

Word Memory Test (WMT) performance and MMPI-2 validity scales and indices

Roger Gervais, Ph.D., Neurobehavioural Associates, Edmonton, Alberta
Dustin B. Wygant, M.A., Department of Psychology, Kent State University
Yossef S. Ben-Porath, Ph.D., Department of Psychology, Kent State University

2005 Annual Meeting of the National Academy of Neuropsychology, Tampa, FL., October 19-22, 2005.

Abstract

Objective: Numerous validity scales and indices for detecting biased responding and feigned/exaggerated symptoms have been developed for the MMPI/MMPI-2. The present study examined the relationship between performance on the Word Memory Test (WMT; Green, 2003; Green, Allen & Astner, 1996; Green & Astner, 1995), an objective measure of cognitive response bias, and scores on the MMPI-2 F scales, the Infrequent Somatic Complaints (F_s) scale (Wygant, 2004), F-K, the Meyers Validity Index (Meyers et al., 2002), Md scale (Steffan et al., 2003) and other indices designed to detect symptom exaggeration/biased responding.

Method: This archival study used MMPI-2 and WMT data from 1212 consecutive non-head injury disability-related referrals to the first author's private practice. Chronic pain (39%) and anxiety-related disorders (23%) were the primary diagnoses. We excluded patients who did not complete the WMT, omitted 30 or more items, or obtained a T score of 80 or greater on VRIN/TRIN. This resulted in a final sample of 775 cases (51% male). The sample was divided into WMT pass (70%) or fail (30%) groups.

Results: The fail WMT group scored significantly higher than the pass WMT group on all MMPI-2 scales and indices ($p = 0.01$ to 0.0005). ANOVA results indicated significant increases in MMPI-2 scores across decreasing WMT score ranges ($p < 0.0005$).

Conclusions: Poor performance on the WMT was associated with increasing scores on MMPI-2 validity scales and indices. This provides external validation and support for interpretation of the MMPI-2 validity indices and scales as indicators of biased responding/symptom exaggeration. Sensitivity and specificity of these scales relative to the WMT is discussed.

Materials and Methods

Participants

This archival study examined the relation between scores on various MMPI-2 validity scales and indices of symptom exaggeration or response bias and the Word Memory Test (WMT; Green, 2003; Green, Allen & Astner, 1996; Green & Astner, 1995), a well validated cognitive symptom validity test. The data was drawn from a database of 1212 consecutive non-head injury disability-related referrals to the first author's private practice. Excluding persons who omitted 30 or more items on the MMPI-2 ($n = 273$), or who obtained TRIN or VRIN T-scores greater than or equal to 80 ($n = 31$) resulted in a final sample of 775 cases. The sample was 40.3 (10.6) years of age, 51% male, and had 12.1 (2.3) years of education. Chronic pain (39%) and anxiety-related disorders (23%) were the primary diagnoses, followed by orthopedic injuries (17%) with no primary psychological diagnosis, and depression (15%). Most clients were receiving or in the process of applying for WBC or medical disability benefits.

Assessment Methods

Depending upon the type of assessment, the participants completed a variety psychological and neuropsychological tests and self-report questionnaires and inventories including the WMT and the full 567-item MMPI-2. Cutoffs for determining failure on the WMT were set in accordance with the test manual (Green, 2003). The raw MMPI-2 data collected by the first author were scored by the second author to provide the standard MMPI-2 validity and clinical scales, as well as F_p (Arbisi & Ben-Porath, 1995), the Infrequent Somatic Complaints (F_s) scale (Wygant, 2004), the F-K index (Gough, 1950), Gough's Dissimulation-revised ($Ds-r$; Gough, 1957), the Fake Bad Scale (FBS, Lees-Haley, English, & Glenn, 1991), Meyers Validity Index (MVI, Meyers, Millis, & Volkert, 2002), and the Malingered Depression (Md, Steffan, Clopton, & Morgan, 2003).

Statistical Analyses

The first level of analysis used T-tests to contrast the means of the MMPI-2 scales and indices between the pass and fail WMT groups. We then used ANOVAs to contrast the mean differences in MMPI-2 scales across three levels of WMT performance: 1. Mean WMT $\geq 90\%$, 2. Mean WMT = 80-89%, and 3. Mean WMT $< 80\%$. WMT failure rate in the second group was 70%. In the final analysis, we calculated sensitivity, specificity, positive predictive power, and hit rate of the MMPI-2 scales in predicting WMT failure.

Results

T-tests indicated that the fail WMT group scored significantly higher than the pass WMT group on all MMPI-2 scales and indices ($p < 0.01$ to 0.0005). Effect sizes (Cohen's d) were in the medium range (.27 - .52). The results of these analyses are reported in Table 1.

Similarly, in the second analysis using ANOVA, there were significant increases in all MMPI-2 scales across WMT levels ($p < 0.0005$). Post hoc pairwise multiple comparisons using Tukey's HSD as detailed in Table 2 indicate the significant comparisons.

Setting cutoffs for invalid or biased responding as recommended for the various MMPI-2 validity scales and indices, we calculated the sensitivity and specificity of these scales in predicting WMT failure. Sensitivity was poor to moderate for all scales, with F_p demonstrating the lowest sensitivity at .01 and the FBS obtaining the highest at .64. Specificity was moderate to high for all scales, ranging from .57 for FBS to .99 for F_p . Using these cutoffs, we calculated positive predictive power (PPP) and negative predictive power (NPP) for each of the MMPI-2 validity scales and indices in predicting WMT failure. The results were mediocre, with PPP ranging from .33 to .48. NPP was stronger, ranging from .70 to .79. Hit rates ranged from 61% for Md to 69% for F_B and F_p . The results of these analyses and the cutoffs employed are presented in Table 3.

Discussion

The present study demonstrated that poor performance on the WMT, an objective measure of response bias and effort, was associated with increasing scores on MMPI-2 validity scales and indices. This pattern was also consistently seen across three levels of WMT performance. The subgroup demonstrating the poorest WMT performance also obtained the highest scores on the MMPI-2 scales. Examining the sensitivities of various MMPI-2 validity indices on Table 3, it appears that scales measuring aspects of dissimulation other than severe psychopathology may be more sensitive to poor performance on the WMT. For instance, the FBS and F_S potentially measure co-occurring somatic exaggeration and depression (Md) may be more sensitive to feigned cognitive impairment than overt psychopathology (e.g., F_p). This coincides with Larrabee's (2003) argument that individuals feigning cognitive impairment "act hurt" rather than "acting crazy." This provides external validation for interpretation of the MMPI-2 validity indices and scales as indicators of biased responding/symptom exaggeration. This interpretation does not rule out, however, that psychopathology could be contributing to WMT failure and MMPI-2 validity scale elevation, as these analyses are essentially correlational in nature and cannot definitively conclude causality.

Although WMT failure was associated with increasing scores on the MMPI-2 validity scales and indices, the converse was not observed. Using commonly accepted cutoffs for biased or otherwise invalid responding, the MMPI-2 scales demonstrated poor PPP in predicting WMT failure. Considering that the MMPI-2 validity scales and indices were designed to detect deviant responding in regard to psychopathology, it is not surprising that they might be generally insensitive to a measure designed to detect predominantly cognitive response bias. These findings suggest that further research into development of MMPI-2 scales to evaluate cognitive response bias, as measured by SVTs such as the WMT, would be a worthwhile endeavor (see Gervais, 2005).

References

- Arbisi, P. & Ben-Porath, Y. (1995). An MMPI-2 infrequency response scale for use with psychopathological populations: the Infrequency Psychopathology Scale, $F(p)$. *Psychological Assessment*, 7, 424-431.
- Gervais, R. O. (2005). Development of a response bias scale for the MMPI-2. Presented at Annual conference of National Academy of Neuropsychology, Tampa, Florida, October 2005. *Archives of Clinical Neuropsychology*, 20 (7), 891.
- Gough, H. G. (1950). The F minus K dissimulation index for the MMPI. *Journal of Consulting Psychology*, 14, 408-413.
- Gough, H. G. (1957). *California psychological inventory manual*. Palo Alto, CA: Consulting Psychologists Press.
- Graham, J. R. (2006). *MMPI-2: Assessing personality and Psychopathology*, 4th edition. New York: Oxford University Press.
- Green, P. (2003). *Green's Word Memory Test for Windows: User's manual*. Edmonton: Green's Publishing.
- Green, P., Allen, L., & Astner, K. (1996). *The Word Memory Test: A User's Guide to the Oral and Computer-Administered Forms, US Version 1.1*. Durham, NC: CogniSyst, Inc.
- Green, P. & Astner K. (1995). *Manual for the Oral Word Memory Test*. Durham, NC: CogniSyst, Inc.
- Larabee, G.J. (2003). Exaggerated MMPI-2 symptoms report in personal injury litigants with malingered neurocognitive deficit. *Archives of Clinical Neuropsychology*, 18, 673-686.
- 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, 203-210.
- Meyers, J. E., Millis, S. R., & Volkert, K. (2002). A validity index for the MMPI-2. *Archives of Clinical Neuropsychology*, 17, 157-169.
- Rogers, R. (1997). *Clinical assessment of malingering and deception*, 2nd edition. New York: The Guilford Press.
- Steffan, J., Clopton, J., & Morgan, R. (2003). An MMPI-2 scale to detect malingered depression (Md Scale). *Assessment*, 10, 383-392.
- Wygant, D. B. (2004). Development and initial validation of a scale to detect infrequent somatic complaints. Unpublished master's thesis, Kent State University, Kent, Ohio.

Table 1
MMPI-2 Validity Scales and Indices in Pass and Fail WMT groups

	WMT Performance Level				<i>t</i> test		Effect Size Cohen's <i>d</i>
	Pass WMT: (<i>n</i> = 543)		Fail WMT: (<i>n</i> = 232)				
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
F	61.9	17.1	67.3	18.8	3.8	< .0005	.31
F _B	64.9	20.0	73.1	23.6	4.6	< .0005	.38
F _P	53.4	12.3	57.6	13.9	3.9	< .0005	.33
F _S	66.8	19.5	75.3	22.8	4.6	< .0005	.41
F - K	-6.2	9.0	-3.7	10.1	3.1	.002	.27
FBS	23.2	5.6	26.1	5.6	6.5	< .0005	.52
MVI	2.70	3.0	4.2	3.6	5.5	< .0005	.47
Ds-r	61.7	15.1	66.0	16.9	3.3	.001	.27
Md	12.8	5.6	15.5	5.6	6.1	< .0005	.48

Note. Pass and Fail WMT as defined in Green (2003). All MMPI-2 scores are T scores except F - K, FBS, MVI, and Md, which are raw scores.

Table 2
ANOVA Results Word Memory Test (WMT) and MMPI-2 Validity Scales and Indices

	WMT Performance Level						ANOVA		Effect Size
	Level 1: > 90% (<i>n</i> = 507)		Level 2: 80-89% (<i>n</i> = 129)		Level 3: < 80% (<i>n</i> = 139)		F (2, 272)	<i>p</i>	η^2
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
F	61.6 _a	17.0	66.0 _b	19.2	68.1 _b	18.3	8.9	< .0005	.023
F _B	64.5 _a	19.9	69.1 _a	22.7	76.0 _b	23.4	16.9	< .0005	.042
F _P	53.2 _a	12.0	56.2 _{ab}	14.5	58.7 _b	13.7	11.3	< .0005	.029
F _S	66.3 _a	19.4	71.8 _b	21.3	78.4 _c	23.0	20.3	< .0005	.050
F - K	-6.3 _a	8.9	-4.2 _{ab}	9.7	-3.2 _b	10.2	7.4	.001	.019
FBS	23.1 _a	5.5	24.5 _b	5.6	27.4 _c	5.4	33.7	< .0005	.080
MVI	2.61 _a	2.9	3.6 _b	3.7	4.6 _c	3.5	23.7	< .0005	.058
Ds-r	61.4 _a	14.9	65.5 _b	17.3	66.9 _b	16.5	8.7	< .0005	.022
Md	12.7 _a	5.5	14.2 _b	5.5	16.3 _c	5.8	23.9	< .0005	.058

Note. Means with different subtext are significantly different (Tukey HSD). WMT performance is mean Immediate Recall, Delayed Recall, and Consistency.

Table 3. Sensitivity, Specificity, Positive Predictive Power, Negative Predictive Power and Hit Rate of MMPI-2 validity scales and indices in predicting WMT failure in sample ($N = 775$).

MMPI-2 Scale	Cutoff	Sensitivity	Specificity	PPP	NPP	Hit Rate
F	15	.17	.88	.38	.71	.67
F _B	90	.27	.87	.48	.74	.69
F _P	8	.01	.99	.33	.70	.69
F _S	80	.41	.75	.40	.74	.64
F-K	6	.19	.89	.42	.72	.68
FBS	25	.64	.58	.39	.79	.60
MVI	5	.38	.80	.45	.75	.67
Md	15	.58	.62	.40	.78	.61
Ds-r	75	.33	.79	.40	.73	.65

Note. Cutoffs for F, F-K, and Ds set at 75th percentile as per Table 9.17 in Rogers (1997). Cutoff for F_B and F_P as per Graham (2006); F_S (Wygant, 2004), FBS and Ds-r as per Meyers et al., (2002); and Md (Steffan et al. 2003). Cutoffs are greater or equal to listed scores. All scores are raw except for F_B, F_S, and Ds-r, which are T-scores.

Contact information:

Roger Gervais, Ph.D.
 Suite 201, 17107-107 Avenue, Edmonton Alberta, Canada
 T5S 1G3
 E-mail: rgervais@shaw.ca