Psychological Services

Patterns of Minnesota Multiphasic Personality Inventory–2–Restructured Form (MMPI-2-RF) Validity Scale Elevation Across Veterans Seen in a Veterans Affairs (VA) Assessment Clinic: The Impact of Compensation Status

Keegan J. Diehl, Paul B. Ingram, Louis A. Pagano, Jr., and Hunter J. Gideon

Online First Publication, February 24, 2025. https://dx.doi.org/10.1037/ser0000946

Psychological Services

ISSN: 1541-1559
https://doi.org/10.1037/ser0000946

Patterns of Minnesota Multiphasic Personality Inventory–2–Restructured Form (MMPI-2-RF) Validity Scale Elevation Across Veterans Seen in a Veterans Affairs (VA) Assessment Clinic: The Impact of Compensation Status

Keegan J. Diehl[1], Paul B. Ingram[1, 2], Louis A. Pagano Jr.[3], and Hunter J. Gideon[1]
[1] Department of Psychological Sciences, Texas Tech University
[2] United States Department of Veteran Affairs, Eastern Kansas Veteran Affairs Health Care System, Leavenworth, Kansas, United States
[3] United States Department of Veteran Affairs, Research and Development Service, St. Cloud Veteran Affairs Health Care System (Veterans Integrated Service Networks 23), St. Cloud, Minnesota, United States

The purpose of this investigation is to provide descriptive information on veteran response styles for compensation and pension (C&P) evaluations Veterans Affairs (VA) referral types using the Minnesota Multiphasic Personality Inventory–2–Restructured Form (MMPI-2-RF), which has well-supported embedded validity scales capturing invalid response styles. The total sample included 356 veterans from a single VA psychological testing clinic who were administered the MMPI-2-RF during a broader psychological evaluation, with 201 veterans undergoing C&P evaluations. This study examines frequencies of protocol invalidity based on the MMPI-2-RF's validity scales and provides comprehensive descriptive findings on validity scale scores across appointment types (i.e., C&P and non-C&P). Three distinct trends emerged: (1) Veterans generally produced valid MMPI-2-RF profiles, (2) when more than one elevation emerges, it is likely to be thematically consistent (e.g., overreporting scales), and (3) overreporting generally captured the highest frequency of validity scale elevations relative to underreporting or noncontent-based invalid responding. Implications and limitations for practice and the utility of the MMPI-2-RF within VA testing clinics are discussed.

Public Significance Statement
This study examined veteran response styles in Veterans Affairs compensation and pension evaluations using the Minnesota Multiphasic Personality Inventory–2–Restructured Form. Results revealed that most veterans produced valid, interpretable test profiles; however, overreporting was the most frequent invalid response style regardless of compensation and pension status (~20%–25% of cases). Findings support constructing context-specific comparison groups to better understand patterns and influences of response validity.

Keywords: veteran health service, assessment, Minnesota Multiphasic Personality Inventory–2–Restructured Form

John R. McQuaid served as action editor.
Keegan J. Diehl https://orcid.org/0009-0007-4510-497X
Paul B. Ingram https://orcid.org/0000-0002-5409-4896
Louis A. Pagano Jr. https://orcid.org/0009-0007-9314-2356

This study was not preregistered. The data that support the findings of this study are available from the corresponding author upon reasonable request. Materials and analysis code for this study are not available. A portion of these results was presented at the 2023 Combat posttraumatic stress disorder Conference.

The authors have no other known conflicts of interest to disclose. Paul B. Ingram serves on the advisory board for the Personality Assessment Inventory, published by Psychological Assessment Resources Inc., a competitor of the Minnesota Multiphasic Personality Inventory family of measures. He has also received research support from Pearson Clinical Assessments and the University of Minnesota Press, distributor and publisher of the Minnesota Multiphasic Personality Inventory, although none were secured or used for this project. The contents of this article do not represent the views of the U.S. Department of Veterans Affairs or the U.S. Government.

Keegan J. Diehl played a lead role in data curation, formal analysis, software, visualization, and writing–original draft and an equal role in conceptualization, investigation, and writing–review and editing. Paul B. Ingram played a lead role in conceptualization, resources, supervision, and writing–review and editing, a supporting role in data curation, and an equal role in methodology. Louis A. Pagano Jr. played a lead role in investigation and project administration, a supporting role in writing–review and editing, and an equal role in conceptualization, data curation, methodology, and resources. Hunter J. Gideon played a supporting role in writing–review and editing and an equal role in methodology.

Correspondence concerning this article should be addressed to Paul B. Ingram, Department of Psychological Sciences, Texas Tech University, 2700 18th Street, Lubbock, TX 79410, United States. Email: pbingram@gmail.com

Veterans have substantial health service needs. Epidemiological estimates consistently find veterans having more frequent posttraumatic stress (Magruder et al., 2005; Wisco et al., 2014), depression (Liu et al., 2019), substance misuse concerns (Eisen et al., 2012), and higher mortality rates from death by suicide (with veteran suicide reaching unprecedented prevalence; Nichter et al., 2021) relative to civilian populations, for instance. Recent wars have further increased this historic area of service need, with evidence suggesting greater mental health (mh) needs relative to those of other service eras (Doran et al., 2017; Seal et al., 2009). Higher prevalences of disorders, more frequent comorbid health issues, and greater functional impairment are, broadly, the unique and critical areas of need facing current and future veterans (e.g., Hoge & Warner, 2014; Inoue et al., 2022). Each of these problems impacts treatment effectiveness and the ability to match needs to available services for veterans. Due in part to noninitiation of services (e.g., Keller & Tuerk, 2016), some estimates suggest that up to 60% of military and veteran populations do not receive needed services despite positive psychiatric need screeners which identify areas of need (see Tanielian et al., 2016).

The Veterans Affairs (VA) medical system is essential for providing health care services to U.S. veterans. One pathway to accessing these benefits and disability payments is through the Veterans Benefit Administration (VBA). To determine eligibility for service-connected health care benefits and disability compensation, the VBA requires comprehensive compensation and pension (C&P) evaluations to determine level of service connection (SC). The primary purpose of C&P examination is to establish diagnosis and offer a medical opinion on nexus to military service (i.e., diagnosed condition is "more likely than not" related to the veteran's military service), guided by the Disability Benefits Questionnaire. Importantly, C&P examinations are distinct from typical psychodiagnostics testing in that testing is designed to aid the VBA in making accurate benefit determinations rather than the provision of health care services to veterans (Worthen & Moering, 2011). However, there are several concerns related to this inherently forensically enmeshed assessment process. First, effective diagnosis is complicated by the enmeshed disability system of the VA (Ray, 2014, 2017), which may result in a misallocation of resources if evaluations are not done effectively (Wandler, 2013). Approximately 5 million veterans receive VA-provided benefits resultant from service, increasing about three-fold over the past 15 years and totaling a yearly cost of over $165 billion (US Department of Veterans Affairs, 2023). Such budgetary estimates represent only 27% of service-connected veterans (US Bureau of Labor Statistics, 2023), suggesting that a majority of fiscal resources are used for a minority of veterans. Moreover, there is also notable concern in the VA over efforts by some individuals to obtain earned disability benefits through symptom misrepresentation (Ray, 2014, 2017). Accordingly, the VA uses C&P evaluations to determine SC benefits and corresponding disability payment (see Worthen & Moering, 2011).

Evidence-based psychological assessment is a critical part of C&P evaluations. In addition to the increased accuracy of these actuarial methods (see Grove, 2005, for discussion of this issue), evidence-based assessment also offers a means to detect invalid responding, which can bias diagnostic and treatment recommendations (Rogers & Bender, 2018). Detection of response invalidity ensures that support resources are allocated correctly, with overreporting being the most common way responses are invalid in the VA (e.g., Denning & Shura, 2019; DeViva & Bloem, 2003; Frueh et al., 1997; Ingram et al., 2020; Shura et al., 2021). Undetected, invalid responding can bias clinicians' diagnostic and interpretive conclusions from testing (see Nelson et al., 2010). Within C&P evaluations, multiscale inventories with embedded validity scales (e.g., Minnesota Multiphasic Personality Inventory [MMPI] family of measures) are common assessment instruments used by practitioners (Russo, 2018), given such measures offer methods for assessing both psychopathological symptom need and potential response distortion (Ray, 2014, 2017; Worthen & Moering, 2011). Accordingly, in a national estimate of overreporting (the most frequent form of invalid responding), data suggest that between 23.3% and 56.8% (Mental Health Integrated Care and posttraumatic stress disorder [PTSD] Clinic Team treatment clinics, respectively) of veterans will produce invalidated test protocols depending upon setting (Ingram et al., 2020).

The MMPI family of instruments is one particularly common broadband measure used in C&P evaluations, as well as other diagnostic exams, within the VA system (Ingram et al., 2020; Russo, 2014). The MMPI–2–Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008/2012; Tellegen & Ben-Porath, 2008/2012) is a recent iteration of the MMPI family of measures available within the VA testing administration portal. Accordingly, the overreporting scales of the MMPI-2-RF are often the most frequently relied upon scales to determine symptom-based invalid responding within neuropsychological clinics (Russo, 2018), comprised of five overreporting scales assessing somatic (infrequent somatic [Fs], Symptom Validity Scale [FBS]), cognitive (Response Bias Scale [RBS], FBS), and psychological (Infrequent Responses [F], Infrequent Psychopathology Responses [Fp]) symptoms (see Burchett & Bagby, 2022, for a review of these scales and research supporting them).

MMPI-2-RF validity scales appear effective in capturing invalid responding across a variety of populations and evaluative contexts, with meta-analytic effect estimates (Cohen's d) ranging from .88 (RBS and Fs) to 1.32 (infrequent psychopathology responses [Fp-r]), falling within the moderate to large classification range (Ingram & Ternes, 2016; Sharf et al., 2017). In a nationwide sample of veterans (n = 17,640), Ingram et al. (2020) provided descriptive data on the frequency of MMPI-2-RF scale elevations observing a few particularly notable trends: (a) Overreporting is the most frequent type of response invalidity within the VA and (b) elevation rates of the validity scales vary substantially by the evaluative context. One notable limitation of Ingram et al.'s (2020) work is that not all types of psychological service clinics were included in their survey. Thus, despite variation across clinical contexts which may impact validity detection accuracy (Boress et al., 2024; Ingram & Ternes, 2016), not all VA clinics are equally equipped with the data needed to appropriately make determinations of probable engagement, effort, or symptom validity. For instance, neuropsychological clinics, where most of the validations of the MMPI-2-RF overreporting scales have been conducted (e.g., Gervais et al., 2007; Schroeder et al., 2012; Sullivan et al., 2013), were not provided (Ingram et al., 2020). Likewise, general psychological assessments where many VA assessments are conducted were not evaluated. Due to limitations in the data available to Ingram et al. (2020), referral reasons and C&P status of the evaluation (e.g., if it was a formal disability evaluation or for another clinical service) were not evaluated. This study evaluates validity scale elevations

between those being seen for a C&P examination and those being seen for other psychological assessments. By addressing this gap in understanding validity considerations within the VA, increasingly accurate and effective context-specific interpretation may provide a more nuanced understanding of veteran needs.

The Present Study

The present study expands research on the MMPI-2-RF in veterans seen at a VA psychological testing clinic in a Midwestern Department of VA Medical Center to (a) analyze validity score elevations on the MMPI-2-RF and contrast those between C&P/non-C&P groups and (b) provide descriptive data of a large novel clinic setting, which expands the available comparative data from other VA service clinics. We seek to accomplish such goals by providing the frequency of elevated MMPI-2-RF validity scale scores across each of the nine validity scales, using standard interpretive thresholds (i.e., Ben-Porath & Tellegen, 2008/2012), for both evaluative contexts (i.e., C&P/non-C&P). This provides important interpretive context as we present and compare the percentage of veterans producing elevated noncontent-based (i.e., variable response inconsistencies/true response inconsistencies [VRIN-r/TRIN-r]) and content-based overreporting (overreported pathology [infrequent responses; F-r/Fp-r], somatic and cognitive overreporting [FBS-r/RBS/Fs]), and underreporting (Uncommon Virtues [L-r]/Adjustment Validity [K-r]) in a large novel VA assessment setting. Furthermore, we provide descriptive data which expand the available comparative data from other VA clinics (see Ingram et al., 2020).

Method

Participants

All procedures were approved by the U.S. Department of Veterans Affairs (IRB#1740048), under the direction of the third author. The study was not preregistered, and materials and analysis code for this study are not available. However, materials will be made available upon reasonable request.

Extracted MMPI-2-RF profiles were a part of assessment batteries from a variety of assessment referrals (e.g., neuropsychological assessments about traumatic brain injury, nonneuropsychological psychodiagnostic examinations for issues such as PTSD or attention deficit disorder, etc.). Nonprobability sampling of records primarily from FY20 to FY23 identified MMPI-2-RF profiles that met the study's inclusion criteria (i.e., veteran status, completed assessment through VA Health Care System, testing battery included the MMPI-2-RF). These cases included both those referred as part of C&P evaluations and those conducted as part of routine clinical care. All MMPI-2-RF testing was administered electronically (or entered for scoring) in the VA Mental Health Assistant and presumed adherent with testing security guidelines as examiners were VA-affiliated assessment psychologists. The participants include 356 veterans who were assessed in a VA psychological testing clinic (refer to Table 1 for participant demographics). Of those, 201 (56.5%) were undergoing C&P evaluations at the time of testing. Demographically, most participants identified as White (90.7%) and non-Hispanic (1.1% Hispanic), with a mean age of 48 (SD = 15.3; range = 20 through 81 years). The trends of SC are as follows: 7.3% were not connected, 26.7% were connected less than 50%, and the remaining 65.2% were 50%–100% service-connected. The rates of SC were approximately equal across C&P and non-C&P with a slightly higher rate of 50%–100% SC in the C&P group (70.1% vs. 58.7%).

Exclusionary criteria exclusively consisted of invalid responding for noncontent-based (NCBR) invalid responding (i.e., VRIN-r $\geq$ 80 and/or TRIN-r $\geq$ 80) based on the measure's interpretive guidelines (Ben-Porath & Tellegen, 2008/2012; Tellegen & Ben-Porath, 2008/2012). This criterion removed approximately 5% (n = 10) of C&P MMPI profiles, while only about 1% (n = 2) of non-C&P profiles were excluded (a statistically significant difference, see Table 2).

Table 1
Demographic Characteristics of Participants

Participant characteristic	Full sample	C&P	Non-C&P
Mean age (*SD*)	48 (15.3)	49.4 (15.17)	47.15 (15.42)
Race and ethnicity			
American Indian or Alaska Native	2 (0.6%)	1 (0.5%)	1 (0.6%)
Black	12 (3.4%)	11 (5.5%)	1 (.6%)
Hispanic	4 (1.1%)	2 (1%)	2 (1.3%)
Multiracial	6 (1.7%)	3 (1.5%)	3 (1.9%)
Native Hawaiian or other Pacific Islander	8 (2.2%)	6 (3%)	2 (1.3%)
White	323 (90.7%)	177 (88.1%)	146 (94.2%)
Sex			
Male	295 (82.9%)	167 (83.1%)	128 (82.6%)
Female	55 (15.4%)	28 (13.9%)	27 (17.4%)
Missing	6 (1.7%)	6 (3%)	0
Period of service			
Vietnam era	63 (17.7%)	36 (17.9%)	27 (17.4%)
Post Vietnam	36 (10.1%)	22 (10.9%)	14 (9%)
Gulf War	257 (72.2%)	143 (71.1%)	114 (73.5%)
Service connection			
NSC	29 (8.1%)	13 (6.5%)	16 (10.3%)
<50%	95 (26.7%)	47 (23.4%)	48 (31%)
50%–100%	232 (65.2%)	141 (70.1%)	91 (58.7%)

Note. C&P = compensation and pension evaluation referral type; non-C&P = noncompensation and pension referral type; NSC = nonservice connected.

Table 2
Descriptive Statistics and Scale Elevation Comparison by Appointment Type

	Total sample (N = 356, 346)			C&P (n = 201, 191)			Non-C&P (n = 155, 153)						
Scale	*M*	*SD*	% ≥ RCS	*M*	*SD*	% ≥ RCS	*M*	*SD*	% ≥ RCS	*t*(*df*)	Hedge's *g* [95% CI]	χ2 (1)	ϕ ≥ RCS
VRIN-r	52.6	10.2	1.7	53.7	10.7	3	51.2	9.3	0	2.31 (354)*	.25 [.04, .46]	4.71*	.12
TRIN-r	58.0	7.7	1.7	58.3	8.0	2	57.6	7.4	1.3	0.86 (354)	.09 [−.12, .30]	0.26	.03
F-r	83.7	25.6	7.6	86.6	24.8	8.6	80.7	26.5	6.5	2.30 (342)*	.25 [.04, .46]	0.41	.04
Fp-r	65.4	19.2	6.4	66.3	19.2	7.3	64.6	19.4	5.2	1.00 (342)	.11 [−.10, .32]	0.63	.05
Fs	76.0	22.5	13.1	79.7	21.8	15.2	71.7	22.5	10.5	3.52 (342)*	.38 [.17, .60]	1.67	.07
FBS-r	72.6	14.2	2.3	75.6	13.6	3.7	69.1	14.0	0.7	4.48 (342)*	.49 [.27, .70]	3.39	.10
RBS	80.0	18.0	15.4	83.0	17.4	18.3	76.8	18.5	11.8	3.48 (342)*	.38 [.16, .59]	2.81	.09
L-r	53.5	9.5	0.3	54.5	9.1	0.5	52.2	9.7	1.3	−2.28 (342)*	.25 [.03, .46]	0.6	.04
K-r	40.8	9.6	0.3	39.7	8.4	0	41.9	10.9	0.7	−2.19 (280.69)*	.24 [.03, .46]	1.25	.06

Note. Two separate *ns* represent changes in cases due to removal of noncontent-based invalid responding. Invalid noncontent-based invalid responding (VRIN-r/TRIN-r) profiles were excluded from calculation of over- and underreporting cumulative frequency. C&P = compensation and pension evaluation referral type; RCS = recommended cut score; CI = confidence interval; VRIN-r = variable response inconsistencies; TRIN-r = true response inconsistencies; F-r = infrequent responses; Fp-r = infrequent psychopathology responses; Fs = infrequent somatic; FBS-r = Symptom Validity Scale; RBS = Response Bias Scale; L-r = Uncommon Virtues; K-r = Adjustment Validity.
* $p < .05$.

After exclusions, the C&P group consisted of 191 participants and the non-C&P group consisted of 153 participants.

Measures

Minnesota Multiphasic Personality Inventory–2–Restructured Form (Ben-Porath & Tellegen, 2008/2012; Tellegen & Ben-Porath, 2008/2012)

The MMPI-2-RF is a 338-item self-report inventory, hierarchically capturing a broad array of personality and psychological traits, aligning to contemporary models of psychopathology. It includes 42 substantive scales that measure a variety of clinical criteria and nine embedded validity indicators that serve to detect response patterns indicative of noncredible responses. The MMPI-2-RF validity scales have demonstrated moderate to large magnitude between-group differences between bona fide psychopathology and invalid responding (see Ingram & Ternes, 2016; see also Sharf et al., 2017). Scales capturing content-based overreporting include F-r, Fp-r, Fs responses, symptom validity (FBS-r), and RBS. F-r includes 32 items that detects infrequent responses endorsed by 10% or less of the normative sample, while Fp-r, a revised version of the MMPI-2 Fp scale (Arbisi & Ben-Porath, 1995), measures infrequent psychopathology responses endorsed by 20% or less of psychiatric patients (Arbisi & Ben-Porath, 1995). Fs comprises 16 items to assess overreporting of somatic responses by using items that capture atypical physical symptom endorsements such as pain that are not otherwise endorsed in medical samples (Wygant et al., 2009). The FBS-r scale contains items that were identified through frequency counts and by observation of malingerer response patterns in civil forensic settings (Lees-Haley et al., 1991). The 28-item RBS correlates with scores that are below published performance validity test (PVT) cutoffs (Gervais et al., 2007). The construct validity and reliability of the MMPI-2-RF have been examined through numerous studies (see Tellegen & Ben-Porath, 2008/2012) and are frequently used in C&P evaluation decisions (see Ray, 2017, for a discussion of the MMPI-2-RF in a variety of veteran-specific context).

Procedure and Planned Analyses

To compare mean scores of each MMPI-2-RF indicator, descriptive analyses of both the C&P ($n_{\text{NCBR-included}}$ = 201, $n_{\text{NCBR-excluded}}$ 191) and non-C&P ($n_{\text{NCBR-included}}$ = 155, $n_{\text{NCBR-excluded}}$ 153) groups were conducted in SPSS. Independent *t* test compared mean scores on each of the MMPI-2-RF validity scales, with small, medium, and large effect sizes classified using Hedge's *g* ($g \geq 0.2$, $g \geq 0.5$, $g \geq 0.8$, respectively; Cohen, 1988). Validity scale domain (noncontent, overreporting, underreporting) elevations were compared across recognized cut scores via chi-square, with Field's (2013) criteria anchoring interpretation of weak ($\phi \geq .05$), moderate ($\phi \geq .1$), and strong effects ($\phi \geq .15$). VRIN-r and TRIN-r data were excluded and analyzed separately to prevent noncontent responses from influencing the data of the other scales.

Cumulative frequencies were examined using independent *t* tests across the total sample of veterans against those in the C&P group. The cumulative percentages of administrations at or above recommended *t* scores were examined to identify drop-off in response frequencies in the C&P group when mirrored against the total sample. Elevation frequencies were identified using the recommended cut score (RCS) for each embedded overreporting scale, respectively (see the Measures section). Coelevation of validity scales across each domain was compared using descriptive analysis across non-C&P and C&P groups. Using cumulative frequency of responses as cut scores, each domain's drop-off can be clearly identified and compared to the total elevated scales across each domain within both groups. Overreporting and underreporting rates were compared after the exclusion of VRIN-r and TRIN-r.

Results

Descriptive statistics for the validity scales, as well as cumulative elevation frequencies associated with various *t*-score cutoffs, are provided in Tables 2–4. Protocols invalidated for exceeding recommended TRIN-r or VRIN-r scores (3% of cases; n = 12) were removed from frequency calculation on the over- and underreporting scales. The overreporting scales show the highest levels indicating protocol

Table 3
Descriptive Statistics and Scale Elevation Frequencies for the Total Sample

	Total sample			Cumulative % of administration at or above a given *t*-score value						
Scale	*M*	*SD*	% > RCS	60	70	80	90	100	110	120
VRIN-r	52.6	10.2	1.7	18.5%	5.9%	**1.7%**	0.3%	0%	0%	0%
TRIN-r	58.0	7.7	1.7	27.5%	9.3%	**1.7%**	0%	0%	0%	0%
F-r	83.7	25.6	7.6	83.7%	62.5%	49.4%	37.2%	25.6%	16.6%	**7.6%**
Fp-r	65.4	19.2	6.4	44.5%	31.1%	18.3%	10.5%	**6.4%**	4.1%	1.2%
Fs	76.0	22.5	13.1	68.0%	54.9%	50.0%	28.2%	**13.1%**	7.0%	3.5%
FBS-r	72.6	14.2	2.3	79.4%	61.1%	28.8%	10.2%	**2.3%**	0%	0%
RBS	80.0	18.0	15.4	84.9%	70.9%	47.7%	26.7%	**15.4%**	4.4%	1.7%
L-r	53.5	9.5	0.3	26.2%	5.8%	**1%**	.03%	0%	n/a	n/a
K-r	40.8	9.6	0.3	4.3%	**0.3%**	0%	n/a	n/a	n/a	n/a

Note. $N = 356$. Invalid noncontent-based invalid responding (VRIN-r/TRIN-r) profiles were excluded from calculation of over- and underreporting cumulative frequency. Bolded values indicate that scores at or above this *t* score invalidate an MMPI-2-RF protocol (Ben-Porath & Tellegen, 2008/2012). RCS = recommended cut score; VRIN-r = variable response inconsistencies; TRIN-r = true response inconsistencies; F-r = infrequent responses; Fp-r = infrequent psychopathology responses; Fs = infrequent somatic; RBS = Response Bias Scale; FBS-r = Symptom Validity Scale; L-r = Uncommon Virtues; n/a = indicates that a score at this level is not possible; K-r = Adjustment Validity; MMPI-2-RF = Minnesota Multiphasic Personality Inventory–2–Restructured Form.

invalidity, regardless of appointment type (ranging from 2.3% [FBS-r] to 15.4% [RBS] in the total sample) using the most conservative cut scores provided within the interpretive manual (i.e., F-r $\geq$ 120, Fp-r $\geq$ 100, Fs $\geq$ 100, RBS $\geq$ 100, FBS-r $\geq$ 100, L-r $\geq$ 80, K-r $\geq$ 70). Underreporting infrequently elevated beyond RCS values (1% and .3% on L-r and K-r, respectively). Similarly, NCBR invalid responding infrequently exceeded interpretable cut scores (1.7% for both VRIN-r and TRIN-r). Table 5 presents the cumulative frequency of validity scale elevations (i.e., not exceeding any of the most stringent RCSs). Three distinct trends emerge within the total sample. First, regardless of appointment type, veterans generally produced valid MMPI-2-RF profiles (see Table 5). Second, when elevations are likely to occur, concurrent elevations within the same content scales (e.g., overreporting scales) are likely relative to other content scales (e.g., overreporting and underreporting). Third, overreporting generally captured the highest frequency of validity scale elevations relative to underreporting or NCBR invalid responding.

Table 2 provides an overview of elevation differences based on appointment type (i.e., C&P vs. non-C&P). Generally, the evaluative context of C&P appointment type coincided with larger validity scale elevation magnitude and a greater frequency of invalidated protocols. C&P evaluations frequently had higher scores across scales, with effect sizes generally in the small to moderate range (Hedge's *g* ranging from .11 [Fp-r] through .49 [FBS-r]). Likewise, those in C&P evaluations were more likely to elevate (i.e., $\phi \geq$ RCS) the MMPI-2-RF validity scales to their interpretive levels between appointment types. However, those elevation frequency discrepancies were generally weak to moderate in magnitude for the overreporting (ϕ ranging from .04 [F-r] through .10 [FBS-r]) and NCBR (ϕ ranging from .03 [TRIN-r] through .12 [VRIN-r]), suggesting minimal clinical difference across evaluation type and those cut scores. Inversely, non-C&P elevations observed larger magnitude elevations on underreporting scales (L-r g = .25; K-r g = .24) and weak associations in exceeding RCS relative to C&P evaluations (L-r ϕ = .04; K-r ϕ = .06).

Table 4
Descriptive Statistics and Scale Elevation Frequencies for Compensation and Pension Evaluations

	C&P only			Cumulative % of administration above a given *t*-score value						
Scale	*M*	*SD*	% > RCS	60	70	80	90	100	110	120
VRIN-r	53.7	10.7	3	21.9%	7.0%	**3.0%**	0.5%	0%	0%	0%
TRIN-r	58.3	8.0	2	28.4%	10.4%	**2.0%**	0.0%	0%	0%	0%
F-r	86.6	24.8	8.6	88.5%	69.6%	56.0%	41.4%	26.2%	18.8%	**8.6%**
Fp-r	66.3	19.2	7.3	46.1%	31.4%	19.4%	9.9%	**7.3%**	4.2%	1.6%
Fs	79.7	21.8	15.2	78.5%	63.4%	45.5%	30.9%	**15.2%**	8.9%	4.7%
FBS-r	75.6	13.6	3.7	86.9%	71.2%	33.5%	13.1%	**3.7%**	0%	0%
RBS	83.0	17.4	18.3	89.5%	79.1%	54.5%	30.9%	**18.3%**	5.8%	2.1%
L-r	54.5	9.1	0.5	30.4%	6.3%	**0.5%**	0%	0%	n/a	n/a
K-r	39.7	8.4	0	2.1%	**0%**	0%	n/a	n/a	n/a	n/a

Note. $n = 201$. Invalid noncontent-based invalid responding (VRIN-r/TRIN-r) profiles were excluded from calculation of over- and underreporting cumulative frequency. Bolded values indicate that scores at or above this *t* score invalidate an MMPI-2-RF protocol (Ben-Porath & Tellegen, 2008/2012). C&P = compensation and pension evaluation referral type; RCS = recommended cut score; VRIN-r = variable response inconsistencies; TRIN-r = true response inconsistencies; F-r = infrequent responses; Fp-r = infrequent psychopathology responses; Fs = infrequent somatic; FBS-r = Symptom Validity Scale; RBS = Response Bias Scale; L-r = Uncommon Virtues; n/a = indicates that a score at this level is not possible; K-r = Adjustment Validity; MMPI-2-RF = Minnesota Multiphasic Personality Inventory–2–Restructured Form.

Table 5
Cumulative Frequency of Validity Scale Elevations Indicating Invalidity

Cumulative frequency	0	1	2	3	4	5	6	7
Non-C&P								
No. of noncontent responding	98.7%	1.3%						
No. of overreporting	77.4%	21.3%	9.0%	3.2%	0.6%	0%		
No. of underreporting	98.1%	1.9%						
Total scales elevated	75.5%	24.5%	10.3%	3.2%	0.6%	0%		
C&P								
No. of noncontent responding	95.0%	5.0%						
No. of overreporting	69.2%	25.9%	12.4%	7.0%	4.5%	0.5%		
No. of underreporting	99.5%	0.5%						
Total scales elevated	68.7%	31.3%	13.9%	7.5%	4.5%	0.5%		

Note. Standard cut scores (Ben-Porath & Tellegen, 2008/2012): Cannot Say ≥ 18, VRIN-r ≥ 80, TRIN ≥ 80, F-r ≥ 120, and Fp-r ≥ 100. Invalid noncontent-based invalid responding (VRIN-r/TRIN-r) profiles were excluded from calculation of over- and underreporting cumulative frequency. Non-C&P = non-compensation and pension referral type; C&P = compensation and pension evaluation referral type; VRIN-r = variable response inconsistencies; TRIN-r = true response inconsistencies; F-r = infrequent responses; Fp-r = infrequent psychopathology responses.

Discussion

This study examines MMPI-2-RF validity scale elevation rates across C&P and non-C&P neuropsychological evaluations in a sample of veterans. Previous research has identified the differential utility of validity measures across differing evaluative contexts (see Boress et al., 2024), including for veteran populations (Ingram & Ternes, 2016). Further elucidating the potential impact of evaluative context on validity instruments, and embedded validity scales within multiscale inventories in particular, has notable implications on assessment and warrants examination (see Whiteside & Basso, 2024; e.g., Armistead-Jehle & Buican, 2012; Nelson et al., 2010). Results from our study broadly suggest three distinct trends: (1) Veterans generally produced valid MMPI-2-RF profiles, (2) when multiple elevations emerge, they are likely to be thematically consistent (e.g., multiple overreporting scales) relative to different content scales (overreporting and underreporting elevations), and (3) overreporting scales generally captured the highest frequency of validity scale elevations relative to underreporting or NCBR, approximately 25%–30% of instances in which a validity scale was elevated.

Our results suggest veterans, regardless of appointment type, more frequently produce valid profiles than observed previously in a national sample, collated and presented across a variety of nonneuropsychological assessment–focused VA service clinics (Ingram et al., 2020). However, this difference is relatively small and broadly supports the trends observed previously and suggests patterns are likely generalizable, even if specific elevation rates differ. Discrepancies between these settings are not themselves unsurprising given the documented impact of evaluative context (e.g., Boress et al., 2024). Discrepancies reinforce assessment psychologists' professional responsibility (see Sweet et al., 2021; see also Krishnamurthy et al., 2022) for considering contextual influences (e.g., appointment type) in validity scale score interpretation (see Ingram & Ternes, 2016). Similarly, one unique assessor factor inherent to the current sample (i.e., VA-affiliated assessment-specialized psychologists) may also explain differential elevation rates relative to Ingram et al. (2020). That is, our sample's combination of assessment psychologists as well as board-certified neuropsychologists' unique and advanced training in assessment (see Grote et al., 2016) may observe increased fidelity with standardized assessment materials relative to other assessors, thus establishing unique testing experience and approach to ensuring differentially valid performance. The influence of the evaluator on test validity requires further study to understand what, if any, educational and approach factors alter the evaluatee's experience and self-reporting behavior.

Observed elevation rates differ in their frequency between C&P and non-C&P focused evaluations; however, these differences were relatively small in effect and unlikely to cause major variations in test taking approach with respect to needed cut scores. One potential reason for a lack of difference between these two groups in score performance, given the generally similar scores and discrepant referral reasons, is the enmeshment of disability processes with routine care within the VA (see Worthen & Moering, 2011). Thus, despite small effects magnitudes, the inherent forensic enmeshment of C&P evaluations remains critical to consider during service provision (Ray, 2017), and inclusion of multiple sources of performance and symptom validity testing to support diagnostic or clinical decision making remains necessary (see Sweet et al., 2021). First, examining rates of NCBR, C&P-specific evaluations observed increased protocol invalidity from NCBR relative to their non-C&P comparison (ϕ .12). While our methods are unable to explicate the potential operative factors for such differential NCBR elevations, these results may reflect common referral reasons across VA neuropsychological clinics, including attention and memory difficulties (Burchett et al., 2016; cf. Frederick et al., 2000, for a thorough discussion regarding impaired testing engagement impacting the assessment process).

Broadly, our base rate for validity scale elevations was somewhat higher relative to the Technical Manual comparison groups (Tellegen & Ben-Porath, 2012)[1]; however, these elevations more closely approximated veteran MMPI validity scale descriptives in Ingram et al. (2020), which is drawn from a more contemporary sample. Within each validity scale's conceptual domain (i.e., somatic, cognitive, psychological), psychopathology invalidity rates occurred less frequently (F-r % > RCS = 7.6, Fp-r % > RCS = 6.4), relative to somatic (Fs % > RCS = 13.1, FBS-r % > RCS = 2.3) and cognitive

[1] The all-male VA comparison groups reported in the technical manual (Tellegen & Ben-Porath, 2008/2012, p. 303) list the following standard deviation for the MMPI-2-RF validity scales: VRIN-r 10.0; TRIN-r = 10.5; F-r = 27.5; Fp-r = 15; Fs = 20.5; FBS-r = 15; RBS = 19; L-r = 10; K-r = 10.

(RBS % > RCS = 15.4) overreporting. Given these trends, Fs and RBS appear particularly useful within neuropsychological settings, which is unsurprising given their derivation and previous research establishing their clinical utility (Gervais et al., 2007; Ingram & Ternes, 2016; Sharf et al., 2017; Wygant et al., 2009). Such trends are also expected given current theoretical assumptions and practice standards related to validity testing (see Sweet et al., 2021). Interestingly, other research on overreporting has found that despite these domains differing in some studies, validity scales across domains tend to perform similarly in their effect size and classification accuracy (Gaines et al., 2013; Ingram et al., 2024; Morris et al., 2022).

In all cases, overreporting is the most common form of invalid responding. Underreporting was slightly higher in the non-C&P groups (ϕs = .04 [L-r], .06 [K-r]); however, no notable differences in rate were evident for the K-r and L-r scales as one might expect in the case of maladjustment-based underreporting (Keen et al., 2023). Although our data do not provide further information regarding the types of C&P examinations (e.g., mental disorder, PTSD, traumatic brain injury), heterogeneous C&P evaluation within the sample may explain the failure to replicate previous work identifying K-r and L-r's empirical precedent in capturing internalizing and externalizing underreporting, respectively (Detrick & Chibnall, 2014; Keen et al., 2023). Though, it should be noted that the use of the terms over- and underreporting does not infer that those are the only potential explanations for the elevation on the scales and is merely a convention given the MMPI-2-RF's scale names. For instance, elevation on overreporting scales may also be caused by severe distress, noncontent responding, or a cry or request for help, to name a few potential nonpathological, bona fide reasons for elevation.

Limitations and Future Directions

This study should be considered within the scope of its limitations. First, as with similar descriptive articles focused on understanding individual service clinic patterns (e.g., Ingram et al., 2020), there were no external criteria (e.g., stand-alone symptom validity tests or PVTs) available to evaluate the performance of MMPI-2-RF validity scales, which deters definitive insights into these elevations. As argued in Ingram et al. (2020), observed invalidity rates may, of course, also result from complex manifestation of bonified psychopathology that may be misidentified as noncredible responding. However, when elevation rates reach RCSs, such an interpretation is less likely because of specificity recommendations used (see Sweet et al., 2021) and corresponding research outcomes (e.g., Ingram & Ternes, 2016). Future studies on compensation status' impact on responding styles would benefit from the use of stand-alone external criteria, both for validity and substantive concerns. Inclusion could provide valuable insight into responding style's impact on interpretive validity for broadband self-report measures and for the assessment process itself (e.g., across other data sources used for C&P evaluations). Research has regularly found that collateral measures are impacted by invalid responding patterns (e.g., over- and underreporting), including for veterans receiving inpatient psychological services (e.g., Forbey et al., 2013). Invalid responding tends also to generalize to other measures of invalid responding with consistent patterns of accuracy emerging after the inclusion of multiple symptom validity test (Gervais et al., 2011) or PVT (Kanser et al., 2024) measures.

Additionally, the study site was highly specific (i.e., a single VA psychological testing clinic), which may limit generalizability. Regardless of the highly specific study site, there is a need to learn more about test administration procedure fidelity across different testing settings, as effective, therapeutic engagement with clients can not only cause therapeutic change itself (e.g., Chen et al., 2020; Neale & Rosenheck, 1995) but also lead to more disclosure (Kelly & Yuan, 2009) generally, in both interview (Garbarski et al., 2016) and self-report surveys (Bowling et al., 2016). Research has suggested that while perceived competency is the greatest predictor of intention to conduct assessment in one's career (Ingram et al., 2022), performance-based competency, however, is best predicted by hands-on training (Bergquist et al., 2023). As such, it is possible that training backgrounds may differentially predict test fidelity.

Additionally, our results observed considerable variability in validity scale scores (see Table 3). Variability across samples, regardless of sample size, appears to be a common occurrence within VA research (Ingram et al., 2020), and validity detection broadly (Ingram & Ternes, 2016). For instance, Nelson et al. (2010) observed notable variability in overreporting validity scale scores within their veteran sample across different evaluation contexts. In forensic contexts, Nelson et al.'s (2010) *SDs* ranged from 13.7 (Fp-r) to 19.5 (Fs). In clinical contexts, *SDs* ranged from 13.7 (Fp-r) to 22.7 (F-r). In research contexts, *SDs* ranged from 10.2 (Fp-r) to 20.9 (F-r). Similarly, Ingram et al.'s (2020) VA data extraction of 17,640 veterans' MMPI-2-RF profiles reported marked variability in validity scale scores (*SDs* ranging from 15.7 [FBS-r] through 25.0 [F-r]), with even greater variability across various service locations (invalidity rate ranging from 23.3% invalid in mental health integration through 56.8% invalid in PTSD clinical team). Complicating our sample further, and as previously mentioned, the specific C&P evaluation type was not made available, which may explain our observed variability and warrants study about how specific referral reasons predict performance.

Together, precedent and the current work convincingly demonstrate the complexity of measuring responding styles in high-need samples broadly, and veterans specifically. Given this complexity, assessment psychologists should be advised against assuming a simple comparative sample from the literature in clinical decision making. Work constructing localized comparison groups remains critical as understanding of patterns and influences of response validity becomes more clearly understood and integrated into interpretation. Rather, direct inclusion of additional symptom validity test and PVTs in assessment battery construction, consistent with best practice (Sweet et al., 2021), and remaining current on emerging theory and literature related to performance and symptom validity appear advisable (see Whiteside & Basso, 2024).

References

Arbisi, P. A., & Ben-Porath, Y. S. (1995). An MMPI-2 infrequent response scale for use with psychopathological populations: The Infrequency-Psychopathology Scale, F(p). *Psychological Assessment*, *7*(4), 424–431. https://doi.org/10.1037/1040-3590.7.4.424

Armistead-Jehle, P., & Buican, B. (2012). Evaluation context and symptom validity test performances in a U.S. military sample. *Archives of Clinical Neuropsychology*, *27*(8), 828–839. https://doi.org/10.1093/arclin/acs086

Ben-Porath, Y. S., & Tellegen, A. (2012). *MMPI-2-RF: Manual for administration, scoring and interpretation*. University of Minnesota Press. (Original work published 2008)

Bergquist, B. K., Keen, M. A., Ingram, P. B., Morris, N. M., & Schmidt, A. T. (2023). Doctoral students' intention to use assessments in their career: The incremental role of self-reported competence. *Journal of Clinical Psychology*, *79*(2), 374–390. https://doi.org/10.1002/jclp.23415

Boress, K., Gaasedelen, O., Kim, J. H., Basso, M. R., & Whiteside, D. M. (2024). Examination of the relationship between symptom and performance validity measures across referral subtypes. *Journal of Clinical and Experimental Neuropsychology*, *46*(2), 162–171. https://doi.org/10.1080/13803395.2023.2261633

Bowling, N. A., Huang, J. L., Bragg, C. B., Khazon, S., Liu, M., & Blackmore, C. E. (2016). Who cares and who is careless? Insufficient effort responding as a reflection of respondent personality. *Journal of Personality and Social Psychology*, *111*(2), 218–229. https://doi.org/10.1037/pspp0000085

Burchett, D., & Bagby, R. M. (2022). Assessing negative response bias: A review of the noncredible overreporting scales of the MMPI-2-RF and MMPI-3. *Psychological Injury and Law*, *15*(1), 22–36. https://doi.org/10.1007/s12207-021-09435-9

Burchett, D., Dragon, W. R., Smith Holbert, A. M., Tarescavage, A. M., Mattson, C. A., Handel, R. W., & Ben-Porath, Y. S. (2016). "False feigners": Examining the impact of non-content-based invalid responding on the Minnesota Multiphasic Personality Inventory-2 Restructured Form content-based invalid responding indicators. *Psychological Assessment*, *28*(5), 458–470. https://doi.org/10.1037/pas0000205

Chen, J. A., Fortney, J. C., Bergman, H. E., Browne, K. C., Grubbs, K. M., Hudson, T. J., & Raue, P. J. (2020). Therapeutic alliance across trauma-focused and non-trauma-focused psychotherapies among veterans with PTSD. *Psychological Services*, *17*(4), 452–460. https://doi.org/10.1037/ser0000329

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.

Denning, J. H., & Shura, R. D. (2019). Cost of malingering mild traumatic brain injury-related cognitive deficits during compensation and pension evaluations in the veterans benefits administration. *Applied Neuropsychology: Adult*, *26*(1), 1–16. https://doi.org/10.1080/23279095.2017.1350684

Detrick, P., & Chibnall, J. T. (2014). Underreporting on the MMPI-2-RF in a high-demand police officer selection context: An illustration. *Psychological Assessment*, *26*(3), 1044–1049. https://doi.org/10.1037/pas0000013

DeViva, J. C., & Bloem, W. D. (2003). Symptom exaggeration and compensation seeking among combat veterans with posttraumatic stress disorder. *Journal of Traumatic Stress*, *16*(5), 503–507. https://doi.org/10.1023/A:1025766713188

Doran, J. M., Pietrzak, R. H., Hoff, R., & Harpaz-Rotem, I. (2017). Psychotherapy utilization and retention in a national sample of veterans with PTSD. *Journal of Clinical Psychology*, *73*(10), 1259–1279. https://doi.org/10.1002/jclp.22445

Eisen, S. V., Schultz, M. R., Vogt, D., Glickman, M. E., Elwy, A. R., Drainoni, M. L., Osei-Bonsu, P. E., & Martin, J. (2012). Mental and physical health status and alcohol and drug use following return from deployment to Iraq or Afghanistan. *American Journal of Public Health*, *102*(S1), S66–S73. https://doi.org/10.2105/AJPH.2011.300609

Field, A. (2013). *Discovering statistics using SPSS* (4th ed.). Sage Publications.

Forbey, J. D., Lee, T. T. C., Ben-Porath, Y. S., Arbisi, P. A., & Gartland, D. (2013). Associations between MMPI-2-RF validity scale scores and extra-test measures of personality and psychopathology. *Assessment*, *20*(4), 448–461. https://doi.org/10.1177/1073191113478154

Frederick, R. I., Crosby, R. D., & Wynkoop, T. F. (2000). Performance curve classification of invalid responding on the validity indicator profile. *Archives of Clinical Neuropsychology*, *15*(4), 281–300. https://doi.org/10.1093/arclin/15.4.281

Frueh, B. C., Gold, P. B., & de Arellano, M. A. (1997). Symptom overreporting in combat veterans evaluated for PTSD: Differentiation on the basis of compensation seeking status. *Journal of Personality Assessment*, *68*(2), 369–384. https://doi.org/10.1207/s15327752jpa6802_8

Gaines, M. V., Giles, C. L., & Morgan, R. D. (2013). The detection of feigning using multiple PAI Scale elevations: A new index. *Assessment*, *20*(4), 437–447. https://doi.org/10.1177/1073191112458146

Garbarski, D., Schaeffer, N. C., & Dykema, J. (2016). Interviewing practices, conversational practices, and rapport: Responsiveness and engagement in the standardized survey interview. *Sociological Methodology*, *46*(1), 1–38. https://doi.org/10.1177/0081175016637890

Gervais, R. O., Ben-Porath, Y. S., Wygant, D. B., & Green, P. (2007). Development and validation of a response bias scale (RBS) for the MMPI-2. *Assessment*, *14*(2), 196–208. https://doi.org/10.1177/1073191106295861

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. https://doi.org/10.1080/00223891.2011.594132

Grote, C. L., Butts, A. M., & Bodin, D. (2016). Education, training and practice of clinical neuropsychologists in the United States of America. *The Clinical Neuropsychologist*, *30*(8), 1356–1370. https://doi.org/10.1080/13854046.2016.1213885

Grove, W. M. (2005). Clinical versus statistical prediction: The contribution of Paul E. Meehl. *Journal of Clinical Psychology*, *61*(10), 1233–1243. https://doi.org/10.1002/jclp.20179

Hoge, C. W., & Warner, C. H. (2014). Estimating PTSD prevalence in US veterans: Considering combat exposure, PTSD checklist cutpoints, and DSM-5. *The Journal of Clinical Psychiatry*, *75*(12), e1439–e1441. https://doi.org/10.4088/JCP.14com09616

Ingram, P. B., Keen, M. A., Greene, T. E., Morris, C., & Armistead-Jehle, P. J. (2024). Development and initial validation of the Scale of Scales (SOS) overreporting scores for the MMPI family of instruments. *Journal of Clinical and Experimental Neuropsychology*, *46*(2), 95–110. https://doi.org/10.1080/13803395.2024.2320453

Ingram, P. B., Schmidt, A. T., Bergquist, B. K., & Currin, J. M. (2022). Coursework, instrument exposure, and perceived competence in psychological assessment: A national survey of practices and beliefs of health service psychology trainees. *Training and Education in Professional Psychology*, *16*(1), 10–19. https://doi.org/10.1037/tep0000348

Ingram, P. B., Tarescavage, A. M., Ben-Porath, Y. S., & Oehlert, M. E. (2020). Patterns of MMPI-2-Restructured Form (MMPI-2-RF) validity scale scores observed across veteran affairs settings. *Psychological Services*, *17*(3), 355–362. https://doi.org/10.1037/ser0000339

Ingram, P. B., & Ternes, M. S. (2016). The detection of content-based invalid responding: A meta-analysis of the MMPI-2-Restructured Form's (MMPI-2-RF) over-reporting validity scales. *The Clinical Neuropsychologist*, *30*(4), 473–496. https://doi.org/10.1080/13854046.2016.1187769

Inoue, C., Shawler, E., Jordan, C. H., & Jackson, C. A. (2022). Veteran and military mental health issues. *StatPearls*. StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK572092/

Kanser, R. J., Rohling, M. L., & Davis, J. J. (2024). Determining whether false positive rates increase with performance validity test battery expansion. *The Clinical Neuropsychologist*. Advance online publication. https://doi.org/10.1080/13854046.2024.2416543

Keen, M. A., Lee, T. T. C., Pscheid, K., & Forbey, J. D. (2023). Examination of the generalizability of underreporting detected by the MMPI-2-RF in a correctional sample. *Assessment*, *30*(4), 1157–1167. https://doi.org/10.1177/10731911221089036

Keller, S. M., & Tuerk, P. W. (2016). Evidence-based psychotherapy (EBP) non-initiation among veterans offered an EBP for posttraumatic stress disorder. *Psychological Services*, *13*(1), 42–48. https://doi.org/10.1037/ser0000064

Kelly, A. E., & Yuan, K.-H. (2009). Clients' secret keeping and the working alliance in adult outpatient therapy. *Psychotherapy: Theory, Research, Practice, Training*, *46*(2), 193–202. https://doi.org/10.1037/a0016084

Krishnamurthy, R., Hass, G. A., Natoli, A. P., Smith, B. L., Arbisi, P. A., & Gottfried, E. D. (2022). Professional practice guidelines for personality assessment. *Journal of Personality Assessment*, *104*(1), 1–16. https://doi.org/10.1080/00223891.2021.1942020

Lees-Haley, P. R., English, L. T., & Glenn, W. J. (1991). A Fake Bad Scale on the MMPI-2 for personal injury claimants. *Psychological Reports*, *68*(1), 203–210. https://doi.org/10.2466/pr0.1991.68.1.203

Liu, Y., Collins, C., Wang, K., Xie, X., & Bie, R. (2019). The prevalence and trend of depression among veterans in the United States. *Journal of Affective Disorders*, *245*, 724–727. https://doi.org/10.1016/j.jad.2018.11.031

Magruder, K. M., Frueh, B. C., Knapp, R. G., Davis, L., Hamner, M. B., Martin, R. H., Gold, P. B., & Arana, G. W. (2005). Prevalence of posttraumatic stress disorder in veterans affairs primary care clinics. *General Hospital Psychiatry*, *27*(3), 169–179. https://doi.org/10.1016/j.genhosppsych.2004.11.001

Morris, N. M., Mattera, J., Golden, B., Moses, S., & Ingram, P. B. (2022). Evaluating the performance of the MMPI-3 over-reporting scales: Sophisticated simulators and the effects of comorbid conditions. *The Clinical Neuropsychologist*, *36*(8), 2361–2369. https://doi.org/10.1080/13854046.2021.1968037

Neale, M. S., & Rosenheck, R. A. (1995). Therapeutic alliance and outcome in a VA intensive case management program. *Psychiatric Services*, *46*(7), 719–721. https://doi.org/10.1176/ps.46.7.719

Nelson, N. W., Hoelzle, J. B., McGuire, K. A., Ferrier-Auerbach, A. G., Charlesworth, M. J., & Sponheim, S. R. (2010). Evaluation context impacts neuropsychological performance of OEF/OIF veterans with reported combat-related concussion. *Archives of Clinical Neuropsychology*, *25*(8), 713–723. https://doi.org/10.1093/arclin/acq075

Nichter, B., Stein, M. B., Norman, S. B., Hill, M. L., Straus, E., Haller, M., & Pietrzak, R. H. (2021). Prevalence, correlates, and treatment of suicidal behavior in US military veterans: Results from the 2019–2020 National Health and Resilience in veterans Study. *The Journal of Clinical Psychiatry*, *82*(5), Article 35870. https://doi.org/10.4088/JCP.20m13714

Ray, C. L. (2014). Feigning screeners in VA PTSD compensation and pension examinations. *Psychological Injury and Law*, *7*(4), 370–387. https://doi.org/10.1007/s12207-014-9210-2

Ray, C. L. (2017). Practical use of MMPI-2-RF validity indicators in VA compensation and pension examinations. *Psychological Injury and Law*, *10*(3), 223–233. https://doi.org/10.1007/s12207-017-9289-3

Rogers, R., & Bender, S. D. (Eds.). (2018). *Clinical assessment of malingering and deception* (4th ed.). Guilford Press.

Russo, A. C. (2014). Assessing veteran symptom validity. *Psychological Injury and Law*, *7*(2), 178–190. https://doi.org/10.1007/s12207-014-9190-2

Russo, A. C. (2018). A practitioner survey of department of veterans Affairs psychologists who provide neuropsychological assessments. *Archives of Clinical Neuropsychology*, *33*(8), 1046–1059. https://doi.org/10.1093/arclin/acx139

Schroeder, R. W., Baade, L. E., Peck, C. P., VonDran, E. J., Brockman, C. J., Webster, B. K., & Heinrichs, R. J. (2012). Validation of MMPI-2-RF validity scales in criterion group neuropsychological samples. *The Clinical Neuropsychologist*, *26*(1), 129–146. https://doi.org/10.1080/13854046.2011.639314

Seal, K. H., Metzler, T. J., Gima, K. S., Bertenthal, D., Maguen, S., & Marmar, C. R. (2009). Trends and risk factors for mental health diagnoses among Iraq and Afghanistan veterans using Department of Veterans Affairs health care, 2002–2008. *American Journal of Public Health*, *99*(9), 1651–1658. https://doi.org/10.2105/AJPH.2008.150284

Sharf, A. J., Rogers, R., Williams, M. M., & Henry, S. A. (2017). The effectiveness of the MMPI-2-RF in detecting feigned mental disorders and cognitive deficits: A meta-analysis. *Journal of Psychopathology and Behavioral Assessment*, *39*(3), 441–455. https://doi.org/10.1007/s10862-017-9590-1

Shura, R. D., Yoash-Gantz, R. E., Pickett, T. C., McDonald, S. D., & Tupler, L. A. (2021). Relations among performance and symptom validity, mild traumatic brain injury, and posttraumatic stress disorder symptom burden in postdeployment veterans. *Psychological Injury and Law*, *14*(4), 257–268. https://doi.org/10.1007/s12207-021-09415-z

Sullivan, K. A., Elliott, C. D., Lange, R. T., & Anderson, D. S. (2013). A known-groups evaluation of the response bias scale in a neuropsychological setting. *Applied Neuropsychology: Adult*, *20*(1), 20–32. https://doi.org/10.1080/09084282.2012.670149

Sweet, J. J., Heilbronner, R. L., Morgan, J. E., Larrabee, G. J., Rohling, M. L., Boone, K. B., Kirkwood, M. W., Schroeder, R. W., Suhr, J. A., & the Conference Participants. (2021). American Academy of Clinical Neuropsychology (AACN) 2021 consensus statement on validity assessment: Update of the 2009 AACN consensus conference statement on neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist*, *35*(6), 1053–1106. https://doi.org/10.1080/13854046.2021.1896036

Tanielian, T., Woldetsadik, M. A., Jaycox, L. H., Batka, C., Moen, S., Farmer, C., & Engel, C. C. (2016). Barriers to engaging service members in mental health care within the US military health system. *Psychiatric Services*, *67*(7), 718–727. https://doi.org/10.1176/appi.ps.201500237

Tellegen, A., & Ben-Porath, Y. S. (2012). *Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF): Technical manual*. University of Minnesota Press. (Original work published 2008)

US Bureau of Labor Statistics. (2023, March 21). *Employment situation of veterans news release* [Press release]. https://www.bls.gov/news.release/vet.htm

US Department of Veterans Affairs. (2023). *US department of veteran affairs FY 2024 budget submission*. https://www.va.gov/budget/docs/summary/fy2024-va-budget-in-brief.pdf

Wandler, H. A. (2013). The role of culture in advocating for accurate diagnosis and rating of veterans' psychological disabilities. *Mental Health Law & Policy Journal*, *2*(1), 1–42. https://doi.org/10.2139/ssrn.2133518

Whiteside, D. M., & Basso, M. R. (2024). Innovations in performance and symptom validity testing: Introduction to symptom validity section of the special issue. *Journal of Clinical and Experimental Neuropsychology*, *46*(2), 81–85. https://doi.org/10.1080/13803395.2024.2346022

Wisco, B. E., Marx, B. P., Wolf, E. J., Miller, M. W., Southwick, S. M., & Pietrzak, R. H. (2014). Posttraumatic stress disorder in the US veteran population: Results from the National Health and Resilience in Veterans Study. *The Journal of Clinical Psychiatry*, *75*(12), 1338–1346. https://doi.org/10.4088/JCP.14m09328

Worthen, M. D., & Moering, R. G. (2011). A practical guide to conducting VA compensation and pension exams for PTSD and other mental disorders. *Psychological Injury and Law*, *4*(3–4), 187–216. https://doi.org/10.1007/s12207-011-9115-2

Wygant, D. B., Ben-Porath, Y. S., Arbisi, P. A., Berry, D. T., Freeman, D. B., & Heilbronner, R. L. (2009). Examination of the MMPI-2 restructured form (MMPI-2-RF) validity scales in civil forensic settings: Findings from simulation and known group samples. *Archives of Clinical Neuropsychology*, *24*(7), 671–680. https://doi.org/10.1093/arclin/acp073

Received June 24, 2024
Revision received October 25, 2024
Accepted December 15, 2024 ■