The European Journal of Psychology Applied to Legal Context 5 (2013) 123-129

The European Journal of Psychology Applied to Legal Context

www.elsevier.es/ejpal

Are patients with chronic pain and fibromyalgia correctly classified by MMPI-2 validity scales and indexes?

Alfonso Palmer[a]*, Carmen Borrás[a], Javier Pérez-Pareja[a], Albert Sesé[a], and Manuel Vilariño[b]

[a]*Universitat de les Illes Balears, Spain*
[b]*Universidade de Santiago de Compostela, Spain*

ARTICLE INFORMATION

Manuscript received: 2/05/2012
Revision received: 08/02/2013
Accepted: 11/02/2013

Key words:
MMPI-2
Malingering
Fibromyalgia
Chronic pain

ABSTRACT

A study was designed to find out whether MMPI-2 validity scales and indexes differentiate between true fibromyalgia sufferers, patients with chronic organic pain and normal people, as well as whether they are correctly classified. 105 subjects participated in the study, 27 diagnosed with fibromyalgia and 44 with chronic organic pain and 34 were healthy people, who answered the MMPI-2 following standard instructions. The results showed that fibromyalgia patients scored higher than the control group in the F, Fb, F-K, Fp, Ds, and FBS scales and indexes and that patients with chronic organic pain scored higher in the Ds and FBS scales than the control group. The case study revealed that the F, Fb, F-K, Fp, Ds, and FBS scales and indexes over-diagnose malingering in patients with fibromyalgia, both in comparison with the clinical population and with the normative group. Likewise, patients with chronic organic pain were over-diagnosed as malingerers by all the scales and indexes in comparison with the normative population and by the Fp and FBS scales in comparison with the clinical population. In addition, it was found that at least one of the scales for measuring defensiveness –L, Wsd and Mp– classified 79.5% of the truly ill patients as faking good. The implications for clinical and forensic practice are discussed, as well as for the definition of decision criteria and the (re)classification as true negatives of genuine cases classified as malingerers by the malingering measuring scales and indexes.

¿Son clasificados correctamente los pacientes con dolor crónico y fibromialgia por las escalas e índices de validez del MMPI-2?

Palabras clave:
MMPI-2
Simulación
Fibromialgia
Dolor crónico
Disimulación

RESUMEN

Se diseñó un estudio con el objetivo de conocer si escalas e índices de validez del MMPI-2 diferencian entre verdaderos enfermos de fibromialgia, pacientes con dolor crónico de etiología orgánica y personas normales, así como si los clasifican correctamente. Participaron en el estudio 105 sujetos, 27 diagnosticados de fibromialgia, 44 de dolor crónico con etiología orgánica y 34 sanos, que respondieron al MMPI-2 bajo las instrucciones estándar. Los resultados mostraron que los pacientes de fibromialgia puntuaban más alto en las escalas e índices F, Fb, F-K, Fp, Ds y FBS que el grupo control y los pacientes de dolor crónico con etiología orgánica en las escalas Ds y FBS. El estudio de casos evidenció que las escalas e índices F, Fb, F-K, Fp, Ds y FBS sobrediagnostican simulación en los pacientes con fibromialgia, tanto en comparación con la población clínica como con la normativa. Asimismo, los pacientes con dolor crónico con etiología orgánica fueron sobrediagnosticados como simuladores por todas las escalas e índices en comparación con la población normativa y por las escalas Fp y FBS en comparación con la población clínica. Además hallamos que al menos una de las escalas de medida de defensividad –L, Wsd y Mp– clasificaba como disimuladores al 79.5% de los verdaderos enfermos. Se discuten las implicaciones para la práctica clínica y forense, así como para la definición de criterios de decisión y la (re)clasificación como verdaderos negativos de aquellos casos genuinos clasificados como simuladores por las escalas e índices de medida de la simulación.

*Correspondence concerning this article should be sent to Alfonso Palmer, Facultad de Psicología, Ctra. Valldemossa Km. 7,5, 07122 Palma (Baleares), Spain. E-mail: alfonso.palmer@uib.es

1889-1861/$ - see front matter © 2013 Colegio Oficial de Psicólogos de Madrid. Todos los derechos reservados

DOI: http://dx.doi.org/10.5093/ejpalc2013a1

The International Association for the Study of Pain (Merskey & Bogduk, 2010) defines pain as an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage, specifying that it is chronic if it persists beyond the normal tissue healing time (usually 3 months). The time that needs to elapse for pain to be specified as chronic varies from one study to another depending on the type of population, research methodology, etc., oscillating between 3 and 6 months, and some authors have also added criteria such as intensity and frequency for pain to be specified as chronic (Johannes, Le, Zhou, Johnston, & Dworkin, 2010). As a result of all this and of potential cultural differences in pain management, prevalence rates are highly fluctuating. Thus, Breivik, Collet, Ventafridda, Cohen, and Gallacher (2006) found a prevalence of 19% among European adults, whereas Johannes et al. (2010) found 30.7% in North American adults. Either way, not only does chronic pain constitute one of the greatest health problems but it is also the most important (Breivik et al., 2006) and the most common cause of disability in the middle-aged population (Scascighini, Toma, Dober-Spielmann, & Sprott, 2008). It also carries high socio-medical costs (Gaskin & Richard, 2012; Van Leeuwen, Blyth, March, Nicholas, & Cousins, 2006). Chronic pain is comorbid with emotional suffering, psychopathology, and with alterations in relational, social, and professional activity (Gatchel, 2004; Ilgen, Zivin, McCammon, & Valenstein, 2008; Tecic, Lefering, Althaus, Rangger, & Neugebauer, 2013; Tsang et al., 2008).

Generally, chronic pain occurs with organically caused pathologies, but this is not always so. Fibromyalgia, which was recognised as a disease by the World Health Organization (1992) with its incorporation in ICD-10, was characterized by the absence of a clear organic etiology, but with the presence of chronic pain (Bellato et al., 2012). Previously the American College of Rheumatology (ACR) had defined the diagnostic criteria: pressure pain in at least 11 out of 18 tender points and the presence of widespread pain (Wolfe et al., 1990). Faced with the difficulties in applying and assessing these criteria whereby the diagnosis depended practically exclusively on patients' self-reports, in 2010 the ACR itself introduced the Widespread Pain Index (WPI), a measure of the number of areas in which the patient complains of pain, and severity scales for fatigue, cognitive symptoms, unrefreshing sleep, and some somatic symptoms, which were to be comprehensively integrated in the Symptom Severity scale (SS). Hence, by combining the WPI and the SS a new definition of fibromyalgia is formulated in the following terms: (WPI ≥ 7 and SS ≥ 5) or (WPI 3-6 and SS ≥ 9). Either way, it is a very complicated diagnosis, which opens the door to misdiagnoses (Wierwille, 2012). The prevalence of fibromyalgia is between 2 and 3% (Branco et al., 2010; Wolfe, Ross, Anderson, Rusell, & Hebert, 1995) and also entails high socio-medical costs (Winkelmann et al., 2011).

The perception of pain constitutes a subjective experience which makes its objective measurement difficult, while at the same time it is a fertile ground for claiming compensation (e.g., damages, disability pensions). In assessing compensation, which falls squarely in the field of forensic evaluation, it is always necessary to make a differential diagnosis of malingering (American Psychiatric Association, 2000), which Greve, Ord, Bianchini, and Curtis (2009) calculated at between 20 and 50%.

The comorbidity of chronic pain with emotional suffering (that is, psychological damage), psychopathology, with alterations in relational, social and professional activity entails the fact that, in addition to medical tests and diagnosis, it also requires a psychological assessment of these comorbid pathologies and the differential diagnosis of malingering. Since the perception of pain is purely subjective, physical tests play a limited role in the differential assessment of malingering; therefore psychological assessment is a good, even indispensable, complement for differential diagnosis. Although instruments have been developed for this purpose, some of them exclusively, such as TOMM, MMPI is the benchmark instrument in psychological assessment and the assessment of malingering (Greene, 2011; Rogers, 2008; Rogers, Sewell, Martin, & Vitacco, 2003), including chronic pain (Aguerrevere, Greve, Bianchini, & Meyers, 2008; Greve et al., 2009; Huss, 2009; Meyers, Millis, & Volkert, 2002).

Therefore we undertook, in the framework of psychological assessment associated to chronic pain and fibromyalgia, a field study in order to find out whether MMPI-2 validity scales and indexes –the benchmark instrument in forensic psychological assessment– differentiate between genuine patients with chronic organic pain, patients with chronic fibromyalgia, and a control group. Likewise, we ascertain whether genuine patients are correctly classified as having chronic pain with an underlying organic etiology and fibromyalgia. Finally, we seek the forensic gold standard definition, as well as new criteria which will enable the correct identification of genuine patients.

Method

Participants

105 subjects participated in the study divided into 3 groups: patients diagnosed with fibromyalgia, patients diagnosed with chronic pain, and healthy subjects. The fibromyalgia group was made up of 27 participants, 24 women (88.9%) and 3 men (11.1%), with a mean age of 46.48 years (95% CI: 42.53-50.43). The chronic pain group was made up of 44 patients, 38 women (86.4%) and 6 men (13.6%) with a mean age of 45.82 (95% CI: 42.97-52.43), and the group of healthy individuals (hereinafter, control group) was made up of 34 healthy people, 31 women (91.2%) and 3 men (8.8%), with a mean age of 48.38 (95% CI: 44.33-52.43).

Measurement instrument

The Spanish adaptation of the MMPI-2 (Hathaway & Mckinley, 1999) was adopted as the measuring instrument. For the differential diagnosis of malingering –a requirement in a forensic evaluation context (American Psychiatric Association, 2000)– all the necessary scales and indexes were taken in order to follow the model by Arce, Fariña, Carballal, and Novo (2006, 2009) of research and knowledge transfer into forensic practice: measure of cooperation with the evaluation, measure of consistency, and measure of negative (i.e., which invalidate the protocol) and positive (i.e., which validate the protocol) criteria for the evaluation of the validity of the protocols. Negative criteria are those that invalidate the protocol due to malingering and positive criteria are the ones that validate it. The measure of collaboration with the evaluation was assessed with the Cannot Say scale (?) and the consistency of the responses by TRIN (acquiescence) and VRIN and F-Fb (random responses). For the study of negative criteria regarding malingering we took the scales and indexes that the literature has systematically related to predictors of this: Infrequency scale (F), Back Infrequency scale (Fb), F minus K index (F-K), Infrequency-Psychopathology scale (Fp), Gough Dissimulation scale (Ds), and Fake Bad scale (FBS) (Friedman, Lewak, Nichols, & Webb, 2001; Graham, 2006; Greene, 2011; Rogers, Sewell, Martin, & Vitacco, 2003). Ds scale was preferred to the abbreviated version Ds-r, because of its large effect size (d = 1.62 vs. 1.49), sophisticated strategy and minimal false positives (Rogers et al., 2003). As measures of potential positive validity criteria we took the scales and indexes of relevance for the study of defensiveness, as these can be effective in identifying true clinical cases since they are not related to malingering and are not detected in malingerer protocols but are found in assessments of psychological damage in genuine victims of crime (Arce et al., 2006, 2009): Lie scale (L), Correction scale (K), F minus K index (F-K), Positive Malingering

scale (Mp), Wiggins Social Desirability scale (Wsd), Edwards Social Desirability scale (Esd), Other Deception scale (Odecep) and Superlative scale (S) (Baer & Miller, 2002; Friedman et al., 2001; Graham, 2006, Greene, 2011). Although Odecep was created as an updated version of the Mp, the results of this study showed a significant incremental validity for chronic pain case classification, $\chi^2(1) = 6.25$, $p < .05$, for Mp in relation to Odecep. Thus, both scales were considered for the study.

Procedure and design

The participants making up the groups of Fibromyalgia and Chronic Pain (non-fibromyalgia) were patients diagnosed as such by the health services in Mallorca (Spain), were under treatment, and had a clinical record (in all cases there was a record of specification > 6 months, high frequency, and intensity). The patients diagnosed with chronic pain had a clinical record of a non-inflammatory objectified pathology of the musculoskeletal system that supported organic etiology. The diagnosis of fibromyalgia was contrasted with rheumatologists who were blind to the original diagnosis, according to classification criteria for the disease (Wolfe et al., 1990; Wolfe et al., 2010). We included only the cases in which, as well as coinciding in the diagnosis, there was a total coincidence in diagnostic criteria. Patients with prior diagnoses of mental illness and those waiting to receive some possible secondary gain (disability pension, economic compensation, etc.) were rejected for the study.

Once the patients had been selected and the cross-diagnosis of fibromyalgia had been verified, the Spanish adaptation of MMPI-2 (Hathaway & McKinley, 1999) was applied in standard response conditions by assessors who were unaware of the patient's diagnosis. Patients gave their written consent to the evaluation; afterwards they were given a debriefing and an interview to learn their degree of task engagement and motivation (recall and comprehension of instructions) and to ensure participants had understood and completed the task correctly (Rogers, 2008). The resulting clinical profile in the MMPI-2 of patients was reviewed by the doctors of reference who confirmed that this was compatible with their records and diagnostic impression.

A design was planned to compare the measurements of the clinical patients diagnosed with fibromyalgia and chronic pain with a control group in the MMPI-2 validity measures. The control group, with no pain or chronic pathology, was randomly selected to fit the socio-demographic characteristics of the patients with chronic pain and fibromyalgia out of the people who attended Primary Care Services with mild health problems. The groups were matched by taking into account the socio-demographic characteristics (e.g., gender, age, job situation) of fibromyalgia patients (smallest group) and the clinical record (e.g., time since diagnosis, confidence of the diagnosing physician). A case study was also planned with the normative and clinical populations as contrast values.

Data analysis

Firstly, the difference in means between groups was studied; then, the assumptions of the analysis of variance of normality and homogeneity of variance of the variables were verified in order to define which statistical test to use. Table 1 presents the results of the test for normality (Shapiro-Wilk's test) and homogeneity of variance (Levene's test) of the measurement variables of the study. The lack of homogeneity of variance and/or normality in most of the variables to be studied, as well as our interest in comparing means, recommended the general use of a non-parametric test, the Kruskal-Wallis test, with Dunn-Bonferroni post hoc tests (Palmer, 2011).

The case study was addressed using classification rates, a test of the proportion observed with the expected observation (chronic pain and fibromyalgia groups vs. control group) using the *Z* statistic, and the effect size using the Odds Ratio (*OR*). As for the interpretation of the *OR*s in Cohen's effect size categories, these were converted, with the formula of Chinn (2000), to Cohen's *d*, which allows an interpretation in terms of *z*-scores.

Table 1
Summary of the Tests for Normality and Homogeneity of Variance of the Measurement Variables of the Study in the Protocols of Participants

		Normality	
		Yes	No
Variance Homogeneity	Yes	K, S, Wsd, VRIN,	L, TRIN
	No	PM, Esd, Odecep	F, Fb, F-K, P, Ds, FBS

Results

There are three steps that Arce et al. (2006, 2009) relate in the model of research and knowledge transfer into forensic practice in the study of MMPI protocols in forensic evaluation. The first step is the analysis of cooperation with the assessment, in such a way that if a lack of cooperation were to be observed, the protocol should be invalidated. In this respect, the results of the mean difference test in the Cannot Say scale (?) –which measures collaboration with the assessment through the number of items left unanswered or double answered– reveal that both patients with chronic pain and fibromyalgia and the control group collaborate with the assessment to the same extent.

The second step is the assessment of the consistency of responses both in terms of (non)acquiescence (TRIN) and random responses (VRIN, |F-Fb|). The results, which can be seen in Table 2, reveal differences between populations in acquiescence (TRIN), which the Dunn-Bonferroni post hoc tests pinpointed as patients diagnosed with chronic pain showing a more acquiescent pattern of responses than the control group, whereby it can be concluded that the responses in all the conditions are not acquiescent, even though the means of the three groups fall within the region of normality (5 < raw score < 13) in TRIN. Meanwhile, in VRIN and |F-Fb| the three groups have equal means that are within the region of normality. As a result, it can be concluded that, on the whole, the protocols of patients diagnosed with chronic pain and fibromyalgia are consistent.

The third step, which is performed if lack of cooperation and inconsistency have been eliminated in the responses, is the analysis of the validity of the protocol. This third step is, in turn, divided, depending on the availability of the control scales, into the original validity control scales and indexes –that is, those that were originally formulated in the MMPI and which are available in the commercial version and, as such, are the ones used by forensic psychologist in their everyday practice– and additional validity control scales which are used scientifically but are not in the commercial version of MMPI-2 and, therefore, are not used by forensic psychologists in their everyday practice. In the original validity scales, the results (see Table 2) showed statistically significant differences between samples in the F and Fb scales, and the F-K index. The Dunn-Bonferroni post hoc tests established that the patients diagnosed with fibromyalgia obtained significantly greater scores than the control group in the F and Fb scales and in the F-K index.

As regards the additional validity scales, the results unveiled differences in the Fp, Ds, FBS, and Esd scales mediated by the population factor (see Table 2). The Dunn-Bonferroni post hoc tests revealed that the patients diagnosed with chronic pain obtained significantly higher scores in the Ds and FBS scales than the control group and that the patients diagnosed with fibromyalgia obtained significantly higher scores in the Fp, Ds, and FBS scales and lower scores for Esd than the control group. Namely, the Ds and FBS scales

Table 2
Evidence of the Mean Difference Tests in the MMPI-2 Validity Indicators for the Population Factor (Fibromyalgia Patients, Chronic Pain Patients and Control Group)

Indicator	M_{CG}	M_{CP}	M_F	$\chi^{2(a)}$	*p*
Item omissions	0.03	0.02	0.00	0.74	.689
Consistency					
VRIN	8.12	8.52	8.85	0.65	.722
TRIN	9.35	10.41	9.96	6.09	.048
|F-Fb|	3.32	4.00	7.07	4.54	.103
Original validity scales					
F	6.94	9.55	21.67	29.11	.001
Fb	4.09	6.09	15.78	31.23	.001
L	6.97	6.84	7.15	0.10	.950
K	1 5.00	13.14	13.29	4.75	.093
F-K index	- 8.06	-3.59	8.37	26.66	.001
Additional validity scales					
Fp	2.15	3.16	8.37	14.09	.001
Ds	11.29	16.45	26.33	34.65	.001
FBS	17.03	20.93	24.56	17.56	.001
S	24.53	22.75	22.48	2.29	.318
Wsd	15.41	16.23	15.22	3.19	.203
Odecep	15.12	16.00	15.67	1.57	.456
PM	11.29	12.00	12.30	1.19	.552
Esd	24.91	21.66	17.63	27.90	.001

Note. *df*(2); [(a)]Kruskal-Wallis test.

report a tendency towards malingering among patients diagnosed with chronic pain, just like the Fp, Ds, and FBS scales did among patients diagnosed with fibromyalgia.

Mean comparison studies are extremely useful in science and research but are insufficient for knowledge transfer to professional practice, which requires *N* = 1 designs. This has led the American Psychiatric Association (2000) to recommend additional analyses to mean comparison and, within the forensic field, Arce et al. (2006, 2009) urged in their model of evaluation and knowledge transfer, in order to be able to transfer knowledge into forensic practice, mean comparison analyses to be conducted along with case studies and the derivation of forensic decision criteria. Therefore, we will immediately proceed with the case study and with the identification of forensic decision criteria.

Going back to the aforementioned steps, as regards cooperation with the assessment, the case study reveals that all the people assessed cooperated with the assessment, as none exceeded the cut-off point for suspected lack of cooperation with the assessment (raw score > 10; Graham, 2006). The analysis of the cases and items showed that patients with fibromyalgia answered all the items and that the only item left unanswered detected in patients with chronic pain is not part of the validity scales.

The inconsistency study shows (see Table 3) significantly higher rates (.02 was taken as the test criterion as these cut-off scores are calculated approximately on the 98th percentile) of cases classified by the inconsistency measures such as in TRIN and VRIN among patients diagnosed with chronic pain and in TRIN, VRIN, and |F-Fb|in the population of patients diagnosed with fibromyalgia. The *OR*s (observed/expected prevalence) exhibit a likelihood of classification of random responses, VRIN, which is 3.4 times greater for chronic pain patients than for the control group and 6.8 times greater for acquiescence, TRIN. Likewise, the *OR*s also indicate a likelihood of classification of acquiescence that is 6.8 times higher for fibromyalgia

Table 3
Classification Rate of the Inconsistency Indicators

		Chronic pain			Fibromyalgia		
Indicator	Cut-off score	*f*(*p*)	*Z*	*OR*	*f*(*p*)	*Z*	*OR*
TRIN	≥ r13[a]	6(.136)	5.50**	6.8	3(.111)	3.38**	5.6
VRIN	≥ r13[a]	3(.068)	2.27*	3.4	3(.111)	3.38**	5.6
|F-Fb|	≥ r10[b]	2(.045)	1.18	----	8(.296)	10.26**	14.8

Note. [a]cut-off score from Butcher, Dahlstrom, Graham, Tellegen, and Kaemmer (1989); [b]cut-off score from Greene (2008) for clinical settings; dashes indicate that *OR*s were not computed because the observed prevalence of cases was not statistically significant; +raw scores in TRIN ≤ 5 are also suggestive of inconsistency (response in the direction of false), but no contingency was recorded under this circumstance; *p < .05, **p < .001.

patients than for the control group and 5.6 times greater likelihood of random responses by TRIN and 14.8 by |F-Fb|.

Since the participants' responses were obtained in a non-forensic context and under standard instructions –that is, they are presumably honest answers– the classification of the protocol validity indicators as typical of malingering is a false positive. By comparing the observed proportion of protocols classified as fake bad (see Table 4) with the expected prevalence, as we were dealing with clinical cases, in the clinical population (the criteria taken from Caldwell are for the 98[th] percentile of the clinical population, in which the expected likelihood of wrong classification of real cases as fake bad is ≤ 2%, whereby the test value is .02) it was found that all the original and additional malingering indicators classified the protocols of patients diagnosed with fibromyalgia as invalid (false positives) beyond what was expected for clinical population, with an over-diagnosis of invalidity which oscillates between 11.1 times more than expected for the clinical population in the Ds scale and 20.4 in the FBS, in so far as the patients diagnosed with chronic pain would be over-diagnosed by the Fp and FBS scales (2 out of 6 indicators), with over-diagnosis of invalidity rates 4.6 times greater than for the clinical population for the Fp scales, and 8 times greater for FBS. Nevertheless, as in forensic practice the object of study are *N* = 1 designs, these results do not allow us to define a forensic decision criterion. At this point, Arce et al. (2006, 2009) urge, in their model of research and knowledge transfer, the definition of a forensic decision criterion, that is, typical of *N* = 1 designs, by studying the frequency of indicators per case. After implementing this, we found in 4 cases of fibromyalgia up to 5 indicators of malingering and in 1 case of chronic pain 5 indicators too, which enabled us to identify as true positives (fake good cases) the protocols in which the presence of 5 or more invalidating criteria is detected.

However, forensic evaluation is not considered a clinical context, wherefore the contrast population would not necessarily be the clinical one, but rather the normative one. The results (see Table 5)

Table 4
Classification Rate of Malingering Indicators with Clinical Cut-off Scores

		Chronic pain			Fibromyalgia		
Indicator	Cut-off score[a]	*f*(*p*)	*Z*	*OR*	*f*(*p*)	*Z*	*OR*
F scale≥	≥ r25	1(.023)	0.14	----	8(.296)	10.26**	14.8
F-K index	≥ r17	1(.023)	0.14	----	8(.296)	10.26**	14.8
Fb scale	≥ r21	0(0)	----	----	8(.296)	10.26**	14.8
Fp scale	≥ r7	4(.091)	3.36**	4.6	8(.296)	10.26**	14.8
Ds scale	≥ r35	2(.045)	1.18	----	6(.222)	8.25**	11.1
FBS	≥ r34	7(.159)	5.17**	8.0	11(.407)	14.39**	20.4

Note. [a]Cut-off scores in clinical settings from Caldwell for the 98[th] percentile (Greene, 2008); **p < .001.

Table 5
Classification Rate of Malingering Indicators with Normal Individual Cut-off Scores

		Chronic pain			Fibromyalgia		
Indicator	Cut-off score[a]	*f*(*p*)	*Z*	*OR*	*f*(*p*)	*Z*	*OR*
F scale≥	≥ r12	14(.318)	14.12**	15.9	20(.741)	26.80**	37.1
F-K index	≥ r4	11(.250)	10.90**	12.5	13(.481)	17.14**	24.1
Fb scale	≥ r9	10(.227)	9.81**	11.4	18(.667)	24.05**	33.4
Fp scale	≥ r4	17(.386)	17.35**	19.3	17(.630)	22.68**	31.5
Ds scale	≥ r24	9(.205)	8.77**	10.3	13(.481)	17.14**	24.1
FBS	≥ r22	20(.455)	20.62**	22.8	18(.667)	24.05**	33.4

Note. [a]cut-off scores in normal individuals (Butcher et al., 1989); **p < .001.

of the comparison of the observed likelihood with the one expected in the normative population (as in the test from the clinical population, we took the 98th percentile as the cut-off score, in this case from the manual for administration and scoring of the MMPI-2 of Butcher et al., 1989) reveal that all the scales and indexes for measuring malingering report this to be significantly greater in both samples than in the normative population, oscillating –in the sample with chronic pain– between an over-invalidation of the protocols 10.3 times higher than for the normative population in the Ds scale to 22.8 in FBS. The quantification of the overestimation of the invalidity of the protocol in patients with fibromyalgia (*OR*s) is even greater than among patients with chronic pain, with a range which oscillates between 24.1 times more than for the normative population in the Ds scale and F-K index to 37.1 times more in the F scale. Briefly, when forensic experts use these malingering indicators and decision criteria they are biased towards malingering. The study of the frequency of indicators per case does not allow us to establish a forensic decision criterion, as 6 malingering indicators were recorded both among those diagnosed both with chronic pain (n = 4) and with fibromyalgia (n = 6).

Continuing with the appraisal system of Arce et al. (2006, 2009), after the study of what they call negative criteria, that is, malingering indicators, comes the analysis of positive criteria, that is, of strategies for minimising false positives (authentic patients classified as malingerers). For our specific context, they propose the analysis of the MMPI-2 defensiveness scales and indexes. The cut-off points for suspected faking good are around the 95th percentile (stricter criteria are not recommendable because in the study assessment context, honest responses are expected and defensiveness is not suspected, and because of their contribution, to a greater extent than more conservative criteria, to the minimisation of false positives), which corresponds approximately with a likelihood, in the normative population, of cases of .05, which, in turn, is the criterion for statistical significance. By contrasting the expected likelihood (.05) with the observed one (see Table 6), we find a significantly higher prevalence of cases reported as typical of faking good than expected in the L, Wsd, MP, and Odecep scales in both the chronic pain and the fibromyalgia sample. For the case study, this means that fibromyalgia patients have a rate of cases classified as typical of defensiveness that is 4.4 times more than the expected likelihood for the Wsd scale, 5.9 for Odecep, 7.4 for Mp, and 9.6 for L. Meanwhile, in the sample of patients with chronic pain the rate of cases classified by Wsd or Odecep as making self-favourable reports is 6.4 times greater than expected, a large effect size, while it is 10.5 times for Mp and 11.4 for L, large effect sizes. Additionally, 79.5% of patients with chronic pain were classified by at least 1 scale as defensive, 50% by 2 scales, 29.5% by 3 scales, and 13.6% by all 4 scales, whereas 59.3% of patients with fibromyalgia were classified as defensive by 1 scale, 44.4% by 2 scales, 29.6% by 3 scales, and 3.7% by all 4 scales. Meanwhile, the K, S and Sd scales are totally insensitive to the labelling of the protocols in both samples, as characteristic of self-favourable reporting.

Table 6
Cut-off Scores, Classification Rate and Contrast of the Prevalence of the Defensiveness Indexes

		Chronic pain			Fibromyalgia		
Indicator	Cut-off score	*f*(*p*)	*Z*	*OR*	*f*(*p*)	*Z*	*OR*
L scale	≥ r7[a]	25(.568)	15.79**	11.4	13(.481)	10.29**	9.6
K scale	≥ r22[a]	0(0)	----	----	1(.037)	----	----
F-K index	< r-21[a]	0(0)	----	----	0(0)	----	----
S scale	≥ r39[a]	0(0)	----	----	1(.037)	----	----
Wsd scale	≥ r18[a]	14(.318)	8.17**	6.4	6(.222)	4.10**	4.4
Esd scale	≥ r30.25[b]	0(0)	----	----	0(0)	----	----
Mp scale	≥ r13[b]	23(.523)	14.42**	10.5	10(.370)	7.64**	7.4
Odecep scale	≥ r19[a]	14(.318)	8.17**	6.4	8(.296)	5.87**	5.9

Note. [a]Cut-off scores from Butcher et al. (1989); [b]cut-off scores from Baer and Miller (2002) are the mean scores of the meta-analysis for the non-coached under-reporting group; *Z* for a proof value (H_0) of .05; dashes for *Z* indicate that *Z* and the *OR* were not computed because the observed proportion of cases was ≤ .05 (H_0); *p < .01, **p < .001.

Discussion

Prior to discussing the results and their implications for forensic practice, it is necessary to warn of the limitations in their generalization and scope, which must be born in mind. First of all, although different steps were taken to ensure that patients with chronic organic pain and with fibromyalgia were genuine, it is not possible to totally guarantee that they were real patients suffering from chronic pain. Secondly, this thoroughness meant that the sizes of the experimental groups were reduced, which has an effect on the statistical estimations. Thirdly, the results cannot be generalised to other pathologies that are not chronic pain associated with fibromyalgia or with an organic etiology. Fourthly, MMPI-2 does not offer diagnoses, but rather diagnostic impressions, therefore it must necessarily be complemented with other evaluations, requiring a multi-method and, in this case, multidisciplinary approach (e.g., physical and medical tests); (Bianchini, Greve, & Glynn, 2005) for the differential diagnosis of malingering (Graham, 2006; Greene, 2011; Polusny & Arbisi, 2006; Pope, Butcher, & Seelen, 2006; Resnick, West, & Payne, 2008; Rogers et al., 2003). In accordance with these observations, we have drawn the following conclusions from our study:

a) Even though the American Psychiatric Association (2002) points towards a lack of cooperation with the evaluation as a criterion for suspected malingering, there is no scientific evidence that this strategy is typical of malingerers in the forensic field (Baer & Miller, 2002; Rogers et al., 2003). Likewise, the results with the clinical samples in this study also reflect total collaboration. Either way, failure to collaborate is exceptional, both in clinical and normative samples (Butcher et al., 1989; Greene, 2011). From a purely cognitive perspective, a non-collaborative attitude in clinical assessments may be related to defensiveness (defence towards an unwanted assessment), a lack of cognitive skills to be able to complete the assessment, or spurious reasons. However, in the forensic evaluation, a motivational aspect must be applied such that if the person evaluated expects to obtain a benefit from the evaluation, s/he will collaborate, but if s/he may be adversely affected, this may not be so (e.g., involuntary commitment evaluations).
b) The study of mean comparisons for the consistency measures showed that the responses from the populations with chronic

pain are more acquiescent but, just like the fibromyalgia group, with a mean within normality, that is, responses that are not biased towards true or false, and for which neither group follows randomised response patterns. However, the case study revealed that acquiescence is over-diagnosed (with the decision criterion ≥ 13, in the sense of responses biased towards true) in populations with chronic pain and fibromyalgia. This over-diagnosis, as well as being significant, is a large effect size (*OR* > 4.25). Likewise, VRIN also over-diagnoses inconsistency linked to random responses in populations with chronic pain and fibromyalgia, with a moderate (2.47 < *OR* < 4.25) and large effect size (*OR* > 4.25), respectively. Equally, the |F-Fb| index over-diagnoses, with a large effect size, inconsistency (random responses) in the population of patients with fibromyalgia. These inconsistency indicators must be taken with caution in the forensic evaluation, as they are interpreted as signs of malingering due to their association with hypotheses such as a lack of cooperation (American Psychiatric Association, 2000); but the over-diagnoses in these populations make it necessary to take into account other alternative hypotheses to malingering that are characteristic of clinical cases (e.g., clinically too confused, noncompliance) (Greene, 2008). The rate of false negatives linked to these inconsistency indicators must be added to this over-diagnosis. Hence, 14.9% of the randomly answered protocols lead to valid scores in VRIN (r < 13; Greene, 2011) in so far as there is no scientific evidence to support inconsistency in the responses of malingerers in the forensic field (Baer & Miller, 2002; Rogers et al., 2003). Thus, the invalidation of a protocol due to inconsistency in these clinical populations is not advisable since, in addition to the fact that it is not characteristic of malingerers (it would, in any case, be a positive criterion, that is, not malingering), it has a significantly higher likelihood of being found in clinical cases, therefore other alternative hypotheses must be studied.

c) The study of the validity of protocols revealed that in 6 of the malingering measuring scales and indexes, 3 original ones (F, Fb, and F-K) and 3 additional ones (Fp, Ds, and FBS), the patients with fibromyalgia reached higher means than the control group. Moreover, the Esd scale, indicative of defensiveness, also differentiated between both groups, with fibromyalgia patients obtaining lower means. Meanwhile, the group of patients with chronic pain reached means greater than the control group in the additional malingering measuring scales: Ds and FBS. In short, the 6 malingering indicators are biased against real fibromyalgia patients by 2 (none of the original ones) of the chronic pain ones. As a result, psychologists are biased in their judgments by these scales and indexes towards malingering, especially in patients with fibromyalgia. Yet, these results have no implications for clinical or forensic practice with $N = 1$ designs.

d) The case study, in a clinical setting, revealed that all the malingering scales and indexes would wrongly over-diagnose real cases of fibromyalgia as malingerers (invalid) with large effect sizes (*OR*s > 4.25), whereas the Fp and FBS would do so with the cases of chronic pain, also with large effect sizes. For the purposes of daily clinical practice, these results have totally different implications for fibromyalgia populations and chronic pain populations as in the original malingering scales and indexes –which are the ones normally used by clinical psychologists– they do not over-diagnose in clinical cases of chronic pain but they do in cases of fibromyalgia. Since up to 5 malingering indicators were recorded in cases in both populations, in order to ensure a subject is faking bad without the commission of false positives, more than 5 indicators (gold standard) of malingering (negative criteria, that is, that invalidate the protocol, in the model of Arce et al., 2006, 2009) will have to be verified in both populations.

e) The case study for the forensic setting (that is, contrasted with the normative population) warns that the 6 evaluation scales and indexes for malingering overestimate this, with very large effect sizes (*OR*s from 10.3 to 37.1, that is, from 1.29 to 2 standard deviates) in both populations. Briefly, forensic experts, when using these malingering indicators and decision criteria, are informed in a biased way. Since these scales and indexes are biased towards malingering and in the forensic evaluation a differential diagnosis of malingering is required (American Psychiatric Association, 2000), the malingering scales and indexes lead forensic psychologists to make mistakes in their classification of true fibromyalgia and chronic pain patients as malingerers (false positives). The results do not allow us to establish a forensic gold standard that will enable the identification of true positives without the commission of false positives, as we have recorded real cases classified as malingerers in both populations for all the malingering measuring scales and indexes.

f) The overestimation of invalidity of the response protocols, both with the standards in a clinical setting and in the normative population, suggests that other alternative hypotheses to malingering must be studied, such as the presence of a severe psychopathology (e.g., Graham, 2006; Greene, 2011). The solutions proposed for the (re)classification as true negatives of the genuine cases classified as malingerers by the validity scales and indexes, are the clinical decision models for establishing malingering (Bianchini et al., 2005; Cunnien, 1997; Resnick et al., 2008) and the presence of positive criteria, that is, the presence of criteria that are not under the subject's voluntary control, or of indicators that are not found in cases of malingering (Arce et al., 2006, 2009). With the application of the MMPI-2, there is only room for the search for criteria that are not found in faking bad, which are the indicators of defensiveness. The base rate of defensiveness in evaluation contexts in which it is the hypothesis to be suspected (e.g., custody litigations, job applications) has been estimated at 25-30% (Baer & Miller, 2002; Fariña, Arce, & Sotelo, 2010), whereas among malingerers in the forensic evaluation of psychological damage, compensation for damages or disability (characteristic fields of forensic evaluation of chronic pain and fibromyalgia), this rate turned out to be zero (Arce et al., 2006, 2009). This is because psychopathology is inversely related to defensiveness (Greene, 2008) and because malingerers do not bias their responses towards social desirability or hide symptoms, but rather quite the opposite, they subtract symptoms or attribute positive characteristics that they do not have (Arce et al., 2006, 2009, Fariña et al., 2010). The results only support differences between means in the Esd scale, which is lower for fibromyalgia patients than in the control group, that is, biased against defensiveness. However, in the classification of defensiveness, the L, Wsd, Mp, and Odecep scales significantly classified –with large (*OR* > 4.25) effect sizes– the protocols of fibromyalgia and chronic pain patients above the normative population. These scales share the impression management measurement (Paulhus, 1991; Strong, Greene, Hope, Johnston, & Olesen, 1999), which is characterised because the people evaluated consciously create a favourable impression of themselves in the evaluator. Cognitively, impression management is contrary to malingering strategies (for the strategies followed by malingerers, see Rogers, 2008), thereby a malingerer is not expected to consciously handle the impression to be caused so that it will be positive. The presence of an indicator of conscious self-favourable reporting enables us to (re)classify 79.5% of the protocols of patients with chronic pain and 44.4% of those with fibromyalgia as true negatives (that is, genuine cases). On the other hand, the Esd, K and S scales are utterly insensitive to the classification of fibromyalgia and chronic pain patients as making self-favourable reports. These are related to self-deceptive positivity (Paulhus, 1991; Strong et al., 1999), an unconscious distortion of the degree of adjustment, mental health, optimism and self-esteem, which detection would indicate that it is a true positive.

References

Aguerrevere, L. E., Greve, K. W., Bianchini, K. J., & Meyers, J. E. (2008). Detecting malingering in traumatic brain injury and chronic pain with an abbreviated version of the Meyers index for the MMPI-2. *Archives of Clinical Neuropsychology, 23*, 831-838. Doi: 10.1016/j.acn.2008.06.008

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Arce, R., Fariña, F., Carballal, A., & Novo, M. (2006). Evaluación del daño moral en accidentes de tráfico: Desarrollo y validación de un protocolo para la detección de la simulación [Evaluating psychological injury in motor vehicle accidents, MVA: Development and validation of a protocol for detecting simulation]. *Psicothema*, *18*, 278-283.

Arce, R., Fariña, F., Carballal, A., & Novo, M. (2009). Creación y validación de un protocolo de evaluación forense de las secuelas psicológicas de la violencia de género [Creation and validation of a forensic protocol to assess psychological harm in battered women]. *Psicothema*, *21*, 241-247.

Baer, R. A., & Miller, J. (2002). Underreporting of psychopathology on the MMPI-2: A meta-analytic review. *Psychological Assessment, 14*, 16-26. Doi: 10.1037/1040-3590.14.1.16

Bellato, E., Marini, E., Castoldi, F., Barbasetti, N., Mattei, L., Bonasia, D. E., & Blonna, D. (2012). Fibromyalgia syndrome: Etiology, pathogenesis, diagnosis, and treatment. *Pain Research and Treatment, .* Doi: 10.1155/2012/426130

Bianchini, K. J., Greve, K. W., & Glynn, G. (2005). On the diagnosis of malingered pain-related disability: Lessons from cognitive malingering research. *The Spine Journal, 5*, 404-417. Doi: 10.1016/j.spinee.2004.11.016

Branco, J. C., Bannwarth, B., Failde, I., Abello-Carbonell, J., Blotman, F., Spaeth, M., … Matucci-Cerinic, M. (2010). Prevalence of fibromyalgia: A survey in five European countries. *Seminars in Arthritis and Rheumatism, 39*, 448-453. Doi: 10.1016/j.semarthrit.2008.12.003

Breivik, H., Collet, B., Ventafridda, V., Cohen, R., & Gallacher, D. (2006). Survey of chronic pain in Europe: Prevalence, impact on daily life, and treatment. European *Journal of Pain, 10*, 287-333. Doi: 10.1016/j.ejpain.2005.06.009

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A. M., & Kaemmer, B. (1989). *MMPI-2: Manual for administration and scoring.* Minneapolis, MN: University of Minnesota Press.

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine, 19*, 3127-3131. Doi: 10.1002/1097-0258(20001130)

Cunnien, A. J. (1997). Psychiatric and medical syndromes associated with deception. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 23-46). New York: Guilford Press.

Fariña, F., Arce, R., & Sotelo, A. (2010). ¿Es efectivo el estudio psicométrico estándar del peritaje del estado clínico y de la disimulación en progenitores en litigio por la guarda y custodia de menores? [Is the standard psychometric forensic evaluation of the mental health and faking good of the partners litigating for the child custody effective?]. *Revista Iberoamericana de Psicología y Salud, 1*, 65-79.

Friedman, A. F., Lewak, R., Nichols, D. S., & Webb, J. T. (2001). *Psychological assessment with the MMPI-2*. Mahwah, NJ: LEA.

Gaskin, D. J., & Richard, P. (2012). The economic costs of pain in the United States. *The Journal of Pain, 13*, 715-24. Doi: 10.1016/j.jpain.2012.03.009

Gatchel, R. J. (2004). Comorbidity of chronic pain and mental health disorders: The biopsychosocial perspective. *American Psychologist, 59*, 795-805. Doi: 10.1037/0003-066X.59.8.795

Graham, J. R. (2006). *MMPI-2: Assessing personality and psychopathology* (4th ed.). New York, NY: Oxford University Press.

Greene, R. L. (2008). Malingering and defensiveness on the MMPI-2. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed., pp. 159-181). New York, NY: The Guilford Press.

Greene, R. L. (2011). *The MMPI-2/MMPI-2-RF: An interpretive manual* (3rd ed.). Boston, MA: Allyn & Bacon.

Greve, K. W., Ord, J. S., Bianchini, K. J., & Curtis, K. L. (2009). Prevalence of malingering in patients with chronic pain referred for psychologic evaluation in a medico-legal context. *Archives of physical medicine and rehabilitation, 90*, 1117-1126. Doi: 10.1016/j.apmr.2009.01.018

Hathaway, S. R., & McKinley, J. C. (1999). *Inventario Multifásico de Personalidad de Minnesota-2. Manual.* [The Minnesota Multiphasic Personality Inventory-2. Manual]. Madrid, Spain: TEA Ediciones.

Huss, M. T. (2009). *Forensic psychology: Research, clinical practice, and applications.* Chichester, UK: Wiley-Blackwell.

Merskey, H., & Bogduk, N. (Eds.) (1994). *Classification of chronic pain: Descriptions of chronic pain syndromes and definitions of pain terms* (2th ed.). Seattle, WA: IASP Press.

Ilgen M. A., Zivin, K., McCammon, R. J., & Valenstein, M. (2008). Pain and suicidal thoughts, plans and attempts in the United States. *General Hospital Psychiatry, 30*, 521-527. Doi: 10.1016/j.genhosppsych.2008.09.003

Johannes, C. B., Le, T. K., Zhou, X., Johnston, J. A., & Dworkin, R. H. (2010). The prevalence of chronic pain in United States adults: Results of an internet-based survey. *The Journal of Pain, 11*, 1230-1239. Doi: 10.1016/j.jpain.2010.07.002

Meyers, J. E., Millis, S. R., & Volkert, K. (2002). A validity index for the MMPI-2. *Archives of Clinical Neuropsychology, 17*, 157-169. Doi: 10.1016/S0887-6177(00)00107-4

Palmer, A. (2011). *Análisis unifactorial de la variancia con SPSS y R* [One-way analysis of variance with SPSS and R]. Palma de Mallorca, Spain: Edicions de la Universitat de les Illes Balears.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.

Polusny, M. A., & Arbisi, P. A. (2006). Assessment of psychological distress and disability after sexual assault in adults. In G. Young, A. W. Kane, & K. Nicholson (Eds.), *Psychological knowledge in courts. PTSD, pain and TBI* (pp. 97-125). New York, NY: Springer.

Pope, K. S., Butcher, J. N., & Seelen, J. (2006). *The MMPI, MMPI-2, and MMPI-A in court. A practical guide for expert witnesses and attorneys*. Washington, DC: American Psychological Association.

Resnick, P. J., West., S., & Payne, J. W. (2008). Malingering of post-traumatic disorders. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 109 -127). New York, NY: Guilford.

Rogers, R. (2008). Researching response styles. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed., pp. 411-434). New York, NY: The Guilford Press.

Rogers, R. Sewell, K. W., Martin, M. A., & Vitacco, M. J. (2003). Detection of feigned mental disorders: A meta-analysis of the MMPI-2 and malingering. *Assessment*, *10*, 160-177. Doi: 10.1177/1073191103252349

Scascighini, L., Toma, V., Dober-Spielmann, S., & Sprott, H. (2008). Multidisciplinary treatment for chronic pain: A systematic review of interventions and outcomes. *Rheumatology, 47*, 670-678. Doi: 10.1093/rheumatology/ken021

Strong, D. R., Greene, R. L., Hope, C., Johnston, T., & Olesen, N. (1999). Taxometric analysis of impression management and self-deception in college student and personnel evaluation settings. *Journal of Personality Assessment, 78*, 161-175. Doi: 10.1207/S15327752JPA7801_10

Tecic, T., Lefering, R., Althaus, A., Rangger, C., & Neugebauer, E. (2013). Pain and quality of life 1 year after admission to the emergency department: Factors associated with pain. *European Journal of Trauma and Emergency Surgery*. Advance online publication. Doi: 10.1007/s00068-013-0271-9

Tsang, A., Von Korff, M., Lee, S., Alonso, J., Karam, E., Angermeyer, M. C., …Watanabe, M. (2008). Common chronic pain conditions in developed and developing countries: Gender and age differences and comorbidity with depression-anxiety disorders. *The Journal of Pain, 9*, 883-891. Doi: 10.1016/j.jpain.2008.05.005

Van Leeuwen, M. T., Blyth, F. M., March, L. M., Nicholas, M. K., & Cousins, M. J. (2006). Chronic pain and reduced work effectiveness: The hidden cost to Australian employers. *European Journal of Pain, 10*, 161-166. Doi: 10.1016/j.ejpain.2005.02.007

Wierville, L. L. (2012). Fybromialgia: Diagnosing and managing a complex syndrome. *Journal of American Academy of Nurse Practitioners, 24*(4), 184-192. Doi: 10.1111/j.1745-7599.2011.00671.x

Winkelmann, A., Perrot, S., Schaefer, C., Ryan, K., Chandran, A., Sadosky, A., & Zlateva, G. (2011). Impact of fibromyalgia severity on health economic costs. *Applied Health Economics and Health Policy, 9*, 125-136. Doi: 10.2165/11535250-000000000-00000

Wolfe, F., Clauw, D. J., Fitzcharles, M., Goldenberg, D. L., Katz, R. S., Mease, P., …Yunus, M. B. (2010). The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. *Arthritis Care & Research, 62*, 600-610. Doi: 10.1002/acr.20140

Wolfe, F., Ross, K., Anderson. J., Russell, I. J., & Hebert, L. (1995). The prevalence and characteristics of fibromyalgia in the general population. *Arthritis and Rheumatism, 38*, 19-28. Doi: 10.1002/art.1780380104

Wolfe, F., Smythe, H. A., Yunus, M. B., Bennett, R. M., Bombardier, C., Goldenberg, D. L., …Sheon, R. P. (1990). The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. *Arthritis and Rheumatism, 33*, 160-172. Doi: 10.1002/art.1780330203

World Health Organization. (1992). *The ICD-10 classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines.* Geneva, Switzerland: Author.