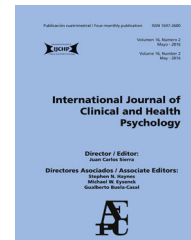




# International Journal of Clinical and Health Psychology

www.elsevier.es/ijchp



## ORIGINAL ARTICLE

# A meta-analytic review of the MMPI validity scales and indexes to detect defensiveness in custody evaluations



Francisca Fariña<sup>a</sup>, Laura Redondo<sup>a</sup>, Dolores Seijo<sup>b</sup>, Mercedes Novo<sup>b</sup>, Ramón Arce<sup>b,\*</sup>

<sup>a</sup> Universidad de Vigo, Spain

<sup>b</sup> Universidad de Santiago de Compostela, Spain

Received 17 December 2016; accepted 10 February 2017

Available online 21 April 2017

### KEYWORDS

Parent child custody disputes;  
Parenting capacity assessment;  
MMPI;  
Defensiveness;  
Meta-analysis

**Abstract** *Background/Objective:* In child custody disputes, one of the remit of the forensic psychologist is to evaluate parental attributes while suspecting defensiveness. The instrument of choice for undertaking this double task is the MMPI. *Method:* As to establish the state of the art on this, a meta-analysis was undertaken with a total of 32 primary studies from which 256 effect sizes were assessed. A meta-analysis was undertaken, effect sizes were corrected for sampling error and criterion unreliability. *Results:* The results revealed a positive, significant, large and generalizable mean true effect size for the *L*, *K*, *S* and *MP* scales, and the *L + K* and *L + K - F* indexes. The *Wsd* was positive, significant and large, but not generalizable. A negative and significant, but not generalizable mean true effect size was found for the *F* and generalizable for *F - K* index. The effect sizes for the *L*, *K*, *S* and *MP* scales, and the *L + K* and *L + K - F* indexes were equal. Both the gender of parents (father vs. mother) and the context of evaluation (parent child custody disputes vs. parenting capacity) were assessed as moderators. *Conclusions:* The results are discussed in relation to forensic practice.

© 2017 Asociación Española de Psicología Conductual. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### PALABRAS CLAVE

Disputa parental por la custodia;  
evaluación de la capacidad parental;  
MMPI;  
disimulación;

### Revisión meta-analítica de las escalas e índices de validez del MMPI para detectar disimulación en la evaluación de custodias

**Resumen** *Antecedentes/Objetivo:* En los casos de disputa por la custodia, el psicólogo forense tiene entre sus cometidos la evaluación de las competencias parentales, así como sospechar disimulación. Para esta doble tarea, el instrumento de referencia es el MMPI. *Método:* Para establecer el estado de la cuestión se llevó a cabo un meta-análisis encontrando 32 estudios primarios de los que se obtuvieron 256 tamaños del efecto. Los tamaños del efecto fueron

\* Autor para correspondencia. Facultade de Psicoloxía. Campus Vida, s/n. 15782 Santiago de Compostela, Spain.  
E-mail address: [ramon.arce@usc.es](mailto:ramon.arce@usc.es) (R. Arce).

separación;  
meta-análisis

corregidos por error de muestreo y falta de fiabilidad del criterio. *Resultados:* Los resultados mostraron un tamaño del efecto medio verdadero positivo, significativo, grande y generalizable para las escalas *L*, *K*, *S* y *MP*, y los índices *L+K* y *L+K-F*. Para *Wsd*, también resultó positivo, significativo y grande, pero no generalizable. Para *F* y el índice *F-K* fue negativo y significativo, pero no generalizable para *F* y generalizable para *F-K*. Los tamaños del efecto de las escalas *L*, *K*, *S* y *MP*, y los índices *L+K-F* y *L+K* resultaron ser iguales. Se estudiaron como moderadores el género del progenitor (padre vs. madre) y el contexto de evaluación (progenitores en disputa por la custodia de los hijos vs. evaluación de la capacidad parental). *Conclusiones:* Se discute la utilidad para la práctica forense de estos resultados.

© 2017 Asociación Española de Psicología Conductual. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Forensic psychological evaluation in child custody disputes is regulated by standards and guidelines established by an array of associations around the world such as the American Psychological Association (2010), the Association of Family and Conciliation Courts (Martindale, Martin, Austin, & Task Force Members, 2007), or the Spanish Psychological Association [Colegio Oficial de Psicólogos] (Chacón, García, García, Gómez, & Vázquez, 2009). Though these standards and guidelines may vary slightly, they all have common aims i.e., to determine the child's psychological best interests, to guide professionals in the evaluation of children, parents and the child-parent interaction in order to identify the child's psychological needs and parental attributes to find the best fit between child's needs and parental attributes.

The primary aim of these standards and guidelines is to evaluate parenting attributes in terms of the knowledge, abilities and skills required to effectively cater for the child's needs, and to detect deficits and psychopathology that may put the child at risk. Both separation and divorce are psychosocial stressors closely linked to clinical symptomatology (Amato & Keith, 1991; Cheng, Dunn, O'Connor, & Holding, 2006; Weaver & Schofield, 2015). Moreover, defensive responding should be suspected (Arce, Fariña, Seijo, & Novo, 2015; Bagby & Marshall, 2004; Strong, Greene, Hoppe, Johnston, & Olesen, 1999), affecting an estimated 30 to 40% of evaluations (Arce, Fariña, Seijo et al., 2015; Baer & Miller, 2002; Fariña, Arce, & Sotelo, 2010; Strong et al., 1999).

To evaluate parental attributes, psychologists employ psychological tests, clinical interviews, behavioural observation (e.g., parent-child interactions), home visits, and collateral contacts (e.g., extended family). The clinical interview, in particular the forensic-clinical interview (Vilariño, Arce, & Fariña, 2013), and psychological tests, primarily the MMPI-2, the psychometric instrument most extensively used worldwide for forensic psychological assessment which has been translated into over 40 languages (Archer, Buffington-Vollum, Stredny, & Handel, 2006; Fariña, Arce, Vilariño, & Novo, 2014; Rogers, Sewell, Martin, & Vitacco, 2003), and used in over 90% of parental evaluations in child custody disputes (Ackerman & Pritzel, 2011; Arch, Jarne, Peró, & Guàrdia, 2011; Fariña et al., 2010), serve to evaluate both parental attributes and defensiveness. When defensiveness or malingering is suspected in the assessment of psychological and personal attributes, the combination of

clinical interview and psychometric evaluation is required (Arce, Fariña, & Vilariño, 2015; Graham, 2011). The MMPI-2 includes the *L*, *F* and *K* original validity scales. The *L* scale was designed to detect the deliberate and overt acknowledgment of uncommon virtues. The *F* scale was initially designed to detect random responding, but empirical research has shown that *F* was also sensitive to intentional attempts to portray one's own negative image. The *K* scale was as a subtle indicator (*F* and *L* are more obvious) of attempts to exaggerate psychopathology and to appear in a very unfavourable way (low scores), or deny psychopathology and to present oneself in a favourable way (high scores). Due to the Restructured Form of the MMPI-2, the MMPI-2-RF, the original validity scales for defensiveness, *L* and *K*, were also reformulated as *L-r* and *K-r*. The *L-r* scale consisted of 14 items, sharing 11 with the original *L* scale and adding three additional items, while the *K-r* scale consisted of 14 items from the original *K* scale (16 were deleted and the scoring direction for one was reversed). No evidence or rationality was provided to support actions in both scales (Greene, 2011). Moreover, the MMPI-2 contains additional scales for measuring defensiveness: Positive Malingering Scale (*MP*); Wiggins's Social Desirability Scale (*Wsd*); Edward's Social Desirability (*Esd*); *O-S* Scale (Obvious-Subtle); Test Taking Defensiveness Scale (*Tt*); Other Deception (*Od*); Superlative Scale (*S*); and the Positive Mental Health Scale (*PMH4*). The *S* scale measures the denial of psychological problems and moral shortcomings, as well as the endorsement of unrealistically positive personal and interpersonal attributes; the *Wsd*, *Od*, *Mp*, *Esd* and *Tt* scales measure social desirability (*Od* is an upgrade version of *MP* and *Wsd*); the *PMH-4*, the denial of various forms of psychological maladjustment; and the *O-S* subscale reports underreporting when subtle items are endorsed more than obvious (negative scores). Finally, three indices, *F-K*, *L+K* and *L+K-F*, were related with defensiveness (Baer & Miller, 2002; Graham, 2011; Lanyon & Lutz, 1984; Posthuma & Harper, 1998).

As for distortions related to defensiveness, two response patterns have been observed i.e., self-deception (*SD*) and positive impression management (*IM*) according to whether the individual is conscious or not of manipulating them (Paulhus, 1984). These response patterns have different legal implications since the *IM* entails a deliberate attempt (volitional component) to wilfully deceive in spite of being fully aware it is illegal (intent, cognitive

component), whereas the SD implies a unreal (volitional component), but honest (cognitive component) responses (Fariña et al., 2010). In the context of forensic evaluation of custody disputes both types of response patterns can be expected. Thus, the SD would be a stable trait of a subject generalizable to all measurement contexts, whereas IM is characteristic to this measurement context, involving approximately 40% of the population under evaluation (Arce, Fariña, & Vilariño, 2015). The MMPI *Wsd*, *L*, *Od* and *MP* scales assessed IM, and the *Esd*, *K*, *S* and *PMH4* scales assessed the SD (Arce, Fariña, & Vilariño, 2015; Bagby & Marshall, 2004; Greene, 2011; Strong et al., 1999; Strong, Greene, & Kordinak, 2002).

These standards and guidelines are enshrined in the professional practice of psychologists (Ackerman & Pritzl, 2011; Arch et al., 2011; Archer & Wygant, 2012; Bow & Quinnell, 2001). Moreover, judges and the courts classify, according to psychological reports, parental attributes as incapacitating characteristics for child custody (e.g., drug addiction, negligence), negative for custody (e.g., parental incompetence, mental disorders), and positive (e.g., parental abilities to cater for the child's needs) (Arce, Fariña, & Seijo, 2005).

Research on defensiveness evaluation has focused mainly on two contexts, personnel selection (Strong et al., 2002) and child custody disputes (Strong et al., 1999), suggesting that these MMPI scales and indexes might perform in a different manner across different assessment contexts (Bagby & Marshall, 2004). Thus, Baer and Miller's (2002) meta-analysis has shown that the mean effect size of MMPI-2 traditional and supplementary indices of underreporting was higher for job applicants ( $d = 1.55$ ) than for child custody litigants ( $d = 0.99$ ). Nevertheless, these and other results of this meta-analysis published in the reference journal of psychological evaluation *Psychological Assessment* are not valid since the results were incorrectly computed, given that the mean effect sizes were not corrected for sampling error (nor corrected for criterion unreliability); notwithstanding, these are used worldwide in forensic setting as valid assessments. For example, the unweighted overall mean effect size (non-corrected for sampling error) reported for the *K* scale was  $d = 1.13$ , whereas the corrected sampling error was  $d = 1.47$ . The gap between corrected and uncorrected effect sizes,  $d = 0.34$ , implies that the *K* scale classified correctly 16.8% ( $r = .16$ ) more defensiveness than Baer and Miller's results. Moreover, all the scales and indexes were inappropriately mixed in a global effect size. Additionally, the systematic conclusions of the literature, based mainly on the classification accuracy or incremental validity, concerning the superiority of certain scales and indexes over others (e.g., Baer & Miller, 2002; Bagby, Nicholson, Buis, Radovanovic, & Fildler, 1999; Butcher, 1997; Carr, Moretti, & Cue, 2005), are not statistically supported. In fact, the application of statistical tools to the data provided by Baer and Miller about the classification accuracy (e.g., the computed 95% CIs, meaning there are no mean differences), does not corroborate the superiority. Bearing in mind these gaps, the time lapsed from the last review of the literature, 2002, a meta-analysis was undertaken to determine the mean true effect size for each of the MMPI scales and indexes of defensiveness and to assess their utility in forensic practice for evaluating parents involved child custody litigation.

## Method

### Search of studies

The search strategy was aimed at detecting studies evaluating parents in child custody disputes using any of the family instruments on the MMPI: MMPI (Hathaway & McKinley, 1940), MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989; Butcher et al., 2001), or MMPI-2-RF (Ben-Porath & Tellegen, 2008/2011). The initial search was intended to locate previous systematic reviews and meta-analyses (i.e., Baer & Miller, 2002; Cooke, 2010; Roma et al., 2014) from which to draw a list of reviewed articles and descriptors for subsequent searches (i.e., MMPI, response styles, validity scales, child custody litigation, child custody dispute, underreporting, child custody evaluations, defensiveness, faking good, parental capacity assessment). These descriptors were used to design search algorithms applied to leading scientific databases: Web of Science, Scopus, PsycInfo and Proquest Dissertation & Theses. Finally, a search was performed in the metasearch-engine 'google scholar'. The search was performed in July 2016. These systems yielded a total of 4,310 publications that were applied the following inclusion criteria: a) participants were parents involved in child custody litigation proceedings; b) empirical studies reporting effect size or sufficient data for its computation (when this contingency was not met but the others were, the authors were contacted to obtain the data); and c) parents were evaluated using the family instrument of the MMPI. Studies in which subjects had been instructed to respond (simulation research) like parents (most instructed students to behave as litigant parents) in custody disputes were excluded because the results of these simulation studies enjoy high face validity, while external validity remains untested (Konecni & Ebbesen, 1992), and real subjects and those in feigning conditions (simulation research) provide significant different results (Amado, Arce, & Fariña, 2015; Amado, Arce, Fariña, & Vilariño, 2016) and have been found to perform different tasks (Fariña, Arce, & Real, 1994). All of the studies published meeting these criteria were included.

After screening, a total of 32 primary studies (21 articles in journals, 2 unpublished studies and 9 doctoral theses) were selected in which the effect sizes of one or more scales measuring defensiveness on the MMPI were obtained. Samples duplicity was controlled, 256 effect sizes were obtained: 67 for the *L* scale, 65 for the *K*, 51 for the *F*, 19 for the *S*, 15 for the *Wsd*, 9 for the *MP*, and 1 for the *Esd* and *Od*, 10 for the *F-K* index, 6 for *L+K*, and 12 for *L+K-F*.

### Coding of primary studies

In order to proceed with the meta-analysis the following data from the studies was codified: a) article reference; b) article source (paper, unpublished data, doctoral thesis); c) sample characteristics (i.e., size, gender); d) design characteristics (evaluation of custody disputes or evaluation of parenting abilities, level of conflict, reports of sexual abuse, physical abuse, negligence or abandonment, family violence, alienation, descriptor favourable or unfavourable); e) the statistics required for computing the effect size. This task was carried out separately by two researchers, with

total concordance (Cohen's  $k = 1$ ) in the coding. The characteristics of the primary studies included in this review are shown in [Appendix 1](#).

## Data analysis

The effect size of the primary studies was obtained with Cohen's  $d$  since the means were systematically reported (no study was correlational) of groups in the custody dispute evaluation condition (i.e., the target population of this meta-analysis). Primary studies compared independent groups of cases-controls, multiples groups, repeated measures, and the experimental group with a test value. Moreover, some studies were found to report their results in raw scores, but others used T scores. Similarly, different versions of the MMPI i.e., MMPI-1, MMPI-2 and MMPI-2-RF were used. When the results were reported in T values, the effect size was obtained as for a single sample using the formula of Glass ([Glass, 1976](#); [Glass, McGraw, & Smith, 1981](#)), where the measure and standard deviation of the 'test value' were 50 and 10, respectively. The use of the normative group was preferred to the particular study control group as the idiosyncrasies of this specific control group were controlled by taking the normative group that represents the general population ([Hunter & Schmidt, 2015](#)). When the results were reported in raw scores, these were transformed into T scores by using the means and standard deviations of the normative population in the MMPI manuals. For the scales and indexes not included in the MMPI manuals, the test value for computing  $d$  was the mean cutting scores for coached participants to be applied to test takers involved in legal proceedings ([Baer & Miller, 2002](#)), and the standard deviation for the experimental group. Having computed the effects sizes the meta-analysis was performed and corrected for sampling error and criterion unreliability (procedure of Hunter and Schmidt, 2015), for each of the scales and measurement indexes of defensiveness. [Amado et al. \(2015\)](#) have shown the utility of three statistics for forensic practice: U1, Binomial Effect Size Display (BESD), and Probability Superiority (PS). Thus, these were computed to derive the measures on the effectiveness of the scales and indexes for detecting defensiveness over the natural tendency for defensiveness i.e., responding defensively even with nothing to hide or to give a positive presentation ([Osuna, López-Martínez, Arce, & Vázquez, 2015](#); [Palmer, Borrás, Pérez-Pareja, Sesé, & Vilariño, 2013](#)).

## Criterion reliability

Criterion reliability for the original validity scales ([Table 1](#)) assessing the MMPI and the MMPI-2 (the original scales remain in both versions with the exception of 4 items on the  $F$  scale that were eliminated from version 2 for being offensive) were taken from a meta-analytical review on the reliability of the  $L$  (70 studies),  $F$  (70 studies), and  $K$  scales (71 studies) of [Hunsley, Hanson, and Parkeret \(1988\)](#); and the MMPI-2-RF *Manual for administration, scoring, and interpretation* ([Ben-Porath & Tellegen, 2008/2011](#)). As for the additional defensiveness scales, the reliability of the  $S$  Scale (Superlative) was taken from its creators ([Butcher & Han, 1995](#)), the  $Wsd$  was taken from the only study

**Table 1** Criterion reliability.

Scale/Index	$\alpha^1$	$\alpha^2$
L	.77	.70
F	.77	.61
K	.82	.68
S	.86	—
Wsd	.51	—
MP	.70	—
F-K	.85	—
L + K	.84	—
L + K-F	.86	—

Note.

<sup>1</sup> MMPI/MMPI-2.

<sup>2</sup> MMPI-2-RF, — Scale/index not available at this instrument.

reporting it ([Paulhus, 1984](#)), and for the  $MP$ , as no study was found reporting reliability, it was calculated on the basis of 892 normative subjects evaluated under standard response conditions (control group in studies) from the Forensic Psychology Institute of the University of Santiago de Compostela (Spain). No meta-analysis calculated  $Esd$  and  $Od$  scales as only one study was identified. Finally, the reliability of the composites (i.e.,  $F-K$ ,  $L + K$ ,  $L + K-F$ ) was calculated using the formula of [Mosier \(1943\)](#).

## Results

### Study of outliers

Initially outliers [ $\pm 1.5 \cdot IQR$ ] in each of the scales and indexes of defensiveness were eliminated. This tool found 2 (3.8%) outliers in 53 effect sizes in the  $F$  scale; 2 (22%) of 9 in  $MP$ ; 1 of 15 (6.7%) in  $Wsd$ ; 1 of 10 (10%) in  $F-K$  index; and 4 (33%) of 12 in  $L + K-F$ . As this technique eliminated many effect sizes of the  $MP$  Scale, the  $F-K$  index and the  $L + K-F$  index ( $\geq 10\%$ , [De Dreu & Weingart, 2003](#); [Hunter & Schmidt, 2015](#); [Tukey, 1960](#)), it is likely they were moderators, not outliers. Moreover, the elimination should not account for an excessive percentage of evaluated subjects ( $N$ ), which would substantially affect the  $MP$  Scale with the loss of 56.34% of participants. Thus, a second screening with the criterion  $M \pm 2SD$  was performed, being the results generalizable to 96% of the future samples, with 1 outlier, the same as with criterion  $\pm 1.5 \cdot IQR$ , in  $F-K$  and in  $L + K-F$ , and none in  $MP$ . Hence, the meta-analysis calculated  $MP$  and  $L + K-F$  with the effect sizes within the region  $M \pm 2SD$ . Nonetheless, given that the elimination of outliers reduces the variance, and in turn the effect size, for the  $L$  and  $Wsd$  scales the mean true effect sizes of the samples obtained with the interquartile range (IQR) criterion were computed. The results showed equivalent results for  $L + K-F$  ( $\delta = 1.24$  and  $1.20$  for the criterion  $\pm 1.5 \cdot IQR$  and  $M \pm 2SD$ , respectively), and similar for  $MP$  (a positive, significant and generalizable mean true effect size), but different in size (medium,  $\delta = 0.48$ , with the  $\pm 1.5 \cdot IQR$  criterion; and large,  $\delta = 1.08$ , with the  $M \pm 2SD$  criterion).

**Table 2** Results of the meta-analyses between parents in child custody disputes and the normative population.

Scale/Index	<i>k</i>	<i>N</i>	<i>d<sub>w</sub></i>	<i>SD<sub>d</sub></i>	<i>SD<sub>pre</sub></i>	<i>SD<sub>res</sub></i>	$\delta$	<i>SD<sub>δ</sub></i>	%Var	95% CI <sub><i>d</i></sub>	80% CI <sub><math>\delta</math></sub>
L+	67	10642	0.87	0.37	0.16	0.34	0.99	0.38	19.47	0.83, 0.91	0.50, 1.49
L++	58	9530	0.93	0.35	0.16	0.31	1.06	0.35	21.79	0.89, 0.97	0.60, 1.52
K+	65	10154	0.82	0.27	0.16	0.22	0.91	0.24	36.46	0.78, 0.86	0.60, 1.23
K++	57	9074	0.80	0.28	0.16	0.22	0.89	0.25	34.77	0.76, 0.84	0.57, 1.21
F+	51	9212	-0.23	0.30	0.15	0.26	-0.27	0.30	23.67	-0.27, -0.19	-0.66, 0.13
F++	43	8132	-0.27	0.29	0.14	0.25	-0.31	0.29	24.94	-0.31, -0.23	-0.68, 0.06
S+++	19	3263	0.85	0.29	0.16	0.24	0.91	0.26	29.85	0.78, 0.92	0.57, 1.25
Wsd+++	14	1244	0.78	0.65	0.22	0.61	1.10	0.86	11.51	0.66, 0.90	-0.01, 2.20
MP+++	9	1088	0.91	0.45	0.19	0.41	1.08	0.49	17.90	0.79, 1.03	0.45, 1.71
F-K+++	9	673	-0.60	0.30	0.23	0.19	-0.65	0.21	59.19	-0.76, -0.44	-0.92, -0.38
L+K+++	6	188	0.76	0.08	0.37	0	0.83	0	100	0.47, 1.05	0.83
L+K-F+++	11	339	1.11	0.56	0.39	0.40	1.20	0.43	48.64	0.87, 1.34	0.64, 1.75

*Note.* +studies from original validity scales of MMPI, MMPI-2 and reformulated scales of MMPI-2-RF; ++studies from original validity scales of MMPI-2; +++ studies from the additional validity scales of MMPI-2; *k* = number of studies; *N* = total sample size; *d<sub>w</sub>* = effect size weighted for sample size; *SD<sub>d</sub>* = observed standard deviation of *d*; *SD<sub>pre</sub>* = standard deviation of observed correlations predicted from all artifacts; *SD<sub>res</sub>* = standard deviation of observed correlations after removal of variance due to all artifacts;  $\delta$  = effect size corrected for criterion unreliability; *SD<sub>δ</sub>* = standard deviation of  $\delta$ ; %Var = variance accounted for by artifactual errors; 95% CI<sub>*d*</sub> = 95% confidence interval for *d*; 80% CI <sub>$\delta$</sub>  = 80% credibility interval for  $\delta$ .

**Defensiveness scales and indexes**

The results of the effect size calculated for each scale and index, the total number of effect sizes obtained (*k*); sample size (*N*); the uncorrected effect size weighted by sample size (*d<sub>w</sub>*), and the standard deviation (*SD<sub>d</sub>*); the effect size corrected for criterion unreliability ( $\delta$ ); the percentage of variance explained by the artifactual errors (%Var), 95% confidence intervals, and 80% credibility interval (when both intervals have no zero, it indicated the estimated effect size was significant and generalizable, respectively), are shown in Table 2. The results for the *L*, *K*, *S*, *MP* scales and *L+K* and *L+K-F* indexes reveal a significant (when the confidence interval has no zero, indicating the effect size was significant), positive (between child custody litigants and defensiveness), generalizable (the credibility interval had no zero, indicating the effect size was generalizable to 90% of other samples), and large ( $\delta > 0.80$ ) mean true effect size ( $\delta$ ). Similar results i.e., a significant, positive and large mean true effect size, was found for *Wsd*, but it was not generalizable. As for the *F* scale and the *F-K* index, a significant and negative mean true effect size was found, small ( $0.20 > \delta < 0.50$ ) and not generalizable (credibility interval had zero) for *F* scale, and medium ( $0.50 > \delta < 0.80$ ) and generalizable for *F-K* index. As only one effect size was found for the *Es<sub>d</sub>* and *Od* scales, the mean true effect sizes could not be estimated, the uncorrected effect size were 1.24 and 1.38, respectively.

Comparatively, the mean true effect sizes in the scale and indexes with a positive and generalizable relationship with defensiveness for *L*,  $\delta = 0.99$ , 95%CI [0.95, 1.03], *K*,  $\delta = 0.91$ , 95% CI [0.87, 0.95], *S*,  $\delta = 0.91$ , 95%CI [0.84, 0.98], and *MP*,  $\delta = 1.08$ , 95%CI [0.95, 1.21], scales and for *L+K* index,  $\delta = 0.83$ , 95%CI [0.53, 1.13] and *L+K-F*,  $\delta = 1.20$ , 95%CI [0.97, 1.43], indexes were equal (if the 95% CIs for  $\delta$  overlap, it indicates no mean differences).

In terms of utility for forensic practice (Table 3), the results revealed that the *L* scale classified as defensiveness

**Table 3** Practical utility indicators.

Scale/Index	U1	<i>r</i>	PS
L	.55	.44	.75
K	.51	.41	.74
S	.51	.41	.74
MP	.58	.47	.77
L+K	.48	.38	.72
L+K-F	.62	.51	.80

*Note.* Only for scales and indexes with generalizable effects sizes; U1 = Cohen's U1 statistic; *r* = correlation for BESD compute; PS = probability of superiority.

44.4% more (BESD) protocols in the population of custody disputes than in the normative group; 55.0% (U1 = .55) of the area covering both populations (normative and custody disputes) did not overlap i.e., they were totally independent; and a probability of .75 (PS) that subjects in custody disputes score higher on the *L* scale than the population normative. In *K*, *S*, *MP*, *L+K* and *L+K-F*, the defensiveness classification rate in the custody disputes population was, respectively, 41.4, 41.4, 47.6, 38.4, and 51.6%, more than in the normative population; the distributions for the normative population and custody disputes were totally independent in 51.9, 51.9, 58.2, 48.7 and 62.2%; and the probability of superiority was .74, .74, .77, .72 and .80, that is, these would be the probabilities for the population under custody disputes scoring higher on these scales than the normative population.

The 75% rule (Hunter & Schmidt, 2015) warrants the study of moderators, except for the *L+K* index (%Var = 100, indicating the primary studies were not entirely randomly distributed, and *N* [*<*400] was insufficient for the study of moderators). The literature suggests the parent's gender could play a relevant role in defensiveness (Roma et al., 2014), as well as the situational factor (parent child custody disputes [PCCDs] vs. parenting capacity assessment in

child protection cases [PCA-CPCs]) (Carr et al., 2005). Other moderators could not be analysed due to insufficient effect sizes or  $N$ s. A last moderator, the version of the MMPI i.e., the original MMPI, the MMPI-2 and the MMPI-2-RF, could not be examined as the studies with the original MMPI and the MMPI-2-RF, are only available for the original validity scales, and were insufficient ( $N < 400$  and/or  $k \leq 3$ ). Thus, results were computed for all versions and only for the MMPI-2 (see Table 2).

### Gender as a moderator

The meta-analysis on the gender of the litigator as a moderator (Table 4), in line with the general meta-analysis, showed for the  $L$ ,  $K$  and  $S$  scales a significant, positive, generalizable and large (or nearly large) mean true effect size for both fathers and mothers. The mean true effect sizes for fathers and mothers were equal (Table 4) in the three scales (the 95% CIs for  $\delta$  overlapped).

In the  $F$  scale, as in the general meta-analysis, a significant and negative mean true effect size for both fathers and mothers was observed. Nevertheless, this negative mean true effect size may be generalised to other samples for fathers, but not for populations of mothers.

The meta-analysis for the  $Wsd$  and  $MP$  scales, and  $F-K$  and  $L+K-F$  indexes are not shown as  $k (\leq 3)$  and/or  $N (< 400)$  were too low to guarantee stability in sampling estimates (Hunter & Schmidt, 2015), which were in line with the general meta-analysis and equal to gender.

### The context of disputes as moderator

The context of evaluation (parent child custody disputes [PCCDs] vs. parenting capacity assessment in child protection cases [PCA-CPCs]) appears in primary studies as a potential moderator of differences in the evaluation of parents/caregivers in custody disputes. To this effect, the  $L$ ,  $K$ ,  $F$  and  $S$  scales were evaluated. The results (Table 5) reveal a positive, significant, generalizable and large mean true effect size for the  $L$  scale for both parents in custody disputes and for parents in PCA-CPCs. Notwithstanding, the effect size was significantly larger in PCA-CPCs,  $\delta = 1.41$ , 95%CI [1.22, 1.60], than in PCCDs,  $\delta = 0.97$ , 95%CI [0.93, 1.01]. As for the  $K$  scale, the results of the meta-analysis showed a positive, significant, generalizable and large mean true effect size for PCCs and of a small size for PCA-CPCs. In contrast to the  $L$ , the effect size for the  $K$  scale was significantly larger in PCCDs,  $\delta = 0.95$ , 95%CI [0.91, 0.99], than in PCA-CPCs,  $\delta = 0.28$ , 95%CI [0.11, 0.45]. In the  $F$  scale the results show an inverse relationship: a *negative*, significant, generalizable and small mean true effect size for PCCDs, and *positive*, significant, generalizable and large mean true effect size in PCA-CPCs. Finally, the results for the  $S$  scale showed a positive, significant, generalizable and large mean true effect size in PCCDs, and a non-significant mean true effect size in PCA-CPCs.

### Discussion

The following conclusions may be derived from the results of this study. First, none of the scales or indexes detected

totally defensiveness. Thus, no indicator of defensiveness was a fully efficacious detector on its own and had to be used in combination or accumulatively to enhance efficacy. Second, in line with the original models, the  $L$ ,  $K$ ,  $S$ ,  $MP$ ,  $Wsd$ ,  $Od$  and  $Esd$  scales and the  $L+K$  and  $L+K-F$  indexes were positively related to defensiveness, whereas the  $F$  scale and the  $F-K$  index were negatively related. Third, the results undermine the findings of studies claiming the superiority of scale over the other on the basis of simply observing the means and classification accuracy (e.g., Bagby et al., 1999; Butcher, 1997; Carr et al., 2005), MMPI reference manuals (Graham, 2011; Greene, 2011), and other meta-analysis (Baer & Miller, 2002), which should be revised. However, the results for  $Wsd$ ,  $F$ , and  $F-K$  were not generalizable i.e., they did not consistently detect inter-study defensiveness. Likewise, the findings of studies reporting the validity of these scales and indexes as detectors of defensiveness should also be reviewed (e.g., Baer & Miller, 2002; Baer, Wetter, & Berry, 1992, 1995; Baer, Wetter, Nichols, Greene, & Berry, 1995; Bagby et al., 1997). Fourth, the  $L$ ,  $K$ ,  $S$  and  $MP$  scales, and  $L+K$  and  $L+K-F$  indexes, whose efficacy in detecting defensiveness was similar, were found to be the best detectors. Fifth, the scales and indexes with generalizable results (i.e.,  $L$ ,  $K$ ,  $S$ ,  $MP$ ,  $L+K$ ,  $L+K-F$ ) add to the classification baseline of defensiveness (normative group), with approximately 40 to 50% more cases; the discrimination rate (independence distributions) between protocols of populations in custody disputes and the normative population (honest response) ranged from 50 to 60%; and the probability that parents in custody disputes obtained higher scores on the scales and indexes with generalizable results ranged approximately from .75 to .80. Sixth, the defensiveness attitudes of men and women in the evaluation of child custody disputes were similar, which disagrees with the findings of studies claiming different attitudes towards the evaluation (defensiveness) in men and women in child custody disputes (Roma et al., 2014). Seventh,  $L$  was a significantly better detector of defensiveness in the PCA-CPC than in PCCD evaluation context, and both  $K$  and  $S$  were in PCCDs. Surprisingly, the  $F$  scale was related, in line with the model (high scores suspect potential feigning), negatively (between parents in child custody disputes and defensiveness) in PCCDs, but positively related in PCA-CPCs (contrary to the model). In short, attitudes towards the evaluation (defensiveness) were measured according the evaluation context i.e., PCCDs vs. PCA-CPCs.

This meta-analysis has several limitations that should be borne in mind such as: a) the results were obtained from studies on parent child custody disputes or parenting capacity assessment in child protection cases, and caution should be exercised in generalizing the findings to other contexts; b) the results of the meta-analysis in certain conditions may be subject to a degree of variability given that  $N$ s  $< 400$  or  $k \leq 3$  is no guarantee of the stability of sampling estimates (Hunter & Schmidt, 2015); c) due to insufficient primary studies in the  $Esd$  and  $Od$  Scales, the effect sizes could not be corrected; and d) the results of the self-deception (SD) and positive impression management (IM) scales cannot be directly generalized to forensic practice since they are mediated by conscious or not manipulation that have different legal implications.

**Table 4** Results of the meta-analyses for the gender of the litigator as moderator.

Scale/Subsample	<i>k</i>	<i>N<sub>T</sub></i>	<i>d<sub>w</sub></i>	<i>SD<sub>d</sub></i>	<i>SD<sub>pre</sub></i>	<i>SD<sub>res</sub></i>	$\delta$	<i>SD<sub>\delta</sub></i>	%Var	95% <i>CI<sub>d</sub></i>	80% <i>CI<sub>\delta</sub></i>	95% <i>CI<sub>\delta</sub></i>
<i>L Scale</i>												
Fathers	24	2783	0.67	0.35	0.19	0.30	0.76	0.34	28.73	0.59, 0.75	0.32, 1.21	0.68, 0.84
Mothers	24	2857	0.81	0.41	0.19	0.37	0.92	0.42	21.21	0.73, 0.89	0.38, 1.46	0.84, 1.00
<i>K Scale</i>												
Fathers	23	2723	0.67	0.22	0.19	0.11	0.74	0.13	72.18	0.59, 0.75	0.57, 0.91	0.66, 0.82
Mothers	23	2801	0.75	0.23	0.18	0.14	0.84	0.15	64.55	0.67, 0.83	0.64, 1.04	0.76, 0.92
<i>F Scale</i>												
Fathers	18	2514	-0.34	0.25	0.17	0.18	-0.39	0.21	45.95	-0.42, -0.26	-0.66, -0.11	-0.47, -0.31
Mothers	17	2499	-0.17	0.27	0.16	0.22	-0.20	0.25	36.19	-0.25, -0.09	-0.52, 0.12	-0.28, -0.12
<i>S Scale</i>												
Fathers	8	1306	0.81	0.22	0.16	0.16	0.87	0.17	51.25	0.69, 0.93	0.65, 1.09	0.75, 0.99
Mothers	8	1418	0.90	0.22	0.15	0.15	0.97	0.16	51.05	0.80, 0.99	0.76, 1.19	0.85, 1.09

Note. Studies only from MMPI-2; 95% *CI<sub>\delta</sub>* = 95% confidence interval for  $\delta$ .

**Table 5** Results of the meta-analyses for the evaluation context as moderator.

Scale/Index	<i>k</i>	<i>N<sub>T</sub></i>	<i>d<sub>w</sub></i>	<i>SD<sub>d</sub></i>	<i>SD<sub>pre</sub></i>	<i>SD<sub>res</sub></i>	$\delta$	<i>SD<sub>\delta</sub></i>	%Var	95% <i>CI<sub>d</sub></i>	80% <i>CI<sub>\delta</sub></i>	
<i>L Scale</i>												
PCCDs	60	10099	0.85	0.37	0.16	0.34	0.97	0.38	18.43	0.81, 0.89	0.47, 1.47	
PCA-CPCs	7	543	1.24	0.14	0.24	0	1.41	0	100	1.06, 1.42	1.41	
<i>K Scale</i>												
PCCDs	58	9611	0.85	0.24	0.16	0.18	0.95	0.20	43.22	0.81, 0.89	0.68, 1.21	
PCA-CPCs	7	543	0.26	0.10	0.22	0	0.28	0	100	0.08, 0.44	0.28	
<i>F Scale</i>												
PCCDs	47	8785	-0.28	0.22	0.14	0.16	-0.32	0.19	43.27	-0.32, -0.24	-0.57, -0.07	
PCA-CPCs	5	446	0.71	0.29	0.21	0.19	0.81	0.22	55.97	0.53, 0.89	0.53, 1.10	
<i>S Scale</i>												
PCCDs	16	3043	0.89	0.24	0.15	0.18	0.96	0.20	40.58	0.81, 0.97	0.71, 1.22	
PCA-CPCs	3	220	0.20	0.15	0.23	0	0.21	0	100	-0.07, 0.47	0.21	

Note. Studies only from MMPI-2; Meta-analysis only for generalized scales and indexes.

Further research is required to assess the defensiveness detection capacity of the *Es<sub>d</sub>* and *Od* scales given the lack of studies in the literature and the insufficient *N<sub>s</sub>*; to evaluate the effects of the evaluation context, and to assess the revised MMPI-2-RF scales that could not be used as a moderator in this study owing to the lack of studies. Thus, more studies with the MMPI-2-RF validity scales are necessary. Nevertheless, as for the substitution of the *L* and *K*, the original defensiveness scales of the MMPI-2, as well as the *F* scale as it was used to compute indexes, for the reformulated ones in the MMPI-2-RF i.e., the *L-r*, *K-r*, and *F-r*, and the subsequent indexes derived from these, a great number of studies with a significantly higher mean true effect size will be required. Hence, a File Drawer Analysis showed that for *L*, *K* and *F* scales would be necessary 615, 498 and 143 studies, respectively, to reverse the results from MMPI-2 to a trivial effect or to attribute them to a sampling bias. Additionally, there is no evidence about the performance of the indexes with

the MMPI-2-RF. Moreover, the *S* and *MP* additional validity scales (results from *Wsd* scale are not generalizable) were not reformulated for MMPI-2-RF. As a combination of all the measures of defensiveness is necessary to classify defensiveness in forensic practice (the wrong classification of a protocol as defensive is not permitted in forensic practice as it supposes a false allegation against assessed person) (Arce, Fariña, & Vilariño, 2015; Fariña et al., 2010), while awaiting further evidence for MMPI-2-RF and for the reformulation of additional validity scales, the MMPI-2 must be preferred.

## Funding

This research has been sponsored by a grant of the Spanish Ministry of Economy and Competitiveness (PSI2014-53085-R).

**Appendix 1. Characteristics of the Primary Studies.**

	Source	Instrument	N	Subsample	Evaluation context
Agüero and Álvarez-Icaza (2014)	Paper	MMPI-2	345	Fathers	PCCD
			342	Mothers	PCCD
Arce, Fariña, and Vilarino (2015)	Paper	MMPI-2	488	All	PCCD
Archer, Hagan, Mason, Handle, and Archer (2012)	Paper	MMPI-2-RF	172	Fathers	PCCD
			172	Mothers	PCCD
Bagby et al. (1999)	Paper	MMPI-2	57	Fathers	PCCD
			58	Mothers	PCCD
Bathurst, Gottfried, and Gottfried (1997)	Paper	MMPI-2	258	Fathers	PCCD
			250	Mothers	PCCD
Butcher (1997)	Paper	MMPI-2	868	Fathers	PCCD
			911	Mothers	PCCD
Caldwell (2004)	Unpublished	MMPI-2	1867	All	PCCD
Carr et al. (2005)	Paper	MMPI-2	73	Fathers	PCA-CPC
			91	Mothers	PCA-CPC
Cooke (2010)	Paper	MMPI-2	50	Fathers	PCCD
			50	Mothers	PCCD
Daskalakis (2004)	Doctoral thesis	MMPI-2	49	All	PCCD
Ezzo, Pinsoneault, and Evans (2007)	Paper	MMPI-2	70	All	PCCD
			205	All	PCCD
Fariña et al. (2010)	Paper	MMPI-2	126	All	PCCD
Gordon, Stoffey, and Bottinelli (2008)	Paper	MMPI-2	79	Fathers	PCCD
			79	Mothers	PCCD
Gordon, Stoffey, and Bottinelli (2008)	Paper	MMPI-2	41	Fathers	PCCD
			41	Mothers	PCCD
			7	Fathers	PCCD
			7	Mothers	PCCD
			31	Fathers	PCCD
			31	Mothers	PCCD
			31	Fathers	PCCD
Gready (2006)	Doctoral thesis	MMPI-2	66	Mothers	PCA-CPC
			116	Fathers	PCCD
			124	Mothers	PCCD
Hopkins (1999)	Doctoral thesis	MMPI/MMPI-2	207	Fathers	PCCD
			219	Mothers	PCCD
Kauffman, Stolberg, and Madero (2015)	Paper	MMPI-2	51	All	PCCD
Leib (2006)	Doctoral thesis	MMPI-2	6	Fathers	PCCD
			18	Mothers	PCCD
			7	Fathers	PCCD
			18	Mothers	PCCD
Mandappa (2004)	Doctoral thesis	MMPI-2	420	All	PCCD
Moreland and Greenberg (1993)	Unpublished	MMPI	201	All	PCCD
			33	Fathers	PCCD
			32	Mothers	PCCD
Normington (2006)	Doctoral thesis	MMPI-2	19	All	PCA-CPC
			19	All	PCCD
Ollendick and Collings (1984)	Paper	MMPI	38	Fathers	PCCD
			38	Mothers	PCCD
Peters (2012)	Doctoral thesis	MMPI-2	68	All	PCCD
			57	All	PCCD
Posthuma and Harper (1998)	Paper	MMPI-2	40	Fathers	PCCD
			40	Mothers	PCCD
			27	Fathers	PCCD
			27	Mothers	PCCD
			27	Fathers	PCCD
			27	Mothers	PCCD
Rehil (2011)	Doctoral thesis	MMPI-2	61	All	PCCD



	Source	Instrument	N	Subsample	Evaluation context
Resendes and Lecci (2012)	Paper	MMPI-2	136	All	PCA-CPC
Roma et al. (2014)	Paper	MMPI-2	194	Fathers	PCCD
			197	Mothers	PCCD
Schenk (1996)	Paper	MMPI-2	60	Fathers	PCCD
			56	Mothers	PCCD
			46	Fathers	PCCD
			34	Mothers	PCCD
Stredny and Archer (2006)	Paper	MMPI-2	127	All	PCA-CPC
Strong et al. (1999)	Paper	MMPI-2	206	Fathers	PCCD
			206	Mothers	PCCD
Wakefield and Underwager (1990)	Paper	MMPI-2	32	Fathers	PCCD
			27	Mothers	PCCD
Wisneski (2006)	Doctoral thesis	MMPI-2	626	All	PCCD

Note. PCCD = parent child custody disputes; PCA-CPC = parenting capacity assessment in child protection cases.

## References

- Ackerman, J. J., & Pritzl, T. B. (2011). Child custody evaluation practices: A 20-year follow-up. *Family Court Review*, 49, 618–628. <http://dx.doi.org/10.1111/j.1744-1617.2011.01397.x>
- Agüero, M. C. P., & Álvarez-Icaza, M. A. V. (2014). Evaluación psicológica con el MMPI-2 a padres en litigio judicial de materia familiar. *Revista Intercontinental de Psicología y Educación*, 16, 71–91.
- Amado, B. G., Arce, R., & Fariña, F. (2015). Undeutsch hypothesis and Criteria Based Content Analysis: A meta-analytic review. *European Journal of Psychology Applied to Legal Context*, 7, 3–12. <http://dx.doi.org/10.1016/j.ejpal.2014.11.002>
- Amado, B. G., Arce, R., Fariña, F., & Vilariño, M. (2016). CBCA reality criteria in adults: A meta-analytic review. *International Journal of Clinical and Health Psychology*, 15, 29–36. <http://dx.doi.org/10.1016/j.ijchp.2016.01.002>
- Amato, P. R., & Keith, B. (1991). Parental divorce and adult well-being: A meta-analysis. *Journal of Marriage and the Family*, 53, 43–58. <http://dx.doi.org/10.1037/0893-3200.15.3.355>
- American Psychological Association (2010). Guidelines for child custody evaluations in family proceedings. *American Psychologist*, 49, 677–680. <http://dx.doi.org/10.1037/a0021250>
- Arce, R., Fariña, F., & Seijo, D. (2005). Razonamientos judiciales en procesos de separación: Análisis cognitivo y de contenido de las motivaciones. *Psicothema*, 17, 57–63.
- Arce, R., Fariña, F., Seijo, D., & Novo, M. (2015). Assessing impression management with the MMPI-2 in child custody litigation. *Assessment*, 22, 769–777. <http://dx.doi.org/10.1177/1073191114558111>
- Arce, R., Fariña, F., & Vilariño, M. (2015). Daño psicológico en casos de víctimas de violencia de género: Un estudio comparativo de las evaluaciones forenses. *Revista Iberoamericana de Psicología y Salud*, 6, 72–80. <http://dx.doi.org/10.1016/j.rips.2015.04.002>
- Arch, M., Jarne, A., Peró, M., & Guàrdia, J. (2011). Child custody assessment: A field survey of Spanish forensic psychologists' practices. *European Journal of Psychology Applied to Legal Context*, 3, 107–128.
- Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment*, 87, 84–94. [http://dx.doi.org/10.1207/s15327752jpa8701\\_07](http://dx.doi.org/10.1207/s15327752jpa8701_07)
- Archer, E., Hagan, L. D., Mason, J., Handle, R., & Archer, R. (2012). MMPI-2-RF characteristics of custody evaluation litigants. *Assessment*, 19, 14–20. <http://dx.doi.org/10.1177/1073191110397469>
- Archer, R. P., & Wygant, D. (2012). Child custody evaluations: Ethical, scientific, and practice considerations. *Journal of Psychological Practice*, 17, 1–70.
- Baer, R. A., & Miller, J. (2002). Underreporting of psychopathology on the MMPI-2: A Meta-analytic review. *Psychological Assessment*, 24, 16–26. <http://dx.doi.org/10.1037/1040-3590.14.1.16>
- Baer, R. A., Wetter, M. W., & Berry, D. T. R. (1992). Detection of underreporting of psychopathology of the MMPI: A meta-analysis. *Clinical Psychology Review*, 12, 509–552. [http://dx.doi.org/10.1016/0272-7358\(92\)90069-K](http://dx.doi.org/10.1016/0272-7358(92)90069-K)
- Baer, R. A., Wetter, M. W., & Berry, D. T. R. (1995). Effects of information about validity scales on underreporting of symptoms. *Assessment*, 2, 189–200. <http://dx.doi.org/10.1177/107319119500200209>
- Baer, R. A., Wetter, M. W., Nichols, D. S., Greene, R., & Berry, D. T. R. (1995). Sensitivity of the MMPI-2 validity scales to underreporting of symptoms. *Psychological Assessment*, 7, 419–423. <http://dx.doi.org/10.1037/1040-3590.7.4.419>
- Bagby, R. M., & Marshall, M. B. (2004). Assessing underreporting response bias on the MMPI-2. *Assessment*, 11, 115–126. <http://dx.doi.org/10.1177/1073191104265918>
- Bagby, R. M., Nicholson, R. A., Buis, T., Radovanic, H., & Fidler, B. J. (1999). Defensive responding on MMPI-2 in family custody and access evaluations. *Psychological Assessment*, 11, 24–28.
- Bagby, R. M., Rogers, R., Nicholson, R. A., Buis, T., Seeman, M. V., & Rector, N. A. (1997). Effectiveness of MMPI-2 validity indicators in the detection of defensive responding in clinical and nonclinical samples. *Psychological Assessment*, 9, 406–413. <http://dx.doi.org/10.1037/1040-3590.9.4.406>
- Bathurst, K., Gottfried, A. W., & Gottfried, A. E. (1997). Normative data for the MMPI-2 in child custody litigation. *Psychological Assessment*, 9, 205–211. <http://dx.doi.org/10.1037/1040-3590.9.3.205>
- Ben-Porath, Y. S., & Tellegen, A. (2008/2011). *MMPI-2-RF (Minnesota Multiphasic Personality Inventory-2 Restructured Form): Manual for administration, scoring, and interpretation*. Minneapolis, MN: University of Minnesota Press.
- Bow, J. N., & Quinnell, F. A. (2001). Psychologists' current practices and procedures in child custody evaluations: Five years after American Psychological Association guidelines. *Professional Psychology: Research and Practice*, 32, 261–268. <http://dx.doi.org/10.1037/0735-7028.32.3.261>
- Butcher, J. N. (1997). Frequency of MMPI-2 scores in forensic evaluations. *MMPI-2 News and Profiles*, 8, 2–4.
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Manual for the restandardized Minnesota Multiphasic Personality Inventory: MMPI-2*. Minneapolis, MN: University of Minnesota Press.

- Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W. G., & Kaemmer, B. (2001). *MMPI-2 (Minnesota Multiphasic Personality Inventory-2): Manual for administration, scoring, and interpretation* (revised edition). Minneapolis, MN: University of Minnesota Press.
- Butcher, J. N., & Han, K. (1995). Development of an MMPI-2 scale to assess the presentation of self in a superlative manner: The S scale. In J. N. Butcher, & C. D. Spielberg (Eds.), *Advances in personality assessment* (10) (pp. 25–50). Hillsdale, NJ: LEA.
- Caldwell, A.B. (2004). [MMPI-2 child-custody dataset]. Unpublished raw data.
- Carr, G. D., Moretti, M. M., & Cue, B. J. H. (2005). Evaluating parenting capacity: Validity problems with the MMPI-2, PAI, CAPI, and ratings of child adjustment. *Professional Psychology: Research and Practice*, 36, 188–196.
- Chacón, F., García, J. F., García, A., Gómez, R., & Vázquez, B. (2009). *Guía de buenas prácticas para la elaboración de informes psicológicos periciales sobre custodia y régimen de visitas de menores*. Madrid: Colegio de Psicólogos de Madrid.
- Cheng, H., Dunn, J., O'Connor, T. G., & Holding, J. (2006). Factors moderating children's adjustment to parental separation: Findings from a community study in England. *Journal of Abnormal Child Psychology*, 2, 239–250.
- Cooke, G. (2010). MMPI-2 defensiveness in child custody evaluations: The role of education and socioeconomic level. *American Journal of Forensic Psychology*, 28, 5–16.
- Daskalakis, K. (2004). *The use of the MMPI-2 in complex issues of high conflict child-custody disputes*. Ontario, Canada: University of Toronto. Doctoral dissertation Retrieved <http://search.proquest.com/docview/305082915/6411BF7FAC0B44F0PQ/5?accountid=17253>.
- De Dreu, C. K., & Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. *Journal of Applied Psychology*, 88, 741–749. <http://dx.doi.org/10.1037/0021-9010.88.4.741>
- Ezzo, F., Pineseault, T. B., & Evans, T. M. (2007). A comparison of MMPI-2 profiles between child maltreatment cases and two types of custody cases. *Journal of Forensic Psychology Practice*, 7, 29–43. <http://dx.doi.org/10.1080/15228932.2012.674469>
- Fariña, F., Arce, R., & Real, S. (1994). Ruedas de identificación: De la simulación y la realidad. *Psicothema*, 6, 395–402.
- 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? *Revista Iberoamericana de Psicología y Salud*, 1, 65–79.
- Fariña, F., Arce, R., Vilariño, M., & Novo, M. (2014). Assessment of the standard forensic procedure for the evaluation of psychological injury in intimate-partner violence. *The Spanish Journal of Psychology*, 17, 1–10. <http://dx.doi.org/10.1017/sjp.2014.30>, e32
- Glass, G. V. (1976). Primary, secondary, and meta-analysis research. *Educational Researcher*, 5, 3–8.
- Glass, G. V., McGraw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.
- Gordon, R. M., Stoffey, R., & Bottinelli, J. (2008). MMPI-2 findings of primitive defences in alienating parents. *The American Journal of Family Therapy*, 36, 221–228.
- Graham, J. R. (2011). *MMPI-2: Assessing personality and psychopathology* (5th ed.). New York, NY: Oxford University Press.
- Gready, P. A. (2006). *Use of the MMPI-2 in child custody evaluations and child protection evaluations: An examination of defensive responding and psychopathology*. Doctoral dissertation, University of Hartford. Retrieved from <http://search.proquest.com/docview/304958565/fulltextPDF/D9A6A25E7F2C4C8EPQ/1?accountid=17253>.
- Greene, R. G. (2011). *The MMPI-2/MMPI-2-RF. An interpretative manual* (3rd ed.). Boston, MA: Allyn and Bacon.
- Hathaway, S. R., & McKinley, J. C. (1940). A multiphasic personality schedule (Minnesota): I. Construction of the schedule. *Journal of Psychology*, 10, 249–254.
- Hopkins, L. (1999). *The role of the K scale as a validity measure in court-ordered child custody MMPI's*. Doctoral dissertation, Kent State University. Retrieved from <http://search.proquest.com/docview/619446221/184D9C7FCF38493DPQ/11?accountid=17253>.
- Hunsley, J., Hanson, R. K., & Parkeret, K. C. H. (1988). A summary of the reliability and stability of MMPI scales. *Journal of Clinical Psychology*, 44, 44–46.
- Hunter, J. E., & Schmidt, F. L. (2015). *Methods of meta-analysis: Correcting errors and bias in research findings* (3rd ed.). Thousand Oaks, CA: Sage.
- Kauffman, C. M., Stolberg, R., & Madero, J. (2015). An examination of the MMPI-2-RF (Restructured Form) with the MMPI-2 and MCMI-III of child custody litigants. *Journal of Child Custody*, 12, 129–151. <http://dx.doi.org/10.1080/15379418.2015.1057354>
- Konecni, V. J., & Ebbesen, E. B. (1992). Methodological issues on legal decision-making, with special reference to experimental simulations. In F. Lösel, D. Bender, & T. Bliesener (Eds.), *Psychology and law: International perspectives* (pp. 413–423). Berlin, Germany: Walter de Gruyter.
- Lanyon, R. I., & Lutz, R. W. (1984). MMPI discrimination of defensive and nondefensive felony sex offenders. *Journal of Consulting and Clinical Psychology*, 52, 841–843.
- Leib, R. (2006). *MMPI-2 family problems scales in child-custody litigants*. Doctoral dissertation, Pacific Graduate School of Psychology. Retrieved from <http://search.proquest.com/docview/304937412/184D9C7FCF38493DPQ/22?accountid=17253>.
- Mandappa, P. (2004). *MMPI-2: The need for specific norms in child custody evaluations*. Doctoral dissertation, The Chicago School of Professional Psychology. Retrieved from <http://search.proquest.com/docview/305057409/184D9C7FCF38493DPQ/6?accountid=17253>.
- Martindale, D. A., Martin, L., Autin, W. G., & Task Force Members. (2007). Model standards of practice for child custody evaluations. *Family Court Review*, 45, 70–91.
- Mosier, C. I. (1943). On the reliability of a weighted composite. *Psychometrika*, 8, 161–168.
- Moreland, K., & Greenberg, S.A. (1993). Unpublished raw data.
- Normington, D. (2006). *Caregiver competency evaluations: An examination of psychological characteristics of caregivers who neglect their children*. Doctoral dissertation, Pacific Graduate School of Psychology. Retrieved from <http://search.proquest.com/docview/304937462/C47F7B0B2ECA4917PQ/1?accountid=17253>.
- Ollendick, D. G., & Collings, R. P. (1984). MMPI characteristics of parents referred of child custody cases. *Journal of Psychology*, 117, 227–232.
- Osuna, E., López-Martínez, M., Arce, R., & Vázquez, M. J. (2015). Analysis of response patterns on the MMPI-2 in psychiatric prison inmates. *International Journal of Clinical and Health Psychology*, 15, 29–36. <http://dx.doi.org/10.1016/j.ijchp.2014.09.002>
- Palmer, A., Borrás, C., Pérez-Pareja, J., Sesé, A., & Vilariño, M. (2013). Are patients with chronic pain and fibromyalgia correctly classified by MMPI-2 validity scales and indexes? *European Journal of Psychology Applied to Legal Context*, 5, 123–129. <http://dx.doi.org/10.5093/ejpalc2013a1>
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598–609. <http://dx.doi.org/10.1037/0022-3514.46.3.598>
- Peters, K. M. (2012). *Marital conflict in child custody disputes and the corresponding psychological variables*.

- Doctoral dissertation, Immaculata University. Retrieved from <http://search.proquest.com/docview/1464395031/184D9C7FCF38493DPQ/24?accountid=17253>.
- Posthuma, A. B., & Harper, J. F. (1998). Comparison of MMPI-2 responses of child custody and personal injury litigants. *Professional Psychology: Research and Practice, 29*, 437–443.
- Rehil, H. (2011). *Clinical usefulness and common profiles of the MMPI-2 RF in child custody assessment*. Doctoral dissertation, Indiana University of Pennsylvania. Retrieved from <http://search.proquest.com/docview/885211308/184D9C7FCF38493DPQ/2?accountid=17253>.
- Resendes, J., & Lecci, L. (2012). Comparing the MMPI-2 scale scores of parents involved in parental competency and child custody assessments. *Psychological Assessment, 24*, 1054–1059. <http://dx.doi.org/10.1037/a0028585>
- 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. <http://dx.doi.org/10.1177/1073191103252349>
- Roma, P., Ricci, F., Kotzalidis, G. D., Abbate, L., Lubrano, A., Versace, G., Pazzelli, F., Malagoli, M., Girardi, P., & Ferracuti, S. (2014). MMPI-2 in child custody litigation: A comparison between genders. *European Journal of Psychological Assessment, 30*, 110–116. <http://dx.doi.org/10.1027/1015-5759/a000192>
- Schenk, P. W. (1996). *MMPI-2 norms for child custody litigants*. *The Georgia Psychologist, 50*, 51–54.
- Stredny, R. V., & Archer, R. P. (2006). MMPI-2 and MCMI-III characteristics of parental competency examinees. *Journal of Personality Assessment, 87*, 113–115. <http://dx.doi.org/10.1207/s15327752jpa8701.10>
- Strong, D. R., Greene, R. L., Hoppe, C., Johnston, T., & Olesen, T. (1999). Taxonometric analysis of impression management and self-deception on the MMPI-2 in child-custody litigants. *Journal of Personality Assessment, 73*, 1–18. <http://dx.doi.org/10.1207/S15327752JPA730101>
- Strong, D. R., Greene, R. L., & Kordinak, S. T. (2002). Taxometric analysis of impression management and self-deception in college student and personnel evaluation settings. *Journal of Personality Assessment, 78*, 161–175. <http://dx.doi.org/10.1207/S15327752JPA7801.10>
- Tukey, J. W. (1960). A survey of sampling from contaminated distributions. In I. Olkin, J. G. Ghurye, W. Hoeffding, W. G. Madow, & H. Mann (Eds.), *Contributions to probability and statistics* (pp. 448–485). Stanford, CA: Stanford University Press.
- Vilariño, M., Arce, R., & Fariña, F. (2013). Forensic-clinical interview: Reliability and validity for the evaluation of psychological injury. *European Journal of Psychology Applied to Legal Context, 5*, 1–21.
- Wakefield, H., & Underwager, R. C. (1990). Personality characteristics of parents making false accusations of sexual abuse in custody disputes. *Issues in Child Abuse Accusation, 2*, 121–136.
- Weaver, J. M., & Schofield, T. J. (2015). Mediation and moderation of divorce effects on children's behavior problems. *Journal of Family Psychology, 29*, 39–48. <http://dx.doi.org/10.1037/fam0000043>
- Wisneski, J. L. (2006). *The MMPI-2 in contested child custody cases: Differences for parents in entrenched disputes*. Doctoral dissertation, Hofstra University. Retrieved from <http://search.proquest.com/docview/305324459/184D9C7FCF38493DPQ/7?accountid=17253>.