EXAMINING THE MMPI-2 INFREQUENCY SCALES: CLINICAL AND NORMATIVE ISSUES

Helen Buzas and Dr. Graeme Senior
Department of Psychology, University of Southern Queensland

INTRODUCTION

The MMPI-2 (Minnesota Multiphasic Personality Inventory-2) is the most commonly used and studied personality inventory in the world (Greene, 1991), and second most widely used of all neuropsychological tests in forensic practice (Lees-Haley, 1991). This self-report inventory is designed to assess a wide of variety of personality constructs as they relate to psychopathology and is used in many clinical contexts. A critical role of the MMPI-2 in the forensic setting is the examination of motivation and effort on the part of the respondent. In the medicolegal assessment of disability this commonly relates to the evaluation of overreporting or exaggeration of physical/psychological symptoms. Specifically, MMPI-2 validity scales such as F (Infrequency), Fb (Infrequency-back), Fp (Infrequency-psychopathology), and FBS (Fake-Bad Scale) are used to examine the validity of the protocol. Individuals who elevate these scales above a criterion level are considered to be exaggerating their symptomatology, and the protocol is thus rendered uninterpretable or invalid.

The F scale was developed by examining items in the test which had low endorsement rates in a normative group (approximately less than 10% of the normative sample endorsed these items in the deviant direction). The F scale is composed of these low frequency items from the first 361 items of the MMPI-2. The items were selected to cover a wide range of content so as not to elevate in individuals with particular disorders. The Fb scale, newly developed for the MMPI-2, was also constructed with items that have a low frequency of endorsement and are derived from the back half of the test (item 281 onwards). A person who endorses a large number of F or Fb scale items in the deviant direction is endorsing a number of items seldom reported by normal individuals. While this is expected in individuals who have a psychopathology, extreme elevations (T>109) of F or Fb have come to be interpreted as overreporting of symptoms.

As indicated, one of the conceptual problems is that individuals who are accurately reporting high levels of distress or disturbance are also expected to elevate F and Fb. Recently, two new scales, Fp and FBS, have been introduced that were designed explicitly for the purpose of assessing symptom exaggeration. The Fp scale was designed to detect ‘infrequency’ in settings where a high F and Fb endorsement rate is known to exist (ie, acute psychiatric inpatient settings). Fp was constructed using items that have a low base rate of endorsement (<20%) in normals and in traditionally high-F populations. The role intended for Fp is to differentiate between high F scores indicative of exaggeration and those reflecting extreme levels of distress or disturbance. The rationale is that if a client endorses an extreme number of F items but an average number of Fp items, the elevated F probably reflects a legitimate reporting of severe symptoms. If, however, F is extreme and Fp is elevated, the client has endorsed an extreme number of infrequent items even for an acute psychiatric patient, and therefore, likely to be exaggerating their symptomatology. The 27 items that constitute the Fp scale can be seen in Table 1 along with their endorsement direction.

The FBS (Fake Bad Scale) for the MMPI-2 was developed by Lees-Haley, English, & Glenn (1991) to serve the same role as the revised Gough Dissimulation Scale (Dr-R) on the MMPI. FBS item content was chosen rationally in that items were selected by Lees-Haley based upon endorsement frequencies of malingerers on

the MMPI. Items were selected with content that is described as a mixture of fake good and fake bad responding. This is explained by the assumption that a malingerer will attempt to exaggerate their problems in identified areas and attempt to appear extremely honest on other content. The 43 items on this scale with their endorsement directions are shown in the FBS results section.

It is unclear exactly when 'infrequency' became a measure of symptom exaggeration. However, it is now standard MMPI-2 interpretation procedure to reject protocols with elevated F, Fb and/or F(p). In recent years concern has grown regarding the assumptions underlying this practice and the validity of the F scales in detection of symptom exaggeration. This study examined four scales, F, Fb, Fp, and FBS, purported to assess overreporting or symptom exaggeration. A database of normal and clinical case MMPI-2 protocols was utilised to examine the base rates of these scales in a variety of diagnostic groups. The degree to which elevations on these scales could be predicted by combinations of Basic or Content scales was also evaluated to address the likelihood that these validity scales are able to measure constructs independent of clinically-related elevations.

METHOD

Two data sets were used in this analysis. The first sample consisted of the MMPI-2 protocols of 221 adults who had elevated F and Fb. The sample consisted of Normals (N=51), Traumatic Brain Injury patients (N=101), Chronic Pain patients (N=53), and Posttraumatic Stress Disorder patients (N=15). The data in this sample included the endorsements for each subject on the 106 items that are used in the F, Fb, and Fp scales. The intention was to conduct a factor analysis on this data set and extract factors that corresponded to meaningful groups of items endorsed by individuals with elevated F scales.

The second sample consisted of the MMPI-2 protocols of 1,144 individuals, 197 or which were part of a normative study and the remaining 947 derived from forensic psychological/psychiatric practices in Brisbane, Queensland. These were utilized to examine the relationship between clinical and content scales and the validity scales (F, Fb, Fp and FBS) and establish base rates for different clinical diagnostic groups.

RESULTS

A Principal Components analysis with varimax rotation using SPSS was performed on the first sample in order to examine the factor structure of the F, Fb, and Fp item pool. Orthoganol rotation was chosen to maximise the likelihood of yielding factors that could be used to generate independent subscales. Attempts at factor analysing the content of MMPI-2 items on the F, Fb, and Fp scales were less than successful. Applying the criterion of "eigenvalues greater than 1" yielded 36 factors, but the constructs underlying the factors were not readily discernible from item content. Examination of the scree plot indicated a smaller and more manageable number of factors (seven). However, here too the underlying constructs were not able to be determined by examination of item content. For example, different items that relate to suicide and suicidal ideation loaded on different factors. It should be noted that mean endorsement rates for items on each factor were similar and were under 20% (with the exception of one factor at 28.5%). This finding is not surprising given the item selection criterion for F and Fb of endorsements in the standardisation sample of less than 20%.

The second sample was then evaluated for validity scale base rates and the influence of Basic and Content scales upon scale performance. In examining these issues, each scale will be considered in turn. Of the 1,142 cases examined, 20 protocols were classified as suspected malingerers based upon clinical impression and supported by poor performance on cognitive tests of malingering. As normal individuals and suspected

malingerers represent, at least theoretically, extreme positions on the exaggeration continuum the data for these two samples has been bolded in the relevant tables to contrast with the clinical groups. In each case multiple regression was conducted on each validity scale using Basic scales and Content scales as predictor variables. In the case of F only Basic scales were employed to be consistent with the item pool which is drawn from the first 370 items only. Similarly, Fb is drawn from the back half of the test which has the greater proportion of Content scale items. For this reason only Content scales were regressed onto Fb. As both FBS and Fp are drawn from the whole test item pool both Basic and Content scales were used in their regressions. In each case where a substantial prediction was achieved a follow-up regression was computed substituting subscales for parent scales where applicable. This was done to achieve a more detailed understanding of the relationships between clinical measures and validity scales.

F Scale

Base Rates of F in a Variety of Clinical Conditions

	Percentage of Cases with Elevations in F			
	Unelevated	Elevated	Markedly Elevated	Extremely Elevated
Group	T <65	T = 65-90	T = 91-109	T = 110+
Adjustment Disorder (N=78)	64	31	3	3
Alcoholism (N = 47)	49	28	11	13
Anxiety (N = 32)	53	41	3	3
Chronic Fatigue Syndrome (N = 6)	67	33	0	0
Chronic Pain (N = 217)	65	29	5	1
Depression (N = 100)	49	36	11	4
Dissociative Disorder (N = 10)	0	50	30	20
Drug Abuse (N = 59)	47	29	15	8
Normal (N = 197)	**72**	**26**	**2**	**0**
Phobia (N = 20)	65	30	0	5
PTSD (N = 91)	55	27	15	2
Somatoform Disorder (N = 15)	60	33	7	0
Suspected Malingerers (N = 20)	**80**	**5**	**0**	**15**
TBI (N = 252)	58	32	6	5
Average (N = 1144)	56	31	8	6

Multiple Regression of Clinical Scales onto F Scale (N = 1,144)

Model	R	R^2	Adj. R^2	SEe
Sc	0.79	0.63	0.63	11.37
Sc, Hy	0.82	0.67	0.67	10.79

Sc, Hy, Pa	0.83	0.70	0.70	10.31
Sc, Hy, Pa, Si	0.84	0.71	0.71	10.10
Sc, Hy, Pa, Si, Pt	0.85	0.72	0.72	9.90
Sc, Hy, Pa, Si, Pt, Ma	0.85	0.73	0.73	9.75
Sc, Hy, Pa, Si, Pt, Ma, Pd	0.86	0.73	0.73	9.68
Sc, Hy, Pa, Si, Pt, Ma, Pd, Hs	0.86	0.74	0.73	9.66

F = Sc*0.71 - Hy*0.13 + Pa*0.27 + Si*0.40 - Pt*0.29 + Ma*0.20 - 6.04

Multiple Regression of Clinical Subscales onto F Scale (N = 1,144)

Model	R	R^2	Adj. R^2	SEe
SC1	0.80	0.63	0.63	11.31
SC1, SC2	0.83	0.70	0.69	10.34
SC1, SC2, PA1	0.85	0.73	0.73	9.81
SC1, SC2, PA1, SC6	0.86	0.75	0.75	9.41
SC1, SC2, PA1, SC6, SC5	0.87	0.75	0.75	9.32
SC1, SC2, PA1, SC6, SC5, SC3	0.87	0.76	0.76	9.25
SC1, SC2, PA1, SC6, SC5, SC3, HY3	0.87	0.76	0.76	9.16
SC1, SC2, PA1, SC6, SC5, SC3, HY3, SI2	0.87	0.77	0.76	9.10

F = Sc1*0.36 + Sc2*0.19 + Pa1*0.28 + Sc6*0.19 - 0.24

Fb Scale

Base Rates of Fb in a Variety of Clinical Conditions

	Percentage of Cases with Elevations in Fb			
	Unelevated	Elevated	Markedly Elevated	Extremely Elevated
Group	T <65	T = 65-90	T = 91-109	T = 110+
Adjustment Disorder (N=78)	60	19	15	5
Alcoholism (N = 47)	34	32	6	28
Anxiety (N = 32)	31	53	9	6
Chronic Fatigue Synd. (N = 6)	67	17	17	0
Chronic Pain (N = 217)	59	27	11	4
Depression (N = 100)	37	42	11	10

Dissociative Disorder (N = 10)	0	40	20	40
Drug Abuse (N = 59)	34	34	12	20
Normal (N = 197)	**85**	**11**	**3**	**1**
Phobia (N = 20)	65	20	5	10
PTSD (N = 91)	35	31	23	11
Somatoform Disorder (N = 15)	73	27	0	0
Suspected Malingerers (N = 20)	**75**	**10**	**5**	**10**
TBI (N = 252)	57	27	8	8
Average (N = 1144)	51	28	10	11

Multiple Regression of Content Scales onto Fb Scale (N = 1,144)

Model	R	R^2	Adj. R^2	SEe
DEP	0.83	0.69	0.69	12.80
DEP, BIZ	0.87	0.75	0.75	11.51
DEP, BIZ, TRT	0.88	0.77	0.77	11.04
DEP, BIZ, TRT, FRS	0.89	0.78	0.78	10.69
DEP, BIZ, TRT, FRS, OBS	0.89	0.79	0.79	10.64
DEP, BIZ, TRT, FRS, OBS, LSE	0.89	0.79	0.79	10.55
DEP, BIZ, TRT, FRS, OBS, LSE, HEA	0.89	0.79	0.79	10.50
DEP, BIZ, TRT, FRS, OBS, LSE, HEA, FAM	0.89	0.79	0.79	10.45

Fb = Dep*0.62 + Biz*0.41 + Trt*0.45 + Frs*0.27 - Obs*0.15 - 30.89

Multiple Regression of Content Component Scales onto Fb Scale (N = 1,144)

Model	R	R^2	Adj. R^2	SEe
DEP1	0.79	0.63	0.63	14.06
DEP1, FRS1	0.86	0.75	0.75	11.59
DEP1, FRS1, DEP4	0.90	0.81	0.81	10.02
DEP1, FRS1, DEP4, BIZ2	0.92	0.85	0.85	8.99
DEP1, FRS1, DEP4, BIZ2, TRT1	0.93	0.86	0.86	8.66
DEP1, FRS1, DEP4, BIZ2, TRT1, LSE1	0.93	0.86	0.86	8.54
DEP1, FRS1, DEP4, BIZ2, TRT1, LSE1, FRS2	0.93	0.87	0.86	8.48

Fb = DEP1*0.16 + FRS1*0.36 + DEP4*0.30 + BIZ2*0.38 + TRT1*0.26 - 20.60

Comparisons Between F and Fb

In evaluating protocol validity, comparisons are commonly made between T scores for F and Fb. A number of heuristics have been suggested for evaluating F and Fb, such as: if F is ≤90 and Fb>90 then do not interpret Content or Supplementary scales, L, F, K and Basic scales alone should be interpreted. Table 3 indicates the percentage of individuals in each group that demonstrated differences between F and Fb in a particular range. Of particular note is that the distribution is not symmetrical and favours a greater likelihood that Fb will exceed F in its elevation, again consistent with the hypothesis that Fb is influenced to a greater degree by distress related responses.

Distribution of Differences Between F and Fb Scales in a Variety of Clinical Groups

	Percent of Differences Between for F and Fb (T Scores)										
Group	<-30	-21 to -30	-11 to -20	-6 to -10	-1 to -5	0	1 to 5	6 to 10	11 to 20	21 to 30	>30
Adjustment Disorder (N=78)	5.1	7.7	17.9	11.5	20.5	5.1	15.4	7.7	9.0	0.0	0.0
Alcoholism (N = 47)	8.5	8.5	23.4	14.9	10.6	12.8	10.6	6.4	2.1	2.1	0.0
Anxiety (N = 32)	0.0	18.8	12.5	15.6	15.6	6.3	18.8	3.1	9.4	0.0	0.0
Chronic Fatigue Synd. (N = 6)	0.0	16.7	0.0	16.7	0.0	0.0	16.7	16.7	33.3	0.0	0.0
Chronic Pain (N = 217)	4.1	9.2	13.4	12.0	20.3	3.2	14.7	11.1	10.6	1.4	0.0
Depression (N = 100)	5.0	10.0	12.0	19.0	15.0	4.0	15.0	11.0	8.0	1.0	0.0
Dissociative Disorder (N = 10)	0.0	10.0	10.0	20.0	20.0	20.0	0.0	20.0	0.0	0.0	0.0
Drug Abuse (N = 59)	10.2	8.5	22.0	8.5	18.6	5.1	5.1	10.2	10.2	1.7	0.0
Normal (N = 197)	**0.5**	**2.0**	**4.1**	**9.1**	**13.7**	**1.0**	**27.4**	**17.8**	**17.8**	**5.6**	**1.0**
Phobia (N = 20)	10.0	0.0	30.0	10.0	15.0	0.0	15.0	5.0	15.0	0.0	0.0
PTSD (N = 91)	13.2	4.4	20.9	15.4	14.3	5.5	6.6	9.9	5.5	3.3	1.1
Somatoform Disorder (N = 15)	0.0	6.7	0.0	13.3	26.7	0.0	13.3	26.7	13.3	0.0	0.0
Suspected Malinger. (N = 20)	**0.0**	**0.0**	**20.0**	**40.0**	**20.0**	**0.0**	**5.0**	**15.0**	**0.0**	**0.0**	**0.0**
TBI (N = 252)	2.4	3.6	18.7	11.9	18.3	3.6	13.9	11.1	13.9	2.4	0.4
Mean	26.4			31.9		4.8	24.9		12.0		

Fp

Multiple Regression of Basic and Content Scales onto Fp

Model	R	R^2	Ad.R^2	SEe
BIZ	.513	.263	.263	12.78
BIZ, SC	.557	.311	.309	12.37
BIZ, SC, PT	.582	.339	.337	12.12

BIZ, SC, PT, CYN	.596	.355	.353	11.97
BIZ, SC, PT, CYN, OBS	.601	.361	.358	11.92
BIZ, SC, PT, CYN, OBS, LSE	.610	.372	.369	11.82
BIZ, SC, PT, CYN, OBS, LSE, FRS	.614	.377	.374	11.78
BIZ, SC, PT, CYN, OBS, LSE, FRS, PA	.618	.382	.378	11.74
BIZ, SC, PT, CYN, OBS, LSE, FRS, PA, ANX	.624	.389	.384	11.68

FBS

Multiple Regression of Basic and Content Scales onto FBS (N = 1,144)

Model	R	R^2	Adj.R^2	SEe
HS	.844	.713	.712	3.80
HS, D	.881	.776	.775	3.36
HS, D, ASP	.898	.806	.805	3.13
HS, D, ASP, ANX	.926	.857	.857	2.68
HS, D, ASP, ANX, HEA	.929	.863	.863	2.63
HS, D, ASP, ANX, HEA, WRK	.931	.866	.866	2.60
HS, D, ASP, ANX, HEA, WRK, HY	.933	.871	.870	2.55
HS, D, ASP, ANX, HEA, WRK, HY, MF	.934	.873	.872	2.53

FBS = Hs*0.187 + D*0.026 - ASP*0.226 + ANX*0.123 + HEA*0.087 + WRK*0.059 – 0.164

Multiple Regression of Content Component Scales onto FBS Scale (N = 1,144)

Model	R	R^2	Adj.R^2	SEe
HS	.844	.713	.712	3.80
HS, ANX	.880	.775	.775	3.36
HS, ANX, ASP1	.925	.856	.855	2.70
HS, ANX, ASP1, WRK	.928	.862	.861	2.64
HS, ANX, ASP1, WRK, ASP2	.931	.866	.866	2.60
HS, ANX, ASP1, WRK, ASP2, HEA1	.932	.869	.868	2.58
HS, ANX, ASP1, WRK, ASP2, HEA1, D3	.933	.870	.869	2.56

FBS Raw = HS*0.259 + ANX*0.146 – ASP1*0.220 + WRK*0.079 – 0.415

DISCUSSION

That a factor analytic approach did not yield interpretable composites of the F, Fb, and Fp item pool is perhaps not surprising. In their construction, items were not chosen for these scales based upon their content but rather for their endorsement infrequency. However, the absence of any components that clearly related to meaningful symptom clusters or severity of symptoms does challenge the assumption that these scales relate to symptom severity.

This study has revealed two critical findings concerning assumptions regarding the roles of F, Fb, and FBS as indicators of exaggerated symptomatology.

1. F, Fb, and FBS can be predicted with a high degree of accuracy using a small subset of MMPI-2 subscales. This is, in part, not surprising given the item overlap of the validity scales with these subscales. However, it clearly indicates that a pattern of responding that reflects a specific pattern of clinical problems will result artifactually in an elevation in these validity scales.
2. Perhaps more compelling is the fact that different subscales are predictive of the different validity scales. F is primarily influenced by scales that would characterise disturbance and may be differentially elevated by neurologically related items on Sc6. Fb in turn is influenced by distress related content particularly relating to depression and anxiety. The FBS scale also seems to best predicted by anxiety related scales and subscales.

The low level of prediction for Fp is also somewhat expected in that the items on this scale were chosen for their low frequency of endorsement even in clinical populations. Examination of the specific items, however, fails to reveal content that can be readily associated either with severe symptomatology or exaggeration related behaviours. For example, it is difficult to understand how not loving your father or not thinking your mother is a good woman could have any other clinical significance other than most people would not be prepared to acknowledge this (even with psychopathology).

These findings challenge any hypothesis that these validity scales are evaluating the same construct let alone one of symptom exaggeration. In each case a pattern of clinical findings not uncommon to many clinical conditions could be easily seen to elevate one of these so-called validity scales. In particular the endorsement of any suicidal ideation or behaviour is virtually guaranteed to result in elevations in Fb. As this is a not uncommon concern in some clinical settings, it would be tragic to misinterpret an open "cry for help" in a suicidal client as "exaggeration of symptomatology" on the MMPI-2.

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.

Greene, R. (1991). The MMPI/MMPI-2. An Interpretative Manual. Boston: Allyn and Bacon.

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.

Wednesday, 8 March 2000

Senior Lecturer
Department of Psychology

University of Southern Queensland
Toowoomba, QLD 4350
Australia