation and the severity/frequency of clinical need, also play a role (Ingram et al., 2019a).

While veterans have similar service experiences as current military personnel, results from studies on veterans are not likely to offer the best comparisons or estimates of effect to those in active duty roles. There are a variety of potential reasons for this, including the recency of traumatic or blast injury events in military personnel or an extended opportunity for psychological care following the event for veterans. Time since injury is, after all, an important component of clinical presentation for some of the very concerns which make the experiences of military service and active-duty populations distinct (see Iverson, 2005). Another possible reason is that a disability evaluation may mean something distinct for someone who is currently serving compared to when a veteran undergoes a similar evaluation (e.g., return to duty decisions rather than monetary and medical benefit compensation). As such, the disability compensation process inherent to the Veteran Affairs system where many veterans receive their services is forensically enmeshed, producing a distinctive evaluation process (Russo, 2013). This setting may, as a result of this process, alter the effectiveness and clinical utility of the validity scales (Ingram & Ternes, 2016; Sharf et al., 2017).

Research utilizing only military personnel to assess the utility of the MMPI-2-RF over-reporting scales is more limited than research on veterans. Jones and Ingram (2011) used an optimal data analysis (ODA) paradigm to assess validity scale classification accuracy and found that the medium effect sizes of FBS-r and RBS were better than the small to medium effects of the F-family of scales. Conversely, when examining mean differences, Jones (2016a) found large effect sizes (d = .85–2.01) differentiating between identified non-feigners and groups comprising a variety of levels of feigning certainty (e.g., probable, probable to definite, and definite) in a sample of Army service personnel undergoing neuropsychological evaluation. Jones also found that MMPI-2-RF scales had high specificity, and that RBS had the largest overall effect size of the over-reporting scales (d = 1.58). RBS was the most effective scale in differentiating between military members who passed and failed PVTs, which is consistent with previous research on military service members undergoing neuropsychological evaluations (Jones, Ingram, & Ben-Porath, 2012).

In short, studies with military personnel have generally found that those exceeding recommended cut scores on the MMPI-2-RF are likely to fail concurrent tests of symptom or performance validity (e.g., Armistead-Jehle, Cooper et al., 2017; Bodapati et al., 2018). Moreover, estimations of classification accuracy

vary somewhat depending upon the analytical approach utilized (e.g., effect sizes are no more than moderate using an ODA approach but are large in magnitude when comparing mean differences in the same sample; see Jones & Ingram, 2011); however, RBS appears to consistently function as the most effective over-reporting scale. Despite this consensus, there remains a relative paucity of research on over-reporting in active duty military samples. Given this, and the broader need for improved neuropsychological testing in active duty personnel (Friedl et al., 2007), further investigation into the efficacy of the over-reporting scales of the MMPI-2-RF is warranted.

Present study

The clinical needs and evaluative context common to military personnel make the efficacy of the over-reporting scales within an Army service personnel population unique. There is also a shortage of research evaluating the efficacy of the MMPI-2-RF over-reporting scales within military samples. Accordingly, continued research is warranted to expand available information on the utility and efficacy of the MMPI-2-RF in military service members. Therefore, this study utilizes a sample of U.S. Army soldiers undergoing a neuropsychological evaluation at a neuropsychology clinic to examine the ability of the over-reporting scales to differentiate between those who passed or failed performance validity tests (PVT). Specifically, we evaluate MMPI-2-RF over-reporting scale score differences, compute the frequency of those exceeding interpretive recommendations, calculate the sensitivity (true positive rate) and specificity (true negative rate) of the validity scales, and provide collaborative neuropsychological testing data between those passing and failing PVT(s).

Method

Participants

This study utilized an initial sample of 216 (88.8% male; 73.1% White) active duty United States Army service members evaluated in an outpatient neuropsychological clinic. Evaluations occurred between June 2016 and August 2018 at a midwestern United States Army Health Center. Participants undergoing medical board (MEB) (n = 13) or Temporary Disability Evaluation (TED) (n = 2) were removed from this sample. Following this exclusion, the study's final sample was composed of 201 (88.6% male; 72.6% White) active duty Army service members. In general, participants were 34.5 years old (SD = 8.5), had an average of 14.9 years of education (SD = 2.5), and composed of approximately equal portions of enlisted and officers. Most of the referrals to this clinic are concussion-related; however, the clinic is responsible for other evaluations as well. In terms of neurological diagnosis, 49.3% were identified as having a history of mild traumatic brain injury (mTBI) and/or concussion. The average time since the last TBI was 72.9 months (SD = 68.7) and since the most significant TBI injury was 87.2 months (SD = 81.0). Approximately eighty percent of the sample was diagnosed with a psychiatric condition (82.6%), with the most common diagnoses being an anxiety disorder (27.9%), a depressive disorder (17.9%), attention-deficit hyperactivity disorder (10.4%), PTSD (8.5%), Adjustment Disorder (5.5%), PTSD and a depressive disorder (3.0%), or another disorder (9%). In general, those passing and failing PVT(s) were descriptively similar. Available demographic information for participants is provided in Table 1, including information separately for those who passed and failed PVT(s).

Table 1. Extended participant demographics.

	Full Sample		Failed PVT(s)		Passed PVT(s)	
	(n = 201)		(n = 49)		(n = 152)	
Variable	*n*	%	n	%	n	%
Gender (Male)	178	88.6%	45	91.8%	133	87.5%
Ethnicity						
White	146	72.6%	28	57.1%	118	77.6%
African American	25	12.4%	10	20.4%	15	9.9%
Hispanic	19	9.5%	10	20.4%	9	5.9%
Asian	7	3.5%	0	0.0%	7	4.6%
Native American/Pacific Islander	3	1.5%	0	0.0%	3	2.0%
Other	1	0.5%	1	2.0%	0	0.0%
Diagnosis						
mTBI/History of Concussion	99	49.3%	30	61.2%	69	45.4%
Psychiatric Issues	63	31.3%	10	20.4%	53	34.9%
Other	15	7.5%	4	8.2%	11	7.2%
None	24	11.9%	5	10.2%	19	12.5%
Marital Status						
Married	152	75.6%	35	71.4%	117	77.0%
Single	28	13.9%	10	20.4%	18	11.8%
Divorced	17	8.5%	2	4.1%	15	9.9%
Legally Separated	4	2.0%	2	4.1%	2	1.3%
Psychiatric Diagnosis						
PTSD	17	8.5%	7	14.3%	10	6.6%
DD	36	17.9%	5	10.2%	31	20.4%
PTSD and DD	6	3.0%	3	6.1%	3	2.0%
SUD	2	1.0%	1	2.0%	1	0.7%
Other Anxiety	56	27.9%	19	38.8%	37	24.3%
Adjustment	11	5.5%	1	2.0%	10	6.6%
ADHD	21	10.4%	3	6.1%	18	11.8%
Other	16	8.0%	4	8.2%	12	7.9%
None	35	17.4%	6	12.2%	29	19.1%
Rank						
Enlisted	104	51.7%	30	61.2%	74	48.7%
Officer	97	48.3%	19	38.8%	78	51.3%

DD = depressive disorder, PTSD = posttraumatic stress disorder, SUD = substance use disorder, ADHD = attention-deficit hyperactivity disorder.

Measures

MMPI-2-restructured form

The MMPI-2-RF (Ben-Porath & Tellegen, 2008/2011) is a 338 true-false item personality measure comprised of 51 scales. The 42 substantive scales measure various clinical constructs and have demonstrated validity in a variety of military and veteran samples (e.g., Arbisi, Polusny, Erbes, Thuras, & Reddy, 2011; Goodwin et al., 2013; Gottfried, Bodell, Carbonell, & Joiner, 2014; Ingram et al., 2019a, 2019b). The nine validity scales are used to determine if a respondent is engaging in some form of non-credible responding (non-content based invalid responding, over-reporting, or under-reporting). The MMPI-2-RF technical manual describes the following T-scores as indicating profile invalidity: VRIN-r ≥ 80, TRIN-r ≥ 80, F-r ≥ 120, Fp-r ≥ 100, Fs ≥ 100, RBS ≥ 100, FBS-r ≥ 100, L-r ≥ 80, and K-r ≥ 70. The validity scales of the MMPI-2-RF have consistently demonstrated large effect sizes between groups passing or failing performance and symptom validity tests (Ingram & Ternes, 2016; Sharf et al., 2017). Within military disability samples specifically, the RBS and FBS-r scales outperform the F-family of scales in classification accuracy (Jones & Ingram, 2011; Jones et al., 2012). In this sample, the MMPI-2-RF was electronically administered resulting in no missing data or elevations on the Cannot Say (CNS) scale.

Performance validity

To establish performance validity against which the efficacy of the MMPI-2-RF over-reported scales could be compared, participants were given at least one of three common measures of performance validity. Performance validity relies on quantifiable test performance, typically on tests of memory and cognition, while symptom validity tests evaluate symptom frequency, intensity, duration, and presentation to determine the probability that those symptoms would occur. In our sample, all were administered at least two PVTs and roughly one quarter (26.4%) were administered all three PVTs. Of those individuals with failed performance validity testing, most failed a single PVT (n = 39; 79.6%) while a few failed two (n = 8; 16.3%) or three PVTs (n = 2; 5.1%). As such, our sample is best classified as possible malingering for those failing PVT testing (Slick, Sherman, & Iverson, 1999).

While the MMPI-2-RF scales evaluate symptom validity, measures of performance validity are often used to evaluate the validity of the MMPI-2-RF over-reporting scales. This is particularly pronounced in clinics and populations with cognitive complaint concerns (e.g., Gervais et al., 2017; Rogers et al., 2011; Wygant et al., 2010). PVTs were also utilized during the development and validation of the RBS scale by guiding item selection processes (Gervais et al., 2007). Each PVT indicator within this study has a lengthy history supporting their use for the detection of low performance effort. A brief summary of the psychometric properties of each PVT utilized within this study is outlined below, along with the portion of individuals identified as having failed that performance indicator.

Test of memory malingering. The Test of Memory Malingering (TOMM; Tombaugh, 1996) is a widely used memory assessment which evaluates inadequate or feigned responding to a typical memory task using a visual recognition task. Across Sollman and Berry's (2011) meta-analytic review, the TOMM demonstrated a very large (d = 1.59) effect size as well as high mean sensitivity (69%) and specificity (90%) across studies on feigned memory complaints. In this study, all participants were administered the TOMM and 11 (5.5%) were classified as having failed using the criteria of a score of less than or equal to 44 on Trial 2. A score of equal to or less than 44 is the recommended guideline for the TOMM because only 2% of those who are non-demented meet or exceed this score (Tombaugh, 1996). In this sample, 11 individuals failed the TOMM based on these criteria.

Effort index of the RBANS. The effort index (EI; Silverberg, Wertheimer, & Fichtenberg, 2007) of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, 1998) is a composite index based on two instrument subtests using a scaling system. Research on the EI index has demonstrated mixed results regarding its effectiveness as an internal validity indicator. In studies drawn from military samples, the EI has demonstrated consistent evidence of good positive and negative predictive values and, in cases of definite malingering, an EI of 1 demonstrated a .89 sensitivity and .97 specificity (Jones, 2016b). Contrasting that universal effectiveness, researchers have noted that the EI may not be ideal because of its lower sensitivity (Armistead-Jehle & Hansen, 2011). However, the rates of specificity (e.g., the rate at which individuals are correctly identified as not having extra-test evidence of performance invalidity) are consistently high (> .80). In this study, 53 participants were administered the RBANS and 9 (21.9%) had an EI score of 1 or greater and were classified as having failed the RBANS EI.

Medical symptom validity test. The Medical Symptom Validity Test (MSVT; Green, 2004) involves a 10-item verbal memory task which assesses performance

validity. After 10 minutes, participants are asked to recall the verbal memory prompts. Research has consistently supported the MSVT's utility to identify simulated memory impairment and inadequate performance effort, including amongst those with mild traumatic brain injury (e.g., Green, Flaro, Brockhaus, & Montijo, 2012; Howe, Anderson, Kaufman, Sachs, & Loring, 2007). Likewise, meta-analytic results have shown the MSVT to have large between group effects (d = .94) as well as high sensitivity (70.0%) and specificity (91.3%) (see Sollman & Berry, 2011). The MSVT assesses effort when performance is less than or equal to 85% on Immediate Recognition, Delayed Recognition, or Consistency subtests. The Manual reports sensitivity of approximately 97% in simulator studies (Green, 2004) and in cases of possible dementia the sensitivity is between 90 and 100% (Howe & Loring, 2009; Singhal, Green, Ashaye, Shankar, & Gill, 2009). In this study, all participants were administered the MSVT and 41 (20.4%) failed the MSVT effort tests.

Procedures and planned analysis

Participants were referred for neuropsychological evaluation largely by primary care or behavioral health providers. Data was collected by a psychometrist under the supervision of a board-certified neuropsychologist (the third author of this paper) who provided clinical interpretations and diagnostic formulations based on testing results. This project received IRB approval from the Madigan Army Medical Center to conduct analyzes on an anonymized version of the clinical database in which testing results from these evaluations were stored. Respondents were grouped based on their PVT performance (pass all/failed at least one). Individuals were identified for the failed group when scores on any administered PVT indicated a pattern of probable or possible invalid responding. Differences between groups were calculated for each over-reporting scale using independent t-tests. Receiver operator characteristic (ROC) analyzes were planned for each of the over-reporting scales to determine sensitivity and specificity for the MMPI-2-RF validity scales at various cut scores. Lastly, consistent with how the MMPI-2-RF is used clinically, positive and negative predictive value (PPV and NPV, respectively) were calculated when utilizing the MMPI-2-RF over-reporting scales conjunctively to determine profile classification accuracy. Effect size estimates were calculated using observed means using Cohen's *d*, as well as Hedge's *g* (a sample size adjusted estimator of effect). By providing both effect size estimates, this study can be contrasted with existing literature (which predominantly use Cohen's *d*) and provides more accurate estimates. An a priori determination of clinically meaningful differences between groups was made using a medium effect size – the equivalent of 5 *T*-score points (Rosnow, Rosenthal, & Rubin, 2000).

Results

Following exclusion for random (VRIN-r > 80) and acquiescent (TRIN-r > 80) responding, participants were grouped according to if they passed all (n = 152; 75.6%) or failed any administered (n = 49; 24.4%) PVTs. Descriptive characteristics were calculated for the sample using available collaborating cognitive and neuropsychological testing data, whichis presented in Table 2. In general, cognitive and neuropsychological testing data suggest that those who failed PVTs also demonstrated moderate to large declines across other psychological tests administered during the same evaluation.

Independent t-tests indicated significant statistical differences between those passing or failing PVTs for each of the MMPI-2-RF over-reporting scales (Table 3). Based on means and bootstrapped estimates of confidence intervals using 1000 simulated samples, none of the five over-reporting scales yielded a large effect ($d \geq |.8|$), while all had at least a small effect ($d \leq |.2|$). Differences between the PVT pass/fail groups are most classified as either small or medium effects and the RBS scale demonstrated the largest effect (d = .66; g = .73). F-r (d = .30), Fs (d = .35), and FBS-r (d = .36) fell below the requisite *a priori* threshold for statistical significance ($d < |.5|$; Cohen, 1988). F-r, Fs, and FBS-r also failed to reach a medium effect size when utilizing Hedges *g*. However, differences for these scales meet clinical significance (5 *T*-score points; Rosnow et al., 2000). In general, the magnitude of these estimates of effect suggested that classification effectiveness is best for Fp-r and RBS and that these are medium effects.

ROC analyzes were then conducted to determine the sensitivity and specificity of each over-reporting scale based on participants PVT pass/fail status. Select scores of each MMPI-2-RF over-reporting scale and their associated classification effectiveness are presented in Table 4. Area under curve (AUC) was calculated for each scale: F-r = .575 (standard error [SE] = .046, 95% confidence interval [CI] = .485– .666); Fp-r = .599 (SE = .049, 95% CI = .502– .695); Fs = .593 (SE = .046; 95% CI = .502– .684); FBS-r = .595 (SE = .045, 95% CI = .506– .584); and RBS = .616 (SE = .046; 95% CI = .525– .707). In general, the over-reporting scales performed similarity and AUC, which approximates the degree to which sensitivity and specificity pairings can distinguish between groups,were alike across most over-reporting scales. In general, the

Table 2. Neuropsychological testing by group.

		Full Sample (n = 201)		Failed PVT(s) (n = 49)			Passed PVT(s) (n = 152)					
Variable	*n*	M	SD	n	M	SD	n	M	SD	*t*	*d*	*g*
WAIS-IV												
FSIQ	144	104.1	12.3	25	92.8	9.48	119	106.5	11.5	5.539***	1.29	1.22
VCI	144	106.4	13.2	25	96.9	11.3	119	108.4	12.8	4.144***	0.95	0.91
PRI	144	105.1	14.7	25	95.1	11.9	119	107.2	14.4	3.929***	0.92	0.86
WMI	146	99.0	13.6	25	91.7	11.2	121	100.5	13.6	3.034**	0.71	0.67
PSI	146	99.9	14.1	25	92.2	12.6	121	101.5	13.9	3.103**	0.70	0.68
COWAT	195	45.5	10.2	49	41.1	10.2	146	47.0	9.9	3.576***	0.59	0.59
RBANS												
Total	53	92.5	13.9	24	85.1	12.5	29	98.6	12.0	4.001***	1.10	1.10
Immediate Memory	53	93.6	13.8	24	88.9	13.3	29	97.4	13.2	2.34*	0.64	0.64
Visuospatial/Construction	53	100.0	14.7	24	96.6	15.8	29	102.9	13.3	1.569^{ns}	0.43	0.43
Language	53	97.4	11.8	24	94.9	11.8	29	99.4	11.6	1.407^{ns}	0.39	0.39
Attention	53	91.6	16.8	24	84.7	15.7	29	97.4	15.7	2.949**	0.81	0.81
Delayed Memory	53	90.3	17.7	24	78.9	18.4	29	99.7	9.9	5.246***	1.41	1.45
Trail Making Test Form A	195	46.2	12.3	49	40.2	11.8	146	48.3	11.8	4.138***	0.68	0.68
Trail Making Test Form B	195	47.3	11.4	49	41.4	10.7	146	49.2	10.93	4.361***	0.73	0.72

WAIS-IV = Weschler Adult Intelligence Scale, Fourth Edition. COWAT = Controlled Oral Word Association Test, and RBANS = Repeatable Battery for the Assessment of Neuropsychological Status. FSIQ = Full Scale IQ, PSI = Processing Speed Index, WMI = Working Memory Index, VCI = Verbal Comprehension Index, and PRI = Perceptual Reasoning Index. COWAT scores are presented as standardized T-scores. T-tests presented are between those in the failed and passed PVT groups. *** p < .001, ** p < .05, *p < .01, ns = non-significant.

Table 3. Differences in the MMPI-2-RF scales according to extra-test grouping criteria.

	Full Sample (n = 201)		Failed PVT(s) (n = 49)		Passed PVT(s) (n = 152)				
Scale	M (95% CI)	SD (95% CI)	M (95% CI)	SD (95% CI)	M (95% CI)	SD (95% CI)	t	d	g
F-r	66.2 (63.5–69.1)	20.1 (17.7–22.4)	71.0 (64.7–77.7)	23.1 (17.3–26.9)	64.7 (61.8–67.7)	18.8 (16.2–21.2)	1.913^{ns}	.30	.32
Fp-r	57.3 (55.2–59.8)	16.4 (13.9–18.7)	64.3 (58.2–70.8)	23.2 (17.3–27.2)	55.1 (53.0–57.2)	12.9 (11.3–14.4)	3.525***	.49	.57
Fs	61.6 (58.9–64.4)	18.4 (15.6–20.7)	66.8 (60.7–73.0)	21.9 (15.7–26.0)	59.3 (57.7–63.0)	16.8 (14.2–19.3)	2.309*	.35	.38
FBS-r	64.3 (62.3–66.3)	13.5 (11.9–15.1)	68.0 (63.9–72.1)	14.2 (10.5–17.8)	63.1 (61.0–65.3)	13.1 (11.5–14.5)	2.119*	.36	.36
RBS	69.8 (67.7–71.8)	15.5 (13.6–17.4)	75.7 (70.6–81.1)	18.1 (13.6–21.5)	67.8 (65.4–70.2)	14.2 (12.4–15.8)	3.155**	.66	.73

Confidence intervals were estimates using 1000 bootstrapped samples. RCS = the percentage of the sample that has a score at or above the T-score value which invalidates the MMPI-2-RF protocol (Ben-Porath & Tellegen, 2008/2011). T-tests presented are between those in the failed and passed PVT groups. *** p < .001, ** p < .05, *p < .01, *ns = non-significant.*

Table 4. Classification estimates for the MMPI-2-RF over-reporting validity scales.

	F-r		Fp-r		Fs		FBS-r		RBS	
Cutoff Score	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
T = 120	**0.00**	**1.00**	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00
T ≥ 115	0.12	0.98	0.00	1.00	0.06	0.99	0.00	1.00	0.06	1.00
T ≥ 110	0.12	0.97	0.06	1.00	0.10	0.98	0.02	1.00	0.08	1.00
T ≥ 105	0.12	0.97	0.08	1.00	0.10	0.98	0.02	1.00	0.08	0.99
T ≥ 100	0.14	0.94	**0.08**	**1.00**	**0.10**	**0.97**	**0.02**	**1.00**	**0.10**	**0.99**
T ≥ 95	0.16	0.93	0.14	1.00	0.12	0.95	0.02	0.99	0.10	0.97
T ≥ 90	0.16	0.88	0.14	1.00	0.12	0.95	0.06	0.96	0.16	0.95
T ≥ 85	0.20	0.85	0.14	0.99	0.16	0.93	0.08	0.95	0.20	0.88
T ≥ 80	0.25	0.82	0.22	0.95	0.16	0.93	0.18	0.91	0.29	0.87
T ≥ 75	0.31	0.77	0.22	0.95	0.18	0.87	0.25	0.83	0.35	0.78
T ≥ 70	0.37	0.72	0.29	0.88	0.35	0.78	0.35	0.73	0.47	0.66

Bolded scores reflect cut-values recommended by the MMPI-2-RF interpretive Manual (Ben-Porath & Tellegen, 2008).

over-reporting scales of the MMPI-2-RF have high specificity (true positive rate) and low sensitivity (true negative rate). Said another way, at their respective recommended cut-scores most individuals meeting or exceeding cut values on the MMPI-2-RF over-reporting scales fail PVTs; however, those who failed PVTs frequently did not elevate the MMPI-2-RF over-reporting scores to a level which indicates profile invalidity. Thus, clinicians may have more confidence in positively identifying over-reported symptoms and less confidence that negatively screened profiles indicate valid responding.

Discussion

This study evaluated differences on MMPI-2-RF over-reporting scale scores by comparing groups of active-duty U.S. Army personnel who passed or failed performance validity testing as part of a comprehensive neuropsychological evaluation. Continued research on the efficacy of the MMPI-2-RF over-reporting scales was needed for two reasons. First, there is a shortage of research on military and veteran groups with the MMPI-2-RF validity scales (Ingram & Ternes, 2016). Second, issues critical to those in military service commonly influence response style and clinical presentation (e.g., Jurick et al., 2018). Results from this study suggest three trends that are particularly relevant for the use of, and research on, the MMPI-2-RF: (a) effect size differences between those passing all or failing PVTs are somewhat different across the scales with the most effective scales performing generally consistent with medium effects, (b) utilizing traditional cut scores, the over-reporting scales have high specificity and low sensitivity, and (c) validity scale mean scores are generally below recommended cut values for all participants, including those who failed PVTs, which is likely a contributing factor for the low sensitivity.

The over-reporting scales of the MMPI-2-RF demonstrated small to medium effect sizes (d mean = .45; g mean = .47) and poor sensitivities in this sample drawn from an Army neuropsychology clinic. The frequently low validity scale scores for those failing PVTs served as a contributing factor to the medium effect size and high false negative rates. In contrast, all the over-reporting scales provided high specificity estimates at recommended cut-scores. Accordingly, using recommended cut-scores result in a low rate of false positives and a modestly high rate of classifying those with failed PVT testing. These results suggest that many military respondents who fail external performance validity tests are unlikely to elevate the MMPI-2-RF over-reporting scales but those that do are likely to invalidate measures of performance effort during that same evaluation. Indeed, findings that respondents' MMPI-2-RF validity scale elevations (both of those failing and passing extra-test performance measures) frequently do not exceed recommended cut-scores is common (Ingram & Ternes, 2016; Sharf et al., 2017). Accordingly, the MMPI-2-RF over-reporting scales are likely not sufficient as the sole method for screening performance misrepresentation or exaggeration within military samples. However, when an elevation occurs on one of the MMPI-2-RF over-reporting scales which would invalidate the MMPI-2-RF protocol, clinicians are likely to find evidence of failed performance validity testing on administered performance testing. As such, clinicians may have confidence that when MMPI-2-RF validity scales are deemed to have invalidated the test protocol, the substantive scales of the MMPI-2-RF should not be interpreted.

In contrast to our findings, some studies have found substantially larger differences on the MMPI-2-RF over-reporting scales when those with possible symptom exaggeration and those identified as honest respondents are compared. For example, Jurick et al. (2018) found that the Fs and RBS scales were frequently at levels indicating profile invalidity in those with possible symptom exaggeration. This variability from the large effect sizes observed in some studies may reflect the different methodologies utilized to identify and group individuals. Slick et al. (1999) recommend feigning classifications which express different magnitudes of certainty about group membership (e.g., possible, probable, and definite malingering). Indeed, the largest effect sizes for the MMPI-2-RF over-reporting scales are observed in cases where effort is defined by the strongest evidence of malingering (i.e., *definite*) is present (Wygant et al., 2011). Most of the sample identified as having failed performance testing had only one failed PVT measure and this may explain the magnitudes of effect observed in this study. While research using methodology similar to this study has sometimes found large effect sizes (Jones, 2016a; Jones et al., 2012), the moderate effects of this study are most consistent with research where respondents undergoing neuropsychological evaluation were deemed to have evidence of possible malingering (e.g., Tarescavage, Wygant, Gervais, & Ben-Porath, 2013).

These differences in detection effectiveness may be due to the use of different extra-test criteria across studies. Because of the inherent differences in the sensitivities of those instruments, classification accuracy in any given study relies on the assumption that classification is accurate. As such, positive and negative predictive values in studies will, therefore, vary along with the true classification accuracy of those instruments. For instance, we used the publication manual's recommended cut score of 44 for the TOMM Trial 2 to determine grouping while others have found 49 to be more preferred (Schroeder et al., 2013). Clinicians should, therefore, be mindful of these methodological considerations to ensure that testing procedures in comparable work are representative of their clinical assessment battery. Likewise, extra-test groups definition has differed across studies and this may further contribute

to variation in validity scale estimates of effect. For a greater discussion of trends in PVT/SVT classification within feigning research and their implications for interpretation, see Larrabee (2012).

Additionally, there is some variability in the effectiveness of the over-reporting scales within this study when considering effect size estimates. While F-r consistently had the smallest effect estimate, RBS was the most effective MMPI-2-RF over-reporting scale at differentiating between those who failed PVTs and those who did not. Likewise, the initial validation of the RBS scale highlighted its utility compared to the F-family of scales (Gervais et al., 2007) and research within military samples has regularly found that RBS outperforms other somatic/cognitive over-reporting scales when utilizing PVTs as criterion measures within the study design (e.g., Jones & Ingram, 2011; Jones, 2016a). Accordingly, these findings are consistent with research which notes that RBS is the preferred over-reporting scale for use within military samples where cognitive functioning was assessed or a primary symptom of the presentation (e.g., Jones et al., 2012). Of note, the MMPI-2-RF validity scales are largely symptom validity measures (e.g., utilizing infrequent symptom responding such as is common within the F-family of scales; Ben-Porath, 2012) rather than performance-based indicators. This may also explain why RBS has the largest effect size, given PVT failure as a criterion for item selection (see Gervais et al., 2007).

Low sensitivity estimates suggest low true positive rates, indicating that many of those with evidence of feigning are undetected. Fortunately, specificity estimates are high at recommended cut scores, suggesting that the over-reporting scales also have low false positive rates and a low rate of false positives is particularly desirable within validity testing (Larrabee, 2012). Jones (2016a) found that no over-reporting scale had less than 95% classification accuracy when using recommended cut scores. While classification accuracy was substantially lower in this study, utilization of the MMPI-2-RF scales conjunctively to screen for invalid responding suggests that the over-reporting scales perform reasonably well in classifying those with failed extra-test performance testing. The observed effectiveness in this study is most consistent with other research in which respondents were classified as demonstrating evidence of "possible malingering". Given the performance of the MMPI-2-RF over-reporting scales within this study, it would be wise for clinicians to utilize the recommended cut-scores of the MMPI-2-RF validity scales to ensure the highest confidence in identifying individuals who are not feigning symptom presentation. Likewise, they should also incorporate several additional psychometrically strong measures of performance and symptom validity in order to aid their assessment efforts.

It is also important to remember that the results of this study reflect testing conducted within a neuropsychological clinic where cognitive issues reflect the most frequent referral concern. As such, generalization of these findings will be most appropriate for settings in which cognitive performance feigning is a primary presentation concern, which may reflect why RBS was the most effective scale. The degree to which this may generalize will, of course, vary by the population being studied. Thus, inferences made from the results of this study rely on the stability of feigning prevalence and extra-test evidence supporting feigning detection from this sample to other similar samples.

Importantly, while the low sensitivity is not ideal for the best classification accuracy, the high specificity and suggests that when participants exceed recommended MMPI-2-RF cut scores they were not likely to have failed extra-test performance validity testing. As such, clinicians may have greater confidence that those exceeding recommended cut-scores are demonstrating evidence consistent with those exaggerating or feigning psychiatric, cognitive, or somatic concerns (Berry & Nelson, 2010), dependent upon the MMPI-2-RF validity scale(s) elevated (see Sharf et al., 2017). However, labeling performance as indicative of feigning should not, of course, be based on a test's scores alone as multimodal assessment is preferred (Ali, Jabeen, & Alam, 2015; Larrabee, 2012). Clinicians should also be mindful that the appearance of non-credible responding may result from not only intentional over-reporting, but also from cultural factors or genuine psychiatric and neuropsychological problems. As research on the MMPI-2-RF over-reporting scales continues, it will be useful to understand more about the role of these contextual factors so that they may be appropriately considered. Additionally, participants were not all administered the same PVTs. As a result, it was not possible to create a meaningful grouping variable that conveyed differential levels of effort (e.g., *fail 1 PVT, fail 2 PVT, etc.*; Gervais, Wygant, Sellbom, & Ben-Porath, 2011) because of differential sensitivity and specificities of the administered validity tests. It would also be useful to provide validation for the MMPI-2-RF scales utilizing SVTs rather than PVTs. While SVTs are included in some studies (e.g., Chmielewski, Zhu, Burchett, Bury, & Bagby, 2017), much of the literature on the MMPI-2-RF over-reporting scales draw from clinical samples where performance validity is utilized as the primary grouping criterion. Doing so in future studies may provide an

improvedunderstanding of validity scale performance and its relationship to failed extra-test performance measures within military samples.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Paul Ingram receives research funding from Pearson Clinical Assessments, the distributor of the MMPI-2-Restructured Form, and the University of Minnesota Press, publisher of the MMPI-2-Restructured Form.

ORCID

Paul B. Ingram http://orcid.org/0000-0002-5409-4896

References

Ali, S., Jabeen, S., & Alam, F. (2015). Multimodal approach to identifying malingered posttraumatic stress disorder: A review. *Innovations in Clinical Neuroscience*, *12*(1–2), 12–20.

Arbisi, P. A., Polusny, M. A., Erbes, C. R., Thuras, P., & Reddy, M. K. (2011). The Minnesota multiphasic personality inventory-2 restructured form in national guard soldiers screening positive for posttraumatic stress disorder and mild traumatic brain injury. *Psychological Assessment*, *23*, 203–214.

Armistead-Jehle, P. (2010). Symptom validity test performance in U.S. veterans referred for evaluation of mild TBI. *Journal of Applied Neuropsychology*, *17*, 52–59.

Armistead-Jehle, P., Cooper, D. B., Grills, C. E., Cole, W. R., Lippa, S. M., Stegman, R. L., & Lange, R. T. (2017). Clinical utility of the mBIAS and NSI validity-10 to detect symptom over-reporting following mild TBI: A multicenter investigation with military service members. *Journal of Clinical and Experimental Neuropsychology*, *40*, 213–223.

Armistead-Jehle, P., & Hansen, C. L. (2011). Comparison of the repeatable battery for the assessment of neuropsychological status effort index and stand-alone symptom validity tests in a military sample. *Archives of Clinical Neuropsychology*, *26*, 592–601.

Armistead-Jehle, P., Soble, J. R., Cooper, D. B., & Belanger, H. G. (2017). Unique aspects of traumatic brain injury in military and veteran populations. *Physical Medicine and Rehabilitation Clinics*, *28*, 323–337.

Ben-Porath, Y. S. (2012). *Interpreting the MMPI-2-RF*. Minneapolis, MN: University of Minnesota Press.

Ben-Porath, Y. S., & Tellegen, A. (2008/2011). *Minnesota multiphasic personality inventory-2 restructured form: MMPI-2-RF*. Minneapolis, MN: University of Minnesota Press.

Berry, D. T., & Nelson, N. W. (2010). DSM-5 and malingering: A modest proposal. *Psychological Injury and Law*, *3*, 295–303.

Bodapati, A. S., Combs, H. L., Pastorek, N. J., Miller, B., Troyanskaya, M., Romesser, J., & Linck, J. (2018). Detection of symptom over-reporting on the neurobehavioral symptom inventory in OEF/OIF/OND veterans with a history of mild TBI. *The Clinical Neuropsychologist*. Advanced Online Publication. doi:10.1080/13854046.2018.1482003

Burchett, D. L., & Ben-Porath, Y. S. (2010). The impact of over-reporting on the MMPI-2-RF substantive scale score validity. *Assessment*, *17*, 497–516.

Chmielewski, M., Zhu, J., Burchett, D., Bury, A. S., & Bagby, R. M. (2017). The comparative capacity of the Minnesota Multiphasic Personality Inventory–2 (MMPI–2) and MMPI–2 Restructured Form (MMPI-2-RF) validity scales to detect suspected malingering in a disability claimant sample. *Psychological Assessment*, *29*, 199–208.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. (RevisedEd.). New York, NY: Academic Press.

Coll, J. E., & Weiss, E. L. (2016). *Caring for the military: A guide for helping professionals*. New York, NY: Routledge.

Dursa, E. K., Reinhard, M. J., Barth, S. K., & Schneiderman, A. L. (2014). Prevalence of a positive screen for PTSD among OEF/OIF and OEF/OIF-era veterans in a large population-based cohort. *Journal of Traumatic Stress*, *27*, 542–549.

Friedl, K. E., Grate, S. J., Proctor, S. P., Ness, J. W., Lukey, B. J., & Kane, R. L. (2007). Army research needs automated neuropsychological tests: Monitoring soldier health and performance status. *Archives of Clinical Neuropsychology*, *22*, S7–S14.

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

Gervais, R. O., Tarescavage, A. M., Greiffenstein, M. F., Wygant, D. B., Deslauriers, C., & Arends, P. (2017). Inconsistent responding on the MMPI-2-RF and uncooperative attitude: Evidence from cognitive performance validity measures. *Psychological Assessment*, *30*(3), 410–415. doi:10.1037/pas0000506

Gervais, R.O, Wygant, D.B, Sellbom, M, & Ben-Porath, Y.S. (2011). Associations between symptom validity test failure and scores on the mmpi–2–rf validity and substantive scales. *Journal of Personality Assessment*, *93*(5), 508–517. doi:10.1080/00223891.2011.594132

Goodwin, B. E., Sellbom, M., & Arbisi, P. A. (2013). Posttraumatic stress disorder in veterans: The utility of the MMPI-2-RF validity scales in detecting overreported symptoms. *Psychological Assessment*, *25*, 671–678.

Gottfried, E., Bodell, L., Carbonell, J., & Joiner, T. (2014). The clinical utility of the MMPI–2–RF suicidal/death ideation scale. *Psychological Assessment*, *26*, 1205–1211.

Green, P. (2004). *Green's Medical Symptom Validity Test (MVST) for Microsoft Windows: User's manual*. Edmonton, Canada: Green's Publishing.

Green, P., Flaro, L., Brockhaus, R., & Montijo, J. (2012). Performance on the WMT, MSVT, and NV-MSVT in children with developmental disabilities and adults with mild traumatic brain injury. In C. R. Reynolds & A. M. Horton Jr. (Eds.), *Detection of malingering during head injury litigation* (2nd ed., 201–219). New York, NY: Plenum Press.

Howe, L. L., & Loring, D. W. (2009). Classification accuracy and predictive ability of the medical symptom validity test's dementia profile and general memory impairment profile. *The Clinical Neuropsychologist*, *23*, 329–342.

Howe, L. L. S., Anderson, A. M., Kaufman, D. A. S., Sachs, B. C., & Loring, D. W. (2007). Characterization of the medical symptom validity test in evaluation of clinically

referred memory disorders clinic patients. *Archives of Clinical Neuropsychology*, *22*(6), 753–761.

Ingram, P. B., Tarescavage, A. M., Ben-Porath, Y. S., & Oehlert, M. E. (2019c). Comparing MMPI-2-Restructured Form scores by service era for veterans assessed within the veteran affairs healthcare system. *Journal of Clinical Psychology in Medical Settings*. Advanced Online Publication. doi:10.1007/s10880-019-09650-2

Ingram, P. B., Tarescavage, A. S., Ben-Porath, Y. S., & Oehlert, M. E. (2019a). Descriptive profiles of the MMPI-2-Restructured Form (MMPI-2-RF) across a national sample of four veteran affairs treatment settings. *Journal of Psychopathology and Behavioral Assessment*, *41*, 329–340. Advanced Online Publication.

Ingram, P. B., Tarescavage, A. S., Ben-Porath, Y. S., & Oehlert, M. E. (2019b). Patterns of MMPI-2-Restructured Form (MMPI-2-RF) validity scale scores observed across veteran affairs settings [Measurement-based care and psychological assessment in Mental Health Services]. *Psychological Services*. Advanced Online Publication. doi:10.1037/ser0000339

Ingram, P. B., & Ternes, M. (2016). The detection of content-based invalid responding: A meta-analysis of the MMPI-2-Restructured Form's (MMPI-2-RF) over-reporting scales. *The Clinical Neuropsychologist*, *30*, 473–496.

Iverson, G. L. (2005). Outcome from mild traumatic brain injury. *Current Opinion in Psychiatry*, *18*, 301–317.

Jones, A. (2016a). Cutoff scores for MMPI-2 and MMPI-2-RF cognitive-somatic validity scales for psychometrically defined malingering groups in a military sample. *Archives of Clinical Neuropsychology*, *31*, 786–801.

Jones, A. (2016b). Repeatable battery for the assessment of neuropsychological status: Effort index cutoff scores for psychometrically defined malingering groups in a military sample. *Archives of Clinical Neuropsychology*, *31*, 273–283.

Jones, A., & Ingram, M. V. (2011). A comparison of selected MMPI-2 and MMPI-2-RF validity scales in assessing effort on cognitive tests in a military sample. *The Clinical Neuropsychologist*, *25*, 1207–1227.

Jones, A., Ingram, M. V., & Ben-Porath, Y. S. (2012). Scores on the MMPI-2-RF scales as a function of increasing levels of failure on cognitive symptom validity tests in a military sample. *The Clinical Neuropsychologist*, *26*, 790–815.

Jurick, S. M., Crocker, L. D., Keller, A. V., Hoffman, S. N., Bomyea, J., Jacobson, M. W., & Jak, A. J. (2018). The Minnesota Multiphasic Personality Inventory-2-RF in treatment-seeking veterans with history of mild traumatic brain injury. *Archives of Clinical Neuropsychology. Advanced Online Publication.*, *34*, 366–380.

Kilpatrick, D. G., Resnick, H. S., Milanak, M. E., Miller, M. W., Keyes, K. M., & Friedman, M. J. (2013). National estimates of exposure to traumatic events and PTSD prevalence using DSM-IV and DSM-5 criteria. *Journal of Traumatic Stress*, *26*, 537–547.

Larrabee, G. J. (2012). Performance validity and symptom validity in neuropsychological assessment. *Journal of the International Neuropsychological Society*, *18*, 625–630.

Nelson, N. W., Hoelzle, J. B., McGuire, K. A., Slim, A. H., Goldman, D. J., Ferrier-Auerbach, A. G., … Sponheim, S. R. (2011). Self-report of psychological function among OEF/OIF personnel who also report combat-related concussion. *The Clinical Neuropsychologist*, *25*, 716–740.

Orazem, R. J., Frazier, P. A., Schnurr, P. P., Oleson, H. E., Carlson, K. F., Litz, B. T., & Sayer, N. A. (2017). Identity adjustment among Afghanistan and Iraq war veterans with reintegration difficulty. *Psychological Trauma: Theory, Research, Practice, and Policy*, *9*(Suppl 1), 4–11.

Randolph, C. (1998). *Repeatable battery for the assessment of neuropsychological status manual*. San Antonio, TX: The Psychological Corporation.

Rogers, R., Gillard, N.D., Berry, D. T. R., & Granacher, R.P. Jr. (2011). Effectiveness of the MMPI-2-RF validity scales for feigned mental disorders and cognitive impairment: A known-groups study. *Journal of Psychopathology and Behavioral Assessment*, *33*, 355–367. doi: 10.1007/s10862-011-9222-0

Rosnow, R. L., Rosenthal, R., & Rubin, D. B. (2000). Contrasts and correlations in effect-size estimation. *Psychological Science*, *11*, 446–453.

Russo, A. C. (2013). Ethical, legal, and risk management considerations in neuropsychological assessments of veterans. *Psychology Injury and Law*, *6*, 21–30.

Russo, A. C. (2018). A practitioner survey of department of veteran affairs psychologists who provide neuropsychological assessments. *Archives of Clinical Neuropsychology*, *33*, 1046–1059.

Russo, A.C. (2017). A practitioner survey of department of veterans affairs psychologists who provide neuropsychological assessments. *Archives of Clinical Neuropsychology*, *33*, 1046–2059. doi: 10.1093/arclin/acx139

Schroeder, R. W., Buddin, W. H., Hargrave, E. J., VonDran, E. B., Campbell, C. J., & Baade, L. E. (2013). Efficacy of test of memory malingering trial 1, trial 2, the retention trial, and the albany consistency index in a criterion group forensic neuropsychological sample. *Archives of Clinical Neuropsychology*, *28*, 21–29.

Sharf, A. J., Rogers, R., Williams, M. M., & Henry, S. A. (2017). The effectiveness of the MMPI-2-RF in detecting feigned metnal disorders and cognitive deficits: A meta-analysis. *Journal of Psychopathology and Behavioral Assessment*, *39*, 441–455.

Silverberg, N. D., Wertheimer, J. C., & Fichtenberg, N. L. (2007). An effort index for the Repeatable Battery for the Assessment of of Neuropsychological Status (RBANS). *Clinical Neuropsychology*, *21*, 841–854.

Simms, L. J., Casillas, A., Clark, L. A., Watson, D., & Doebbeling, B. N. (2005). Psychometric evaluation of the restructured clinical scales of the MMPI-2. *Psychological Assessment*, *17*, 345–358.

Singhal, A., Green, P., Ashaye, K., Shankar, K., & Gill, D. (2009). High specificity of the medical symptom validity test in patients with very severe memory impairment. *Archives of Clinical Neuropsychology*, *24*, 721–728.

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

Sollman, M. J., & Berry, D. T. R. (2011). Detection of inadequate effort on neuropsychological testing: A meta-analytic update and extension. *Archives of Clinical Neuropsychology*, *26*, 774–789.

Tarescavage, A. M., Wygant, D. B., Gervais, R. O., & Ben-Porath, Y. S. (2013). Association between the MMPI-2 Restructured Form (MMPI-2-RF) and malingered neurocognitive dysfunction among non-head injury disability claimants. *The Clinical Neuropsychologist*, *27*, 313–335.

Tombaugh, T. N. (1996). *The Test of Memory Malingering*. Los Angeles: Western Psychology Corporation.

Wershba, R. E., Locke, D. E., & Lanyon, R. I. (2015). Analysis of Minnesota Multiphasic Personality Inventory-2 Restructured Form response bias indicators as suppressors or moderators in a medical setting. *Psychological Assessment*, *27*, 733–737.

Wiggins, C. W., Wygant, D. B., Hoelzle, J. B., & Gervais, R. O. (2012). The more you say the less it means: Over-reporting and attenuated criterion validity in a forensic disability sample. *Psychological Injury & Law*, *4*, 12–23.

Wright, C. V., Beattie, S. G., Galper, D. I., Church, A. S., Bufka, L. F., Brabender, V. M., & Smith, B. L. (2017). Assessment practices of professional psychologists: Results of a national survey. *Professional Psychology: Research and Practice*, *48*, 73–78.

Wygant, D. B., Anderson, J. L., Sellbom, M., Rapier, J. L., Allegeier, L. M., & Granacher, R. P. (2011). Association of the MMPI-2 Restructured Form (MMPI-2-RF) validity scales with structured malingering criteria. *Psychological Injury and Law*, *4*, 13–23.

Wygant, D. B., Ben-Porath, Y. S., & Arbisi, P. A. (2004). Development and initial validation of a scale to detect somatic over-reporting. In 39th Annual Symposium of Recent Developments in the Use of MMPI-2 and the MMPI-A, Minneapolis, MN.

Wygant, D. B., Sellbom, M., Gervais, R. O., Ben-Porath, Y. S., Stafford, K. P., Freeman, D.B., & Heilbronner, R. L. (2010). Further validation of the MMPI-2 and MMPI-2-RF response bias scale: Findings from disability and criminal forensic settings. *Psychological Assessment*, *22*, 745–756. doi:10.1037/a0020042

Young, J. C., Roper, B. L., & Arentsen, T. J. (2016). Validity testing and neuropsychology practice in the VA healthcare system: Results from recent practitioner survey. *The Clinical Neuropsychologist*, *30*, 497–514.