Challenging the Validity of the MMPI-2 Infrequency Scales

by

Dr. Graeme Senior
Department of Psychology, University of Southern Queensland

and

Dr. Lucille Douglas
Medical Consultants Australia

Poster Presented at the 21st Annual Conference of the
National Academy of Neuropsychology
San Francisco, California, USA
October 31 – November 3, 2001

INTRODUCTION

The Minnesota Multiphasic Personality Inventory-2 (MMPI-2) is one of the most commonly used and studied personality inventories in the world and enjoys widespread use in the assessment of personal injury claimants. 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), and Fp (Infrequency-psychopathology) are used to examine the validity of the protocol. 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, exaggeration of symptomatology, or “faking bad” and the protocol is rendered uninterpretable or invalid.

Recently, a new scale, Fp has been introduced designed explicitly for the purpose of assessing symptom exaggeration (Arbisi & Ben-Porath, 1995). The Fp scale was designed to detect ‘infrequency’ in settings where a high F and Fb endorsement rate is known to exist (i.e., 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.

It is unclear exactly when ‘infrequency’ became equated with symptom exaggeration, although the hypothesis dates back to the origins of the MMPI itself. However, it is now standard MMPI-2 interpretation procedure to reject protocols with elevated F, Fb and/or F(p).

This presentation seeks to challenge the assumption that the Infrequency scales of the MMPI-2 are capable of detecting symptom exaggeration at least in the context of medicolegal assessment of personal injury claimants. A database of normal and clinical 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.

The Revised MMPI-2 Manual

Additionally, new guidelines for the use of F, Fb, and Fp have been published in the revised MMPI-2 manual (Butcher, Graham, Ben-Porath, Tellegen, Dahlstrom, & Kammer, 2001) and are summarised below:

F Scale: Implications of Scores in Outpatient Clinical Settings

T-Score	T-Score	Interpretation
F ≥ 90	VRIN or TRIN ≥ 79	Invalid and uninterpretable profile
F ≥ 90	Fp ≥ 100	Overreporting psychopathology
F≥ 90	Fp = 70 – 99	Likely exaggerated, but may be valid
F = 70 – 89	Fp = 70 – 99	May be exaggerated but likely is valid
F = 55 – 69		Likely valid
F ≤ 54		May be defensive

Adapted from Butcher et al. (2001)

Interestingly, the role of Fb seems to have been fundamentally changed with the new revision:

> "T scores on Fb should only be used to determine whether a substantial change has occurred in the individual's approach to the MMPI-2. . . . when the MMPI-2 is administered in clinical settings, such a change is indicated when the T score on Fb exceeds 109 and is at least 30 points greater than the T score on F."
>
> (Butcher et al., 2001, pg. 18-19)

The impact of these new interpretative guidelines will also be examined on the detection of invalid or exaggerated protocols in a personal injury claimant database.

METHOD

Participants:

A clinical sample was derived from the MMPI-2 protocols of 2,241 (1278 males, 963 females) personal injury claimants who were administered the MMPI-2 as part of their psychological assessment at a forensic psychiatric practice in Brisbane, Queensland, Australia. The age of this clinical group ranged from 16 to 78 years with a mean of 39.38 (SD = 11.35). The mean number of formal years of education was 11.82 (SD = 3.12) and ranged from 4 to 26 years. The number of different diagnostic groups in the database are indicated in the tables below. Some groups have not been represented individually due to insufficient numbers: Anorexia Nervosa (N=1); Bereavement (N=6); Delusional Disorder (N=1); Dementia (N=2); Dysthymia (N=17);

Obsessive Compulsive Disorder (N=5); Paraphilia (N=1); Pedophilia (N=1); Sexual Harassment (N=1); and whiplash injuries (N=16). There were a further 126 cases for which diagnoses were currently unavailable. Base rates for these groups are combined in the Total group. Diagnoses were made by three experienced forensic psychiatrists using DSM-IV criteria, where applicable.

Forty-nine cases in the database had been identified through surveillance as attempting to misrepresent or exaggerate their injuries. This subsample was analysed separately in the current study and have been termed malingerers. This does not mean that no instances of exaggeration or malingering occurred in the remainder of the database. However, this sample of confirmed malingerers is designed to be representative of those individuals that the MMPI-2 infrequency measures are presumed to be able to detect. A further 112 cases were removed from the database before analyses were conducted due to elevations in VRIN or TRIN that suggested inconsistent responding on the test. This left a total of 2080 cases in the clinical sample. For the purposes of developing multiple regression equations for predicting infrequency scale T-scores, this sample was further divided into a development sample of 1009 cases (Sample 1) and a cross-validations sample of 1027 cases (Sample 2).

A normal sample was compiled of 1,202 (856 males, 342 females) applicants for civil service positions who were administered the MMPI-2 as part of their application process. The mean age for this group was 27.43 years (SD = 5.79; range 18 to 50) and the mean number of years of education was 14.35 (SD = 2.82; range = 8 to 26).

RESULTS

Tables 1 through 3 present base rate data for the MMPI-2 infrequency scales for personal injury claimants, the subsample of confirmed malingerers, and the normal sample.

Four types of information are presented in each table:

(a) Base rates for each of the major diagnostic groups in the database as a function of level of elevation for which interpretative guidelines have been presented in the revised MMPI-2 manual.

(b) Regression equations for predicting F, Fb, and Fp T-scores from Basic and Content scales and subscales. These equations were generated using stepwise regression with all Basic and Content subscales of sample 1 entered. Basic and Content scales were also included where no subscales existed. The selection of the final regression equation was based upon that combination of scales and subscales that were significantly correlated with the infrequency scale, accounted for the greatest amount of variance, and for which there was no evidence of multicollinearity. The retained equation is presented along with the multiple correlation, squared-correlation, and standard error of estimate.

(c) Cross-validation using sample 2 was then conducted to determine the degree of shrinkage when applied to other cases. Included are the means, standard

deviations, and multiple correlations for the development, cross-validation, malingerer, and normal samples.

(d) Difference scores were then generated for each observed and predicted infrequency scale. The differences associated with a variety of percentile ranks were then computed.

Tables 4 and 5 more explicitly examine the new interpretative guidelines from the revised MMPI-2 manual in terms of the distribution of difference scores between F and Fb (Table 4) and the test operating characteristics of the two rules for detecting exaggerated protocols in discriminating between Malingerers and Clinical or Normal cases (Table 5). Finally, a principal axis factor analysis with oblique rotation was conducted on the clinical sample to examine whether or not the infrequency scales would load on a single factor consistent with their hypothesised role of detecting "faking bad".

Table 1a. Base Rates of F in a Personal Injury Claimant Sample

MMPI-2 Manual (2001) Criteria		**Percentage of Cases with Elevations in F**			
		May be Defensive	**Likely Valid**	**May be exaggerated**	**May be invalid**
Group	**N**	**T ≤54**	**T = 55-69**	**T = 70-89**	**T >=90**
Adjustment Disorder	223	32.3	41.7	21.5	4.5
Anxiety Disorder	65	16.9	33.8	40.0	9.2
Bipolar Disorder	21	23.8	38.1	14.3	23.8
Chronic Fatigue Syndrome	27	51.9	33.3	11.1	3.7
Depression	199	23.6	38.7	25.1	12.6
Dissociative Disorders	12	0.0	0.0	41.7	58.3
Medical Conditions	52	32.7	51.9	9.6	5.8
Nil Dx	125	46.4	33.6	12.8	7.2
Pain	383	37.6	36.8	16.7	8.9
Panic Disorder	33	21.2	36.4	30.3	12.1
Personality Disorder	30	26.7	20.0	33.3	20.0
Phobia	36	61.1	16.7	16.7	5.6
Post-traumatic Stress Disorder	181	13.8	35.4	34.3	16.6
Schizophrenia	12	16.7	33.3	33.3	16.7
Somatoform Disorder	63	46.0	28.6	20.6	4.8
Substance Abuse	128	18.0	32.0	30.5	19.5
Traumatic brain injury	313	31.3	37.1	21.4	10.2
Total	2080	30.8	35.8	22.8	10.7
Malingerer	49	32.7	20.4	28.6	18.4
Normals	1202	96.8	2.8	0.2	0.2

Table 1b. Predicting F from Basic and Content Scales

F = Sc1*.293 + Sc6*.11 + Sc2*.197 + Biz1*.234 + Fam2*.156 + Sc3*.129 + Asp2*.151 + Hea1*.073 + Sod1*.096 – 21.249			
N = 990	r = .901	r^2 = .812	SEe = 8.22

Table 1c. Cross-Validation of Subscale Regression Equation

	Sample 1	(N = 1009)	Sample 2	(N = 1027)
	F	Predicted F	F	Predicted F
Mean	65.92	65.36	63.93	64.17
SD	19.39	17.42	18.09	16.51
R	.901		.889	
	Malingerers	(N = 49)	Normals	(N = 1202)
	F	Predicted F	F	Predicted F
Mean	70.31	71.64	42.64	43.47
SD	24.89	21.58	5.75	5.26
R	.922		.674	

Table 1d. Distribution of Difference Between Obtained and Predicted F Scores

	Fp-Fpred Differences Associated with Specific Percentiles										
Group	1	5	10	15	25	50	75	85	90	95	99
Clinical	-20.5	-14.1	-10.3	-8.3	-5.3	-0.1	5.5	8.6	10.4	13.9	21.5
Malingerers	-18.5	-13.7	-12.2	-10.7	-7.9	-3.0	5.1	8.7	9.1	16.0	38.3
Normals	-11.8	-7.7	-6.0	-5.2	-3.9	-1.0	1.9	3.7	4.6	7.0	10.8

Table 2a. Base Rates of Fb in a Medicolegal Sample

MMPI-2 Manual (2001) Criteria		Percentage of Cases with Elevations in Fb		
		Unelevated	Elevated	May Be Invalid
Group	N	T ≤64	T = 65-109	T = 110+
Adjustment Disorder	223	64.4	33.8	1.9
Anxiety Disorder	65	40.0	56.9	3.1
Bipolar Disorder	21	421.9	42.9	14.3
Chronic Fatigue Syndrome	27	70.4	29.6	0.0
Depression	199	45.2	47.2	7.6
Dissociative Disorders	12	0.0	66.7	33.3
Medical Conditions	52	74.0	20.0	6.0
Nil Dx	125	73.9	21.8	4.2
Pain	383	61.1	33.1	5.8
Panic Disorder	33	45.5	45.5	9.1
Personality Disorder	30	44.8	31.0	24.1
Phobia	36	65.7	25.7	8.6
Post-traumatic Stress Disorder	181	30.3	56.7	12.9
Schizophrenia	12	41.7	50.0	8.3
Somatoform Disorder	63	62.3	36.1	1.6
Substance Abuse	128	40.0	42.4	17.6
Traumatic brain injury	313	60.1	32.5	7.5
Total	2080	53.5	36.9	7.5
Malingerer	49	43.8	27.1	29.2
Normals	1202	99.7	0.1	0.2

Table 2b. Predicting Fb from Basic and Content Scales

FB = .068*DEP1 + .314*FRS1 + .291*DEP4 + .202*BIZ2 + .226*TRT1 + .204*SC1 + .109*SC6 – 19.333			
N= 990	r = .939	r^2 = .881	SEe = 7.87

Table 2c. Cross-Validation of Subscale Regression Equation

	Sample 1 (N = 1009)		Sample 2 (N = 1027)	
	Fb	Predicted Fb	Fb	Predicted Fb
Mean	68.76	68.65	67.14	67.23
SD	23.11	21.63	22.41	20.92
r	.939		.937	
	Malingerers (N = 49)		Normals (N = 1202)	
	Fb	Predicted Fb	Fb	Predicted Fb
Mean	79.35	80.26	43.24	42.50
SD	29.52	27.38	4.12	5.08
R	.957		.841	

Table 2d. Distribution of Differences Between Obtained and Predicted FB Scores

	FB-FBpred Differences Associated with Specific Percentiles										
Group	1	5	10	15	25	50	75	85	90	95	99
Clinical	-19.2	-12.4	-9.7	-7.6	-4.7	-0.3	4.7	7.7	10.3	13.9	21.5
Malingerers	-18.1	-16.2	-12.3	-10.1	-6.9	-1.2	4.2	6.9	13.2	15.5	16.8
Normals	-9.1	-4.2	-2.2	-1.3	-0.2	1.1	1.7	2.6	3.1	4.9	6.9

Table 3a. Base Rates of Fp in a Personal Injury Claimant Sample

MMPI-2 Manual (2001) Criteria		Percentage of Cases with Elevations in Fp		
Group	N	Likely Valid **T ≤69**	Likely Exaggerated **T = 70-99**	Likely Invalid Faking Bad **T = 100+**
Adjustment Disorder	223	90.3	9.3	0.5
Anxiety Disorder	65	86.2	13.8	0.0
Bipolar Disorder	21	76.2	19.0	4.8
Chronic Fatigue Syndrome	27	100.0	0.0	0.0
Depression	199	82.7	16.8	0.5
Dissociative Disorders	12	33.3	58.3	8.3
Medical Conditions	52	84.0	16.0	0.0
Nil Dx	125	81.5	15.1	3.4
Pain	383	83.9	14.8	1.3
Panic Disorder	33	78.8	18.2	3.0
Personality Disorder	30	79.3	13.8	6.9
Phobia	36	88.6	5.7	5.7
Post-traumatic Stress Disorder	181	79.2	18.5	2.2
Schizophrenia	12	83.3	16.7	0.0
Somatoform Disorder	63	88.5	11.5	0.0
Substance Abuse	128	75.2	21.6	3.2
Traumatic brain injury	313	80.8	16.9	2.3
Total	2080	81.1	15.1	1.7
Malingerer	49	75.0	16.7	8.3
Normals	1185	97.1	2.7	0.2

Table 3b. Predicting Fp from Basic and Content Scales

Fp = .245*Biz1 + .228*Sc1 + .484*Ma3 + .195*FRS1 + .2*FAM2 - .262*Pd3 - 2.981			
N = 1009	r = .694	r^2 = .481	SEe = 11.08

Table 3c. Cross-Validation of Subscale Regression Equation

	Sample 1 (N = 1009)		Sample 2 (N = 1027)	
	Fp	Predicted Fp	Fp	Predicted Fp
Mean	56.93	56.72	55.55	56.20
SD	15.54	10.85	14.56	10.28
r	.694		.673	
	Malingerers (N = 49)		Normals (N = 1185)	
	Fp	Predicted Fp	Fp	Predicted Fp
Mean	60.33	60.70	48.47	50.40
SD	20.24	13.31	8.69	4.62
r	.852		.479	

Table 3d. Distribution of Difference Between Obtained and Predicted Fp Scores

	Fp-Fpred Differences Associated with Specific Percentiles										
Group	1	5	10	15	25	50	75	85	90	95	99
Clinical	-23.2	-16.5	-13.0	-10.8	-7.9	-1.4	6.5	11.0	14.3	18.9	30.2
Malingerers	-17.1	-15.9	-12.3	-11.7	-8.6	-1.7	2.9	11.3	17.8	21.4	36.0
Normals	-15.6	-12.5	-10.6	-9.3	-7.4	-3.0	2.6	6.0	8.5	13.0	19.6

Table 4. Distribution of Differences Between F and FB

	F-FB Differences Associated with Specific Percentiles										
Group	1	5	10	15	25	50	75	85	90	95	99
Clinical	-41	-28	-21	-17	-10	-1	6	9	12	16	25
Malingerer	-48	-42	-33	-25	-14	-6	0	4	9	12	15
Normal	-10	-7	-6	-5	-3	-1	2	3	6	9	13

Table 5. Test Operating Characteristics of Two Rules for Detecting Exaggerated or Invalid Protocols in Personal Injury Claimants

	Malingerers Compared to Normals (prevalence = 0.039)					Malingerers Compared to Clinicals (prevalence = 0.023)				
Comparison	SENS	SPEC	**PPP**	NPP	OPP	SENS	SPEC	**PPP**	NPP	OPP
Fb≥110 and Fb-F≥30	0.10	1.00	**1.00**	0.97	0.97	0.10	0.99	**0.17**	0.98	0.97
F≥90 and Fp>100	0.08	1.00	**0.67**	0.96	0.96	0.08	0.98	**0.11**	0.98	0.96

Table 6. Exploratory Factor Analysis – Personal Injury Claimants (N = 2080)

Factor 1		Factor 2		Factor 3		Factor 4		Factor 5		Factor 6	
SC4	0.73	SI2	0.71	PD3	0.92	**FP**	**0.45**	RE	0.49	ES	0.42
DEP1	0.71	SOD1	0.68	HY1	0.86	FRS1	0.40	GF	0.42	HY3	-0.54
DEP2	0.69	R	0.46	MA3	0.67			MAC-R	-0.42	HEA1	-0.56
D5	0.68	P5POS	-0.53	SOD1	-0.42			AAS	-0.49	SC6	-0.59
D4	0.67			SOD2	-0.82			ASP2	-0.87	D3	-0.66
D1	0.67			SI1	-0.84			PD2	-0.87	HEA3	-0.73
SC2	0.64									HEA2	-0.83

PD5	0.61									HY4	-0.85
TRT1	0.60									HS	-0.98
PT	0.60										
MT	0.59										
HY3	0.54										
DEP4	0.52										
ANX	0.52										
LSE1	0.52										
A	0.50										
DEP3	0.49										
PS	0.48										
PK	0.48										
D2	0.47										
WRK	0.47										
SC3	0.47										
FB	**0.41**										
Pa2	0.36										
43.5%		10.3%		4.8%		3.4%		2.9%		2.3%	
Factor 7		Factor 8		Factor 9		Factor 10		Factor 11			
PA3	0.90	PD1	0.97	ANG2	0.60	FRS2	0.52	PA1	0.60		
HY2	0.70	FAM2	0.73	TPA1	0.55	GM	-0.49	BIZ1	0.55		
S	0.46	FAM1	0.64					P5PSY	0.54		
SI3	-0.41	MDS	0.59					PD4	0.46		
CYN2	-0.49							MA4	0.43		
ASP1	-0.79										
CYN1	-0.90										
2.0%		1.8%		1.5%		1.4%		1.2%			

A number of observations can be made from the data presented in these tables:

1. While elevations of infrequency scales are a rare occurrence in normal individuals, they are far from uncommon in either clinical groups or malingerers.
2. The majority of clinical cases and malingerers do not generate extreme elevations in infrequency scales.
3. F Scale: Malingerers are almost twice as likely to produce extreme elevations than the clinical group. However, a number of diagnostic groups are just as likely to produce extreme elevations.
4. Fb Scale: On average, malingerers are more than three times as likely to generate extreme elevations than the clinical group. However, as with the F scale at least two diagnostic groups exhibit the same tendency.
5. Fp scale: Malingerers are more than four times as likely to generate extreme elevations as the clinical group. The clinical diagnostic groups show a similar score distributions to the malingerers in the “Likely Exaggerated” range but more closely resemble the normal group in the “Faking Bad” range.
6. Both F and Fb scales can be predicted highly accurately from specific combinations of Basic and Content subscales . 81% of the variance in F can

be accounted for by a weighted combination of scales that relate to alienation, psychotic symptomatology, antisocial behaviour, cognitive difficulties, and introversion. 88% of the variance in Fb can be accounted for by a combination of scales that relate to low motivation, suicidal ideation, anxiety, and alienation.

7. Prediction of Fp is far less accurate and accounts for 48% of the variance with a weighted combination of scales that relate to psychotic symptomatology, alienation, and generalised anxiety.
8. Cross-validation reveals little to no shrinkage in the predictability of the equations in a second sample of personal injury claimants. These equations seem to be as accurate and applicable to the malingerer sample. Lower correlations were observed for the normal sample, but this is not surprising as the low frequency of elevations in F, Fb, and Fp results in a substantial restriction of range.
9. The applicability of the equations for the malingerer group is also reflected in the similarity between the difference score distributions of this group and the clinical sample.
10. The recommended difference between F and Fb of 30 or more points is a rare event occurring in less than 5% of clinical cases. With the added requirement of Fb greater than 109, the base rate of this occurrence drops to only 1.2%.
11. While it is clear that the malingerer sample is more likely to generate extreme elevations on infrequency scales than members of the clinical sample, the positive predictive power of the interpretative guidelines (from 0.11 to 0.17) is insufficient to permit any accurate detection of malingering.
12. The exploratory factor analysis indicates that Fb and Fp share more in common with other clinical scales than they do with each other. F achieved no loadings greater than 0.3 on any of the 11 retained factors. This challenges the assumption of any common role for these scales.

DISCUSSION

This study has revealed a number of concerns regarding the roles of F, Fb, and Fp as indicators of exaggerated symptomatology.

1. F, Fb, and to a lesser extent Fp 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 does indicate 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. This is also reflected in their differential loadings in the exploratory factor analysis.

3. 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.

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 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 many clinical settings, it would be tragic to misinterpret an open acknowledgement in a suicidal client as "exaggeration of symptomatology" on the MMPI-2.

RECOMMENDATIONS

Scores on F and Fb are highly predictable based upon a combination of MMPI-2 scales and subscales. It is recommended that before attempting to interpret elevations in F, Fb, or Fp the degree to which they are a likely result of the commonplace combinations of subscales be examined. In this way the clinician can determine how likely it is that the obtained elevations in the infrequency scales are a consequence of known and robust correlates. This is achieved by computing the predicted F, Fb, and Fp scores using the regression equations provided. The differences between the observed and predicted scores for F, Fb, and Fp can be evaluated by consulting the distribution of difference score tables to determine the frequency with which this difference occurs in personal injury claimants, malingerers, or civil service applicants. The implication of an infrequent or abnormal difference score is that the observed infrequency scale elevation is unlikely to be a consequence of the known clinical correlates that contribute to elevations on these scales.

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.

Butcher, J.N., Graham. J.R., Ben-Poarth, Y.S., Tellegen, A., Dahlstrom, W.G., & Kaemmer, B. (2001). Minnesota Multiphasic Personality Inventory-2. Manual for administration, scoring, and interpretation. Revised edition. Minneapolis: University of Minnesota Press.

Tuesday, 13 November 2001

Senior Lecturer
Department of Psychology
University of Southern Queensland
Toowoomba, QLD 4350

Australia