Psychosocial Intervention 22 (2013) 153-160

Psychosocial Intervention

www.elsevier.es/psi

Batterer intervention programmes: A meta-analytic review of effectiveness

Esther Arias, Ramón Arce* y Manuel Vilariño

Departamento de Psicología Social, Básica y Metodología, Universidad de Santiago de Compostela

INFORMACIÓN ARTÍCULO

Historia artículo:
Recibido: 14/09/2012
Aceptado: 28/02/2013

Key words:
Batterer
Intervention programme
Recidivism
Reoffending
Effectiveness
Meta-analysis.

ABSTRACT

A meta-analysis of the state-of-the-art on the efficacy of batterer treatment programmes was conducted from the year 1975 to 2013. A total of 19 Spanish and English language research articles were retrieved yielding 49 effect sizes from a sample of 18,941 batterers. The results revealed that the recidivism rate as measured by couple reports (CR) was significantly higher than the rate based on official reports (OR), since the recidivism as measured by OR is underestimated. Overall, treatment showed a non significant positive weighted mean effect, δ = 0.41. Nevertheless, the counternull effect size, $ES_{counternull}$ =0.82, suggested a null effect was as probable as a treatment efficacy rate of 38%. The intervention type was not a significant moderator of recidivism, but the counternull effect sizes, $ES_{counternull}$ = 0.82 and 0.94, revealed an efficacy rate of 38% and 42% based on ORs, for Duluth Model and behavioral-cognitive treatment, respectively. The long-term treatment interventions had a significantly positive medium effect size, δ = 0.49. The implications of these findings for the design and assessment of future intervention programmes are discussed.

Programas de intervención con maltratadores: Una revisión meta-analítica de su efectividad

Palabras clave:
Maltratador
Programa de intervención
Recidiva
Recaída en el maltrato
Eficacia
Metaanálisis

RESUMEN

Este artículo presenta la revisión metaanalítica llevada a cabo con el fin de conocer el estado actual de la eficacia de los programas de tratamiento a maltratadores según los trabajos publicados desde 1975 a 2013. Del total de 19 artículos en inglés y en español recuperados se extrajeron 49 tamaños de efecto a partir de una muestra total de 18.941 maltratadores. Los resultados muestran que el índice de recaída según reflejan los informes de parejas era significativamente superior que el de los informes oficiales, dado que en estos últimos está subestimado. En general el tratamiento presentaba un tamaño del efecto medio ponderado positivo pero no significativo (δ = 0.41). Sin embargo el valor contranulo del tamaño del efecto, $ES_{contranulo}$ = 0.82, indicaba que el efecto nulo era tan probable como un índice de eficacia del tratamiento del 38%. El tipo de intervención no moderaba significativamente la recaída, aunque los valores contranulos del tamaño del efecto $ES_{contranulo}$ = 0.82 y 0.94 indicaban un índice de eficacia del 38% y 42% respectivamente, de acuerdo a los informes oficiales, para el tratamiento con el modelo Duluth y el cognitivo conductual respectivamente. Las intervenciones a largo plazo tenían un tamaño del efecto medio significativo positivo de δ = 0.49. Se comenta la implicación que estos resultados pueda tener para el diseño y evaluación de programas de intervención futuros.

Prior to the first meta-analysis on the efficacy of batterers' treatment programmes, the reviews of intervention programmes yielded contradictory results. Thus, while Hamberger and Hastings (1993), and Rosenfeld (1992) concluded that treatments did not work, Davis and Taylor (1999) found that the effect of treatment was substantial, *h* = 0.41. Though several authors (Babcok, Green, & Robie, 2004) assert these effect sizes are modest in terms of Cohen's (1988) classification categories, i.e., a small effect size (*h* < 0.50), it should be noted that Cohen himself indicated that the magnitude of the effect size should not be taken as an absolute value, but should rather be interpreted according to the effects anticipated in a given context. Thus, if this effect size is contrasted with the effect size between cognitive distortions and violence as measured in *d* = 0.82 (Chereji, Pintea, & David, 2012), corresponding to a 68.5% improvement with treatment in contrast to 32.5% of the control group (a large effect size), the results do not appear to be promising. Notwithstanding the foregoing, in comparison to the treatment efficacy of delinquents as

*e-mail: ramon.arce@usc.es

1132-0559/$ - see front matter © 2013 Colegio Oficial de Psicólogos de Madrid. Todos los derechos reservados

DOI: http://dx.doi.org/10.5093/in2013a18

measured by the recidivism rate with a small effect size of *d* ranging from 0.23 to 0.42 (Redondo, Sánchez-Meca, & Garrido, 1999, 2001, 2002), the data obtained by Davis and Taylor represent an increase similar to the treatment efficacy expected in this context. Moreover, the effect size of Davis and Taylor (1999) entails a 20% increase in the recidivism rate, which is equated to 40% of batterers reoffending in comparison to the 60% of non-treated offenders. Previous meta-analyses have identified moderators with significant positive effects with a small treatment efficacy effect size (Babcok et al., 2004; Feder & Wilson, 2005; Levesque & Gelles, 1998). Furthermore, these authors have obtained inconsistent results or even negative treatment effects. Though the effect sizes were small, they were comparable to those obtained for the treatment of delinquents and were indicative of treatment efficacy. To put it another way, a woman is 5% less likely to be reassaulted by a man who was arrested, sanctioned, and sent to a batterers' programme than by a man who was simply arrested and sanctioned (Babcock et al., 2004, p. 1044). Working with an estimated population of 100,000 batterers, this would equal 5,000 fewer batterers; in relation to the hypothetical recidivism base rate of approximately 25% (Bennett, Call, Flett, & Stoops, 2005; Gondolf, 2004) this would imply a recidivism rate of around 20,000 of treated batterers.

The most consistent results highlight that the recidivism base rate of non-treated batterers as measured by Official Reports (ORs) is lower (21%) than the measure based on Couple Reports (CRs), which is estimated to be 35% (Babcock et al., 2004; O'Leary et al., 1989; Rosenfled, 1992). Surprisingly, previous meta-analyses have tended to focus on design variables, e.g., experimental vs. quasi-experimental variables, which have a methodological-scientific significance but do not provide specific guidelines aimed at enhancing treatment efficacy. With this purpose in mind, Arce and Fariña (2010), Lila, Oliver, Galiana, and Gracia (2013) and McGuire, Mason, and O'Kane (2000) have defined variables which are considered to be fundamental for the implementation of batterers' treatment, specifically for those ordered in the community such as: contents, length (number and interval between sessions), duration, intervention level, risk assessment, treatment adherence and progress, and the rationale underlying intervention. First, programmes that are tailored to the specific needs of each batterer enhance treatment efficacy (Holtzworth-Munroe, Meehan, Herron, Rehman, & Stuart, 2000), whereas standard programmes with similar content across the board for all batterers not only lack efficacy but may even prove counterproductive due to the failure to adapt the intervention to the needs of each batterer (Bowen, Gilchrist, & Beech, 2005). Thus, programmes should seek to address the specific needs of each individual batterer though in practice this strategy is almost always neglected. Second, gender violence is entrenched in a culture of violence that has been described as a form of toxic cognition that is essentially internal, stable, and global (Maruna, 2004). Consequently, brief interventions are less effective than long-term programmes since the duration of each session as well as the number of and interval between sessions have a decisive impact on the acquisition and consolidation of socio-cognitive skills given that domestic violence is grounded in internal, stable, and global cognition associated to ongoing recidivism and violent behaviour (Collie, Vess, & Murdoch, 2007; Hutchings, Gannon, & Gilchrist, 2010), which is highly resistant to treatment and hinders adherence (Isorna, Fernández-Ríos, & Souto, 2010; Wormith & Olver, 2002). Third, conventional interventions, of which multimodal interventions (cognitive-behavioural) have proven to be most effective (Beelman & Lösel, 2006; Redondo et al., 1999, 2001, 2002), have focused exclusively on the batterer and have often neglected other aspects that are crucial for social integration and competence through social bonding and employment. Thus, alienation or unemployment foster the continuity of a cycle of violence (Fariña, Arce, & Novo, 2008; Gracia, Herrero, Lila, & Fuente, 2009). Moreover, multimodal interventions involving individual (cognition) and group (behavioural) sessions achieve better outcomes than group-only sessions (Arce & Fariña, 2010; Novo, Fariña, Seijo, & Arce, 2012). An exhaustive control of treatment adherence and progress is not feasible in group sessions, and they fail to stengthen responsibility taking among batterers. Thus, multimodal and multilevel interventions involving individual and group sessions are more effective than exclusively individual sessions. Forth, a treatment requires an ongoing means of measuring the effects of treatment –in this case, teatment progress. In general contexts, such as clinical evaluation, the aim of the assessment is to determine treatment outcomes, but forensic or prison contexts require a differential diagnosis of feigning (American Psychiatric Association, 2000) focused on ensuring treatment adherence and progress. In other words, the feigning of treatment adherence and progress is prevalent among convicted batterers and sexual offenders who seek to gain prison benefits, hence the high risk of recidivism. Fifth, the therapeutic rationale underlying most batterer treatment programmes undermines treatment efficacy in two ways. Treating batterers as patients implies batterers are not responsible for their own behaviour owing to exogenous causes, which hinders treatment adherence and progress, and justifies the persistance of a culture of violence (Maruna & Copes, 2005). Furthermore, the professional implementing the treatment programme may be conceived as an unwitting accomplice aiding the batterer. An alternative is a rationale whereby the role of the professional is to apply the law and serve the wider interests of society to ensure the batterers become fully aware that they are the only ones to be held directly accountable for their behaviour.

The assessment of batterers' treatment programme efficacy has been the source of much controversy regarding reliability of measures. Though recidivism in domestic violence is the most extensively used criterion for measuring treatment efficacy, a wide range of measures have been employed to assess recidivism rates such as police or court reports, trial convictions, prison sentencing, victim reports, partner reports, or even batterer self-reports. Due to the considerable amount of overlapping between police, court, and prison databases, these data are often jointly referred to as Official Reports/Registers. Nevertheless, the reliability of these sources as an estimate of recidivism remains a controversial issue in the literature (Novo et al., 2012). For instance, meta-analysis (Babcock et al., 2004; O'Leary et al., 1989; Rosenfeld, 1992) have found a 21% recidivism rate based on ORs and 35% rate based on CRs (= 0.42), i.e., CRs report 0.42 standard deviation more recidivism than ORs (a medium effect size). A further instance concerns the treatment effects on cognition that sanction and forerun (in comparison to cognitive distortions that have not been shown to precede violence) violence (Maruna & Mann, 2006) –in this case intimate partner violence, what Novo et al. (2012) referred to as the internal mechanisms underlying violence. Though the reliability of these measures based on psychometric instruments has been attested, recidivism continues to be the standard measure of batterers' treatment efficacy, both in the field of science and in terms of socio-political assessment.

Ever since the advent of batterer re-education programmes, two models have been the most extensively used for the treatment of batterers, i.e., the Duluth Model and interventions that have been encompassed under the umbrella term of Cognitive-Behavioural Treatment programmes (CBT). The former, which is currently the most prominent of the two models, takes its name from the pioneering programme set up in Duluth (Minnesota) and combines a gender (feminist) approach with a psychoeducational approach grounded in the assumption that the primary cause of gender violence is patriarchal and sexist ideology that sanctions male dominance and relegates women to submissive obedience. Hence, the goal of treatment is to challenge male dominance and to foster egalitarian relationships. On the other hand, CBT programmes

envisage violence as a learned behaviour, which is best offset by promoting and reinforcing non-violent alternatives aimed at developing social skills and anger management (Babcock et al. 2004). A further option arising from literature reviews and meta-analysis is the creation of another treatment category referred to as "Other Types of Intervention" (OTI) covering a wide variety of treatment programmes such as Psychodynamic counselling, Anger Management, and Mind Body Bridging.

Thus, the aim of this study was to perform a meta-analysis to learn the state-of-the-art of the efficacy of batterer treatment programmes from the 1975 to 2013 by assessing studies measuring treatment efficacy in terms of the recidivism rate.

Method

Database search

The search was restricted to studies assessing batterers' treatment programmes efficacy from 1975, one year after Martinson's (1974) doctrine suggesting that "nothing works" in relation to the treatment of delinquents, to the present date (2013). A review of the batterers' treatment literature was undertaken using the following search strategies: a) search in broad spectrum databases (both small databases and specialized databases with quality control such as Scopus and the Web of Knowledge were included), such as PsycInfo, ERIC, Scirus, Google, and Google Academia; b) search in gender violence observatories (e.g., www.work-with-perpetrators.eu; www.VAWnet.org; www.mincava.umn.edu; www.courtinnovation.org; www.cienciaspenales.net; www.iresweb.org); c) researchers in the field were contacted (i.e., the corresponding authors of both retrieved and excluded articles were contacted); and d) the reference sections of previous meta-analysis were reviewed and cross-referenced.

The list of keywords was generated through a system of successive approximations whereby relevant keywords cited in the articles and previous meta-analysis were cross-referenced. The most productive keywords (other keywords overlapped with the search results) were: batterer, intervention program, evaluation, assessment, effectiveness, intimate partner violence, partner-violent men, recidivism, reoffending, attrition, domestic violence, batterers' reeducation programmes, gender violence aggressors, recidivism, programmes evaluation, prison treatment, and efficacy.

Criteria for inclusion in the study

Bearing in mind the objectives of the meta-analysis, in order to be selected for the study the articles retrieved from the database search should meet the following criteria: a) report sample size; b) report recidivism rate for treatment completers; c) recidivism measured by ORs (official reports, e.g., police, court, or prison reports) and couple reports (the aggressor self-reports were excluded since batterers tend to underreport the true incidence of abuse which would contaminate the results); d) describe the treatment theoretical approach, contents, and duration of the intervention programme; and e) measure recidivism during the follow-up period (studies with a follow-up shorter than 6 months were discarded). In studies where relevant data were lacking, the authors were contacted to request additional data to be subsequently added to the meta-analysis. By applying these criteria, 19 articles from Spanish and English authors were retrieved, yielding 49 effect sizes from a sample of 18,941 batterers.

Data analysis

The procedure consisted of a bare-bones meta-analysis. As the measure of recidivism is often expressed as percentages/proportions and in the studies where this was not the case it was converted into proportions, the measure of recidivism adopted in this meta-analysis was the proportion of reoffending batterers (data on recidivism in other offences were excluded) during the follow-up period. The measure of the effect size was calculated on the basis of the difference in proportions. This involves a previous non-linear transformation of proportions since the simple difference in proportions is not an accurate estimate of effect size –the difference in proportions does not provide a scale of equal detectable units. The effect size in terms of proportions was calculated using Cohen's *h* (1988) and Hedges and Olkin's δ (1985) based on the procedure of Kraemer and Andrews (1982). The results of both methods of analysis were similar, with almost equivalent sizes in the low values and a slightly larger size for the higher for the δ statistics. Nonetheless, this did not affect the qualitative evaluation of the effect size.

In the *h* index, the percentages were transformed into Φ by the formula $2\arcsin\sqrt{p}$. The substraction of the transformed proportions was *h*. For the δ statistics, in line with the procedure of Kraemer-Andrews, the pre- post test effect size was estimated by the difference of the inverse of the normal cumulative distribution function, Φ^{-1}. Thus, δ is the difference of the inverse function of the probability of the experimental group minus the control, $\hat{\hat{\delta}}_1 = \Phi^{-1}(\hat{p}_1^E) - \Phi^{-1}(\hat{p}_1^C)$. The difference of the inverse function in percentages (δ) or of the Φ (*h*), that is, an effect size of 0.20, 0.50, and 0.80 was considered to be small, medium, and large respectively. The δ was the index of choice for the results of the meta-analysis. The studies that met the inclusion criteria were classified as either experimental or quasi-experimental. The experimental design studies (see Table 1 for the list of retrieved articles and the selection criteria) show two recidivism rates, one for the experimental group, i.e., batterers who had completed treatment and another for the control group, i.e., non-treated batterers. The batter intervention studies with a non-equivalent control group design, e.g., studies comparing treatment completers with treatment dropouts, were classified as quasi-experimental (see Table 2 for the list of the 13 retrieved articles and the selection criteria). Given that treatment non-compliance is associated to recidivism, i.e., recidivism rates among treatment dropouts are higher or even doubled the rate among non-treated batterers (Bennett & Williams, 2001; Dutton, Bodnarchuk, Kropp, Hart, & Ogloff, 1997), studies contrasting the recidivism rates of non-equivalent control groups artificially amplify treatment efficacy. As for these designs, the recidivism rate contrast values were .21 for ORs, and .35 for CRs, which are in accordance with the base rates that have been consistently reported in the literature (Babcock et al., 2004; O'Leary et al., 1989; Rosenfeld, 1992). Once the effect sizes had been calculated, the following were computed: the weighted mean δ for the entire sample size; the observed weighted mean variance ($S^2_{\delta w}$); standard deviation ($SD_{\delta w}$); the true variance ($S_{\delta 2}$); standard error (SE_δ); and the confidence interval (90% CI). If the interval contained zero, it indicated heterogeneity (no significant effect) and further analysis was conducted to successively examine other moderators.

To estimate the practical utility of treatments, the Binomial Effect Size Display (BESD) was applied (Rosenthal & Rubin, 1982) transforming δ into *r* by means of the formula $r = \delta/\sqrt{\delta^2 + 4}$. The *r* was converted to a BESD by means of the formula $(.50 \pm r/2) * 100$. The measure of overlapping distributions was performed by *U*1 statistic (Cohen, 1988).

Most of the effect sizes were not significant (the confidence intervals contained 0), indicating the acceptance of H_0. However, the confidence intervals were not exactly precise for accepting the null hypothesis (Cortina & Dunlap, 1997; Frick, 1996). For the effects that were not significant (the confidence interval contained 0) with a medium or large effect size, the hypothesis of a null effect (0) was contrasted by means of $ES_{counternull}$ (Rosenthal & Rubin, 1994), the formula for the size in terms of the correlation being, $r_{counternull} = \sqrt{(4r^2)/(1 + 3r^2)}$.

Table 1
Quasi-experimental designs

Study	*n*	Treatment type and intervention format	Duration and length of intervention	Measure of recidivism during follow-up	Recidivism rate %		Recidivism rate δ	
1. Saunders (1996)	G1: 61 G2: 68	G1: Duluth + CBT G2: Psychodynamic process Group	12 sessions + 20 support (32 weeks)	24 months (OR -CR)	OR G1: 23.2% G2: 20.3%	CR G1: 34% G2: 33.3%	OR -0.07 0.02	CR 0.03 0.05
2. Dobash, Dobash, Cavanagh and Lewis (1996)	40	Duluth Group	6-7 months	12 months (OR -CR)	OR: 7%	CR: 33%	OR: 0.67	CR: 0.05
3. Murphy, Musser, and Maton, (1998)	235	Duluth (format not specified)	22 sessions	12-18 months (OR)	OR:15.7%		0.2	
4. Babcock and Steiner (1999)	106	Duluth Group	36 weeks	24 months (OR)	OR: 8%		0.6	
5. Jones and Gondolf (2002)	P1:213 P2:208 P3:215 P4:217	Duluth Group	P1: 3 months P2: 3 months P3: 5.5 months P4: 9 months	15 and 30 months (OR)	OR P1 P2 P3 P4	15m 31.5% 31.8% 26.2% 24.7%	30m 41.2% 38.6% 34.2% 28.2%	OR P1 P2 P3 P4	15m -0.32 -0.33 -0.17 -0.12	30m -0.58 -0.52 -0.4 -0.23
6. Jenkins and Menton (2003)	114	CBT Group	9 weeks	30 months (OR)	OR: 10%		0.47	
7. Bowen, Gilchrist, and Beech (2005)	86	Duluth Group	24 weeks + 5 sessions	11 months (OR)	OR: 21%		0	
8. Bennet, Call, Flett, and Stoops (2005)	384	Duluth Group	24 weeks	18 months (OR)	OR: 15.4%		0.21	
9. Labriola, Rempel, and Davis (2005)	157	Duluth Group	26 weeks	12 months (OR)	OR: 6%		0.75	
10. Tolleffson and Gross (2006)	102	Duluth Group	20 sessions	7-58 months (OR)	OR: 18%		0.11	
11. Tollefson, Webb, Shumway, Block, and Nakamura (2009)	57	Mind-body Bridging Group	8-10 sessions	9-27 months (OR)	OR: 9%		0.53	
12. Coulter and VandeWeerd (2009)	G1:1424 G2: 9386 G3:1712	G1 and G2 Duluth Group G3 specialized Treatment	G1: 8-12 weeks G2: 26 weeks G3: 26 sem-1year	1-10 years (OR)	OR G1: 8.8% G2: 8.3% G3: 8.6%		OR G1: 0.55 G2: 0.58 G3: 0.56	
13. Pérez, Giménez-Salinas, and Juan (2012)	598	CBT Group	25 weeks	12 months (OR)	OR: 4.6%		0.88	

Coding

For the analysis of moderators the following were coded: the recidivism variables (OR, n = 18,148, k = 33; CR, n = 1,456, k = 13); follow-up time (less than 12 months, k = 13; more than 12 months, k = 35); duration of treatment (< 16 sessions/weeks and > 16 sessions/weeks); intervention level (individual vs. multilevel –the multilevel intervention contingency was not registered); type of session (individual, group, or combined –no type of intervention [individual or combined] contingency was registered. Although Stith, Rosen, & McCollum, 2004 defined a couple intervention as an individual intervention [experimental group 1], actually this is not an individual intervention); contents (adapted to the needs of each batterer vs. homogeneous for all batterers –only one contingency was registered that may be attributed to adapted content, but it referred to clinical cases [group 3 of the study of Coulter & VandeWeerd, 2009], not to batterer treatment); treatment adherence and progress (control of treatment adherence and progress: yes vs. no –no contingency was registered in the measurement of this variable); rationale behind the intervention (therapeutic vs. re-educational –information was not available for coding this variable in a reliable way); risk control (yes vs. no –information was not available for coding this variable in a reliable way); and treatment type (Duluth, k = 29; CBT, k = 8; and OTI, k = 9).

Coding reliability. The coding was carried out separately by two researchers who agreed on all of the coding of the different categories. Thus, coding was reliable.

Results. The results reveal a significantly higher rate (+.156), z = 13.0, p < .001, with a very substantial difference (< 13 *SD*) in recidivisms as measured by CRs in comparison to ORs. Thus, the ORs entail covert recidivism given that this rate may be higher as many couples refused to report their partner's recidivism due to the threat of re-victimization (i.e., a woman may fear that the disclosure of recidivism in the presence of her partner may lead to subsequent retribution), a recidivism rate of .156 that was significantly greater, z = 51.23, p < .001, than the statistically admissible margin of error (.05).

Outlier analysis

Prior to performing the meta-analysis, an outlier analysis was performed to avoid contaminating the results. As treatment efficacy varies according to the variable under assessment, an outlier analysis was conducted for each measure, with the decision criterion being ±2*SD* of the mean effect size δ. The results found that three studies of Stith et al. (2004) were more than 2 standard deviations above the mean for treatment efficacy and were thus eliminated.

Global analysis

Figure 1 illustrates the procedure for calculating the 17 meta-analysis, as well as the resulting effect sizes (δ), the number of studies (k) included in each analysis**,** and the sample size (n). Of the

Table 2
Experimental designs

Study	*n*	Treatment type and intervention format	Duration and length of intervention	Measure of recidivism during follow-up	Recidivism rate %		Recidivism rate δ	
1. Davis, Taylor, and Maxwell (1998)	376 G1: 129 G2:61 GC: 186	Duluth Group	39 hours GE1: 8 weeks GE2: 26 weeks GC: TBC	6 and 12 months (OR-CR)	OR 6m G1: 7% G2: 15% GC: 22% CR 6m G1: 23% G2: 19% GC: 21%	OR 12 m G1: 10% G2: 25% GC: 26% CR 12m G1: 14% G2: 18% GC: 22%	OR 6m G1: 0.7 G2: 0.26 CR 6m G1: -0.07 G2: 0.07	OR 12 m G1: 0.64 G2: 0.03 CR 12m G1: 0.31 G2: 0.14
2. Dunford (2000)	861 G1: 168 G2: 153 G3: 173 GC: 150	CBT G1 and G2 Group G3 Individual	G1: 12 months (6 session weekly and 6 month) G2: 26 weeks + 6-month G3: 12-month	6 months (CR) 12 months (OR)	OR G1: 4% G2: 3% G3: 6% GC: 4%	CR G1:29 % G2: 30% G3: 27% GC: 35%	OR G1: 0 G2: 0.13 G3: -0.19	CR G1:0.17 G2: 0.14 G3: 0.22
3. Feder and Dugan (2004)	404	Duluth GE: Programme GC: Conditional	26 weeks	12 months (OR)	GE: 24% GC: 21%		GE: -0.1	
4. Stith, Rosen, and McCollum (2004)	39 GE1: 14 GE2: 16 GC: 9	Duluth + CBT GE1 Partner individual GE2: Partners Group GC: Pretest and follow-up	6 weeks	6 and 24 months (CR)	6 months GE1: 43% GE2: 25% GC: 67%	12 months GE1: 0% GE2: 13% GC: 50%	6 months GE1:0.62 GE2:1.11	12 months GE1:1.53 GE2:1.13
5. Lin et al. (2009)	301 GE: 70 GC: 231	Duluth + CBT Group	12-18 weeks	6 and 9 months (CR)	6 months GE:34.3% GC:34.2%	9 months GE:27.1% GC:21.2%	6 months -0.003	9 months -0.19
6. Taylor and Maxwell (2009)	629 GE:317 GC: 312	Duluth Group	5 days	6 and 12 months (OR)	6 months GE:65.9% GC:65.7%	12 months GE:68.6% GC:69.6%	6 months -0.005	12 months 0.03

46 initial effect sizes, with a sample of 18,941 subjects, the δ weighted mean was 0.41, 90% CI [-0.12, 0.94], that is, the results revealed a non significant positive treatment effect. What is more, treatment may have considerable negative effects: as much as a 6% increase in the recidivism rate. Notwithstanding this, the effect size is not necessarily null, $ES_{counternull}$ = 0.82, that is, there is as much evidence to support a un null treatment effect as there is to show a 38% intervention efficacy rate. Nevertheless, the credibility interval of the effect size suggested the existence of further moderators. Thus, the studies were classified according to the duration of the follow-up period,

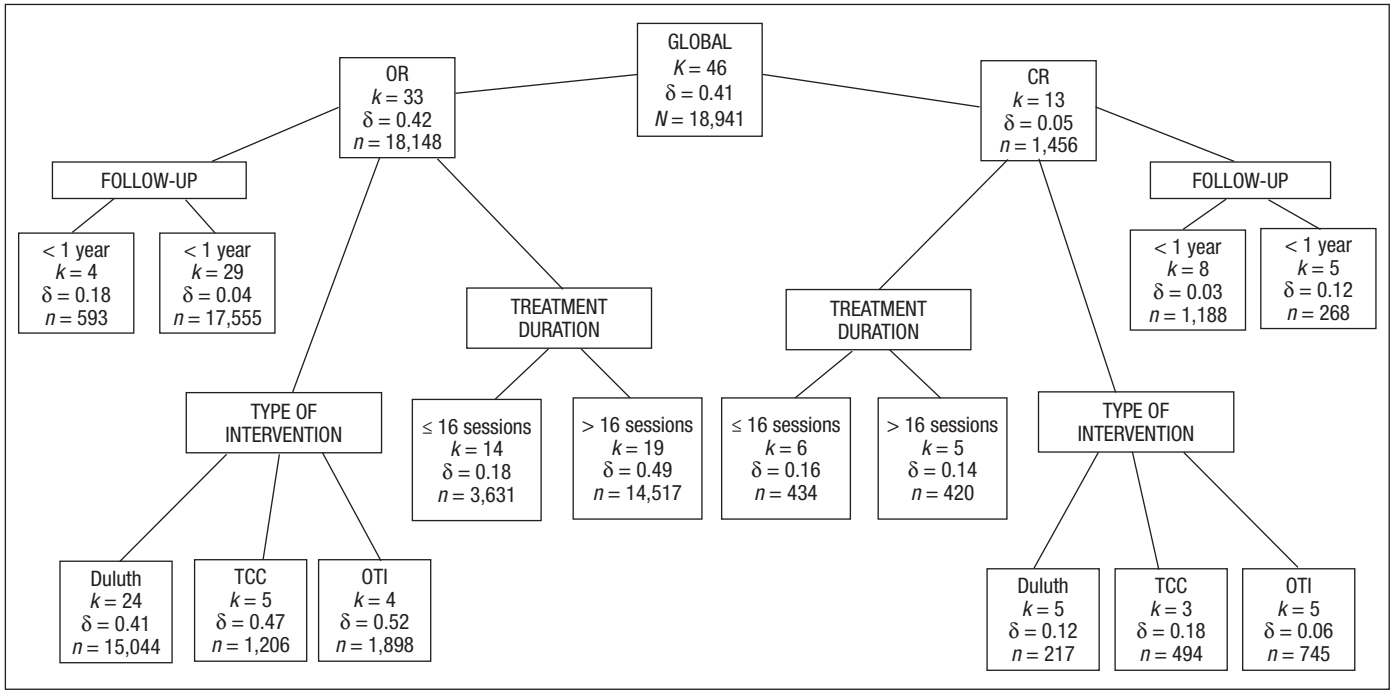

Figure 1. Meta-analytical model for examining recidivism reports, duration of follow-up, and intervention type as moderators
Note. OR = Official reports; CR = Couple reports; δ = weighted mean effect size; *k* = number of effect sizes for each analysis; CBT = cognitive-behavioural treatment programmes; OTI = other types of intervention. In the analysis of the moderator treatment duration the results of the study of Lin et al. (2009) were not included since the duration of the intervention was not accurately specified (from 12 to 18 weeks).

since previous reviews claim it is one of the main moderators of criminal recidivism (Gondolf, 2000, 2002; Redondo et al., 2001) and since prior analysis have found differences in recidivism as measured by ORs or CRs in relation to the follow-up period.

Effects due to the variable of the measure of recidivism

The meta-analysis of the ORs with an *n* of 18,148 batterers found a non significant positive weighted mean treatment effect, δ = 0.42, 90% CI [-0.07, 0.91]. However, as the magnitude of the size was close to medium, the counternull effect was computed, $ES_{counternull}$ = 0.84, indicating that the probability of finding a null recidivism treatment effect is equal to getting a 39% success rate. As for the meta-analysis of the CRs with a total population of 1,456 batterers, treatment effect was not significant, δ = 0.05, 90% CI [-0.52, 0.63]. Moreover, the CI showed that treatment might have negative or even detrimental effects leading to increased recidivism rates reaching 27.8% (δ = -0.52).

Given that the confidence intervals for both the ORs and CRs measures of recidivism had a negative lower limit, i.e., though treatment had a positive weighted mean effect it may also have had a negative effect, further search for moderators was undertaken to identify the variables underlying the difference in effects.

Effects due to the measure of recidivism and follow-up time. In line with previous studies that assert that recidivism occurs primarily within the first two years, and in the case of domestic violence in the first six months (Gondolf, 2000, 2002; Redondo et al., 2001), the two follow-up categories were coded: < 12 months and ≥ 12 months. The results revealed an effect size for recidivism as measured by ORs at 12-month follow-up (*k* = 4) of δ = 0.18, 90% CI [-0.36, 0.65], that is, a positive but non significant mean effect size that may even be highly negative (succinctly, the recidivism rate may rise to 17.7%), whereas for a follow-up period longer than 12 months (*k* = 29) of δ = 0.04, 90% CI [-0.45, 0.53] there is a not significant or near null effect that can be very negative (i.e., it can lead to a 22.0% increase in the recidivism rate). Likewise, the treatment effects in the measure of recidivism in CRs revealed a non significant mean effect close to 0, with negative effects of 28.3% at 12-month follow-up (*k* = 8), δ = 0.03, 90% CI [-0.59, 0.65] and non significant positive and potentially negative effects reaching 18.2% in the follow-up period longer than 12 months, (*k* = 5), δ = 0.12, 90% CI [-0.37, 0.61]. These results indicated the confidence intervals of δ had a negative lower limit both in the OR and CR measures and in both follow-up periods, thus an analysis of the moderators was conducted. At this point, the moderator type of analysis was the best candidate for analysis, since numerous studies have reported that treatment type has effects on recidivism, the highest effects being observed in cognitive-behavioural treatments (Redondo et al., 1999, 2001, 2002). Bearing in mind that the effects size were not significant and overlapped (*U*1 = .00 and .04, for the overlapping distribution of the short and long-term follow-up in the ORs and CRs, respectively), the variance for the short and long-term follow-up in the ORs and CRs was small (S^2 = 0.11, and 0.14 for short-term follow-up in the ORs and CRs, respectively, and 0.09 and 0.09 for the long-term follow-up in the ORs and CRs, respectively), and the distribution of treatment types for each of the follow-up periods would entail several cells with insufficient studies, the ORs and the CRs were aggregated for the analysis of treatment types (cognitive-behavioural, Duluth, and others).

Effects due to the measure of recidivism and the type of intervention. The results of the meta-analysis exhibited a non significant positive mean effect in the ORs for the Duluth Model treatment type (*k* = 24), δ = 0.41, 90% CI [-0.09, 0.92]; a non significant positive mean effect for the CBT programmes (*k* = 5), δ = 0.47, 90% CI [-0.20, 1.14]; and a significant positive mean effect and a moderate size for the OTI (*k* = 4), δ = 0.52, 90% CI [0.29, 0.75]. As the effect sizes were not significant for the Duluth Model or the CBT programmes, but were approximately a medium size, they were contrasted with a null effect. Counternull effect size was 0.82 for the Duluth Model, that is, data suggested that the probability of the null effect for the Duluth Model in reducing recidivism was equal to a 38% efficacy rate. As for the CBT programmes, a $ES_{counternull}$ = 0.94 supported as much a null effect as a 42% success rate. In the measure of recidivism based on CRs the results found a non significant positive mean treatment effect of the Duluth Model (*k* = 5), δ = 0.12, 90% CI [-0.06, 0.30]; a non significant positive mean effect of the CBT programmes (*k* = 3), δ = 0.18, 90% CI [-0.08, 0.44]; and a non significant negative mean effect of the OTI, (*k* = 5), δ = -0.06, 90% CI [-0.81, 0.69]. In other words, treatment may even have negative effects on recidivism rates reaching 37.5%.

Effects due to the measure of recidivism and the duration of the intervention. The measure of recidivism based on ORs in the brief interventions had a non significant positive weighted mean effect (*k* = 14), δ = 0.18, 90% CI [-0.58, 0.94]; and long-term programmes had a statistically significant weighted mean positive effect (*k* = 19), δ = 0.49, 90% CI [0.05, .93), i.e., a medium positive effect size.

The treatment effects as measured by CRs in brief interventions had a non significant weighted mean positive effect (*k* = 6), δ = 0.16, 90% CI [-0.07, 0.39]; long-term treatment programmes had a non significant weighted mean positive effect (*k* = 5), δ = 0.14, 90% CI [-0.09, 0.37].

Discussion

This meta-analysis has certain limitations that should be borne in mind when extrapolating or generalizing the results to other populations. First, the effects of a meta-analysis may be inadvertently contaminated by other variables that preclude the estimate of an effect size due to treatment. Second, details of several of the moderators initially selected for this study were not fully reported or were not accurately measured by the studies selected for this meta-analysis. Third, the measures for batterer treatment efficacy based on ORs and CRs were not entirely accurate since they entailed a margin error in the estimates of the recidivism rates (hidden victimization/undetected delinquency). Most of the interventions were evaluated by the authors themselves who were conscious that the continuity of their intervention programme depended on positive outcomes which may undermine the reliability of the evaluation (thus, the detected outliers were the 3 effect sizes of the same author with a positive effect 2*SD* > *M*, whereas the interventions with the highest negative effect sizes corresponded to the external assessments of Jones and Gondolf, 2002). Taking these limitations into account in generalizing the results, the following conclusions may be drawn:

On the whole, the treatment of batterers had a positive but non statistically significant effect. As for some specific treatments, it may also have had considerably negative effects both in ORs and CRs. Nevertheless, this does not imply that the batterers' treatment efficacy rate is null, given that the probability of an effect being null is equal to a 38% efficacy rate which is quite a respectable efficacy. Hence, the evidence remains inconclusive and sharp conclusions cannot be drawn (Eckhardt et al., 2013; Smedslund, Dalsbo, Steiro, Winsvold, & Clench-Aas, 2011).

Treatment efficacy was not sensitive to the moderator duration of the follow-up (short-term vs. long-term) in the ORs or CRs. In other words, the follow-up period was not a differential indicator of treatment efficacy, which contradicts the findings of Gondolf that the greatest recidivism rate occurs during the first months (Gondolf, 2000, 2002).

The 'type of intervention' moderator (Duluth Model, CBT or OTI Programmes) had no significant effects in CRs or ORs for the Duluth Model and the CBT Programmes though the effects were significant for the OTIs. The lack of a significant treatment effect in the Duluth Model and CBT Programmes corroborated the findings of Babcok et

al. (2004). However, the contrast of the observed effect size with a null effect size (0 efficacy rate) showed that the evidence for supporting a null recidivism efficacy rate in the ORs for the Duluth Model and CBT treatment programmes was the same as a 38% and 42% efficacy rate, respectively. As for the positive effects of the OTIs, these rest on psychological-psychiatric treatment (Coulter & VandeWeerd, 2009) in which the main aim of treatment is psychopathology, not gender violence. Thus, this intervention is the most apt for addressing the needs of batterers. Consequently, further studies are required on the effects of treatment type to identify those variables that are mitigating its potential effects, such as the control of treatment adherence (Arce, Fariña, Carballal, & Novo, 2009), the psychological adjustment (Lila, Gracia, & Murgui, 2013) or the motivation for change (Eckhardt et al., 2013).

The 'duration of the intervention' moderator (brief vs. long) had no significant effect in the CRs whereas in ORs, long-term interventions had a significant mean effect size though no significant mean effect size was found in brief interventions. Thus, long-term interventions were more efficacious in the ORs, i.e., they officially reduce recidivism, but do not appear to do so in the daily life of couples (CRs).

Though the treatment of batterers may have negative effects on recidivism rates, treatment should be aimed at achieving positive effects, i.e., the implementation of treatment programmes that entail negative effects is entirely unacceptable. This underscores the need to identify the characteristics of treatment efficacy studies with considerable negative effect sizes, which in this meta-analysis were as follows: Group 3 of The San Diego Navy Experiment (Dunford, 2000), that was characterized for being brief (< 16 sessions) and CBT being applied individually in a military base; some of the Jones and Gondolf's (2002) studies that were characterized as group interventions based on the Duluth Model; and the study by Lin et al. (2009) that was defined as a mixed (Duluth Model and CBT) group treatment programme. In short, no nexus was found between the treatment programmes which would indicate other causes are responsible for the negative treatment effects. These findings suggest that further research is required to ascertain the causes underlying these negative effects.

In conclusion, overall, the treatment of batterers is not efficacious, though some programmes were (k = 16 for the positive effect of a small effect size or larger than ≥ 0.20) or had negative effects (k = 7). Of the moderators, only the type of intervention (i.e., OTIs) and the duration of the intervention (long-term) were significant, i.e., interventions adapted to the batterers' needs (OTI: psychological-psychiatric programme for batterers with psychopathology) and long-term interventions, which would indicate that (toxic) cognition that sanctions domestic violence is highly resistant to treatment. Nonetheless, the results remain inconsistent and further studies are required to assess the efficacy of batterer treatment programmes, i.e., to examine moderators that may explain why some batterers respond to treatment yet others fail to do so under similar treatment programmes. This calls for authors, reviewers, and editors to provide explicit details regarding the treatment contents, techniques, and methods. This study has focused on certain variables that are crucial for the assessment of treatment, but have often been neglected in the literature, since initially they have been considered of minor importance though the results of this meta-analysis have shown they are robust, e.g., the techniques and methods applied that involve active, focused, collaborative learning (the principle of responsibility), the implementation of treatment programmes by specialized and trained staff, and the implementation of additional judicial measures (Lila, García, & Lorenzo, 2010; McGuire et al., 2000).

Conflicts of interest

The authors of this article declare no conflicts of interest.

Acknowledgements

This research has been sponsored by a grant of the Spanish Ministry of Science and Innovation to the project "Reeducación de penados por violencia de género: Implementación y evaluación de programas de tratamiento" (Ref.: EDU2011-24561).

References

[References marked with an asterisk indicate studies included in the meta-analysis]

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

Arce, R., & Fariña, F. (2010). Diseño e implementación del programa Galicia de reeducación de maltratadores: Una respuesta psicosocial a una necesidad social y penitenciaria [Design and implementation of the Galician program for batterers' re-education: A psychosocial answer to a social and penitentiary need]. *Intervención Psicosocial, 19*, 153-166.

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.

Babcock, J. C., Green, C. E., & Robie, C. (2004). Does batterers' treatment work? A meta-analytic review of domestic violence treatment. *Clinical Psychology Review, 23*, 1023-1053.

*Babcock, J., C., & Steiner, R. (1999). The relationship between treatment, incarceration and recidivism of battering: A program evaluation of Seattle's coordinated community response to domestic violence. *Journal of Family Psychology, 13*, 46-59.

Beelmann, A., & Lösel, F. (2006). Child social skills training in developmental crime prevention: Effects on antisocial behavior and social competence. *Psicothema, 18*, 603-610.

*Bennett, L., Call, C., Flett, H., & Stoops, C. (2005). *Program completion, behavioral change and re-arrest for the batterer intervention system of Cook County, Illinois*. Chicago, IL: Illinois Criminal Justice Information Authority. Retrieved from http://tigger.uic.edu/~lwbenn/DVPEP/cjia.pdf

Bennett, L., & Williams, O. (2001). Intervention program for men who batter. In C. Renzetti, & J. Edleson (Eds.), *Sourcebook on violence against women* (pp. 261–277). Thousand Oaks, CA: Sage.

*Bowen, E., Gilchrist, E. A., & Beech, A. R. (2005). An examination of the impact of community-based rehabilitation on the offending behaviour of male domestic violence offenders and the characteristics associated with recidivism. *Legal and Criminological Psychology, 10*, 189-209.

Chereji, S. V., Pintea, S., & David, D. (2012). The relationship of anger and cognitive distortions with violence in violent offenders' population: A meta-analytic review. *The European Journal of Psychology Applied to Legal Context, 4*, 59-77.Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: LEA.

Collie, R. M., Vess, J., & Murdoch, S. (2007). Violence-related cognition: Current research. In T. A. Gannon, T. Ward, A. R. Beech, & D. Fisher (Eds.), *Aggressive offenders' cognition: Theory, research, and practice* (pp. 179-197). Chichester, UK: John Wiley and Sons.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: LEA.

Cortina, J. M., & Dunlap, W. P. (1997). On the logic and purpose of significance testing. *Psychological Methods, 2,* 161-172.

*Coulter, M., & VandeWeerd, C. (2009). Reducing domestic violence and other criminal recidivism: Effectiveness of a multilevel batterers intervention program. *Violence and Victims, 24*, 139-152.

Davis, R. C., & Taylor, B. G. (1999). Does batterer treatment reduce violence? A synthesis of the literature. *Women and Criminal Justice, 10*, 69-93.

*Davis, R. C., Taylor, B. G., & Maxwell, C. D. (1998). Does batterer treatment reduce violence? A randomized experiment in Brooklyn. *Justice Quarterly, 18*, 171-201.

*Dobash, R., Dobash, R. E., Cavanagh, K., & Lewis, R. (1996). Reeducation programs for violent men: An evaluation. *Research Findings, 46*,309-322.

*Dunford, F. W. (2000). The San Diego Navy experiment: An assessment of interventions for men who assault their wives. *Journal of Consulting and Clinical Psychology, 68*, 468-476.

Dutton, D. G., Bodnarchuk, M., Kropp, R., Hart, S. D., & Ogloff, J. P. (1997). Client personality disorders affecting wife assault post-treatment recidivism. *Violence and Victims, 12*, 37-50.

Eckhardt, C. I., Murphy, C. M., Whitaker, D. J., Sprunger, J., Dykstra, R., & Woodard, K. (2013). The effectiveness of intervention programs for perpetrators and victims of intimate partner violence. *Partner Abuse, 4*, 196-231.

Fariña, F., Arce, R., & Novo, M. (2008). Neighborhood and community factors: Effects on deviant behavior and social competence. *The Spanish Journal of Psychology, 11*, 78-84.

*Feder, L. R., & Dugan, L. (2004). *Testing a court-mandated treatment program for domestic violence offenders: The Broward experiment*. Washington, DC: National Institute of Justice. Retrieved from https://www.ncjrs.gov/pdffiles1/nij/199729.pdf

Feder, L. R., & Wilson, D. B. (2005). A meta-analytic review of court-mandated batterer intervention programs: Can courts affect abusers' behavior? *Journal of Experimental Criminology, 1*, 239-262.

Frick, R. W. (1996). The appropriate use of null hypothesis testing. *Psychological Methods, 1,* 379-390.

Gondolf, E. W. (2000). Reassault at 30-months after batterer programs intake. *International Journal of Offender Therapy and Comparative Criminology, 44*, 111-128.
Gondolf, E. W. (2002). *Batterer intervention systems: Issues, outcomes, and recommendations*. Thousand Oaks, CA: Sage.
Gondolf, E. W. (2004). Evaluating batterer counseling programs: A difficult task showing some effects and implications. *Aggression and Violent Behavior, 9*, 605-631.
Gracia, E., Herrero, J., Lila, M., & Fuente, A. (2009). Perceived neighborhood social disorder and attitudes toward domestic violence against women among Latin-American immigrants. *The European Journal of Psychology Applied to Legal Context, 1*, 25-43.
Hamberger, L. K., & Hastings, J. E. (1993). Court-mandated treatment of men who assault their partner: Issues, controversies, and outcomes. En N. Z. Hilton (Ed.), *Legal responses to wife assault* (pp. 96-121). Newbury Park, CA: Sage.
Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA. Academic Press.
Holtzworth-Munroe, A., Meehan, J. C., Herron, K., Rehman, U., & Stuart, G. L. (2000). Testing the Holtzworth-Munroe and Stuart (1994) batterer typology. *Journal of Consulting and Clinical Psychology, 68*, 1000-1019.
Hutchings, J. N., Gannon, T. A., & Gilchrist, E. (2010) A preliminary investigation of a new pictorial method of measuring aggression-supportive cognition among young aggressive males. *International Journal of Offender Therapy and Comparative Criminology, 54*, 236-249.
Isorna, M., Fernández-Ríos, L., & Souto, A. (2010). Treatment of drug addiction and psychopathology: A field study. *The European Journal of Psychology Applied to Legal Context, 2*, 3-18.
*Jenkins, J. A., & Menton, C. (2003). The relationship between incarcerated batterers' cognitive characteristics and the effectiveness of behavioral treatment. *The Peer-Reviewed Journal of the American Correctional Association, 28*, 1-27.
*Jones, A. S., & Gondolf, E. W. (2002). Assessing the effect of batterer program completion on reassault: An instrumental variables analysis. *Journal of Quantitative Criminology, 18*,71-97.
Kraemer, H. C., & Andrews, G. (1982). A non-parametric technique for meta-analysis effect size calculation. *Psychological Bulletin, 91*, 404-412.
*Labriola, M., Rempel, M., & Davis, R. C. (2005). *Testing the effectiveness of batterer programs and judicial monitoring. Results from a randomized trial at the Bronx Misdemeanor Domestic Violence Court*. New York, NY: Center for Court Innovation. Retrieved from http://www.courtinnovation.org/sites/default/files/battererprogramseffectiveness.pdf
Levesque, D. A., & Gelles, R. J. (1998, July). *Does treatment reduce recidivism in men who batter? A meta-analytic evaluation of treatment outcome*. Paper presented at Program Evaluation and Family Violence Research: An International Conference, Durham, NH.
Lila, M., García, A., & Lorenzo, M. (2010). *Manual de intervención con maltratadores* [Intervention with batterers. Manual]. Valencia: Universitat de València.
Lila, M., Gracia, E., & Murgui, S. (2013). Psychological adjustment and victim-blaming among intimate partner violence offenders: The role of social support and stressful life events. *The European Journal of Psychology Applied to Legal Context*, *5*, 147-153.
Lila, M., Oliver, A., Galiana, L., & Gracia, E. (2013). Predicting success indicators of an intervention programme for convicted intimate-partner violence offenders: The Contexto Programme. *The European Journal of Psychology Applied to Legal Context*, *5*, 73-95.
*Lin, S., Su, C., Chou, F. H., Chen, S., Huang, J., Wu, G. T., … Chen, C. (2009). Domestic violence recidivism in high-risk Taiwanese offenders after the completion of violence treatment programs. *Journal of Forensic Psychiatry & Psychology, 20*, 458-472.
Martinson, R. (1974). What works? Questions and answers about prison reform. *The Public Interest, 10*, 22-54.
Maruna, S. (2004). Desistance and explanatory style: A new direction in the psychology of reform. *Journal of Contemporary Criminal Justice, 20*, 184-200.
Maruna, S., & Copes, H. (2005). What have we learned in five decades of neutralization research? *Crime and Justice: A Review on Research, 32*, 221-320.
Maruna, S., & Mann, R. E. (2006). A fundamental attribution error? Rethinking cognitive distortions. *Legal and Criminological Psychology, 11*, 155-177.
McGuire, J., Mason, T., & O'Kane, A. (2000). Effective interventions, service and policy implications. In J. McGuire, T. Mason, & A. O'Kane (Eds.), *Behavior, crime and legal processes. A guide for forensic practitioners* (pp. 289-314). Chichester, UK: John Wiley and Sons.
*Murphy, C. M., Musser, P. H., & Maton, K. I. (1998). Coordinated community intervention for domestic abusers: Intervention system involvement and criminal recidivism. *Journal of Family Violence, 13*, 263-284.
Novo, M., Fariña, F., Seijo, D., & Arce, R. (2012). Assessment of a community rehabilitation programme in convicted male intimate-partner violence offenders. *International Journal of Clinical and Health Psychology, 12*, 219-234.
O'Leary, K. D., Barling, J., Arias, I., Rosenbaum, A., Malone, J., & Tyree, A. (1989). Prevalence and stability of physical aggression between spouses: A longitudinal analysis. *Journal of Consulting and Clinical Psychology, 57*, 263-268.
*Pérez, M. Giménez-Salinas A., & Juan, M. (2012). *Evaluación del programa "Violencia de Género: Programa de Intervención para Agresores", en medidas alternativas* [Evaluation of the program "Violence against women: Program for the intervention with offenders", in alternative measures]. Madrid: Instituto de Ciencias Forenses y de la Seguridad (ICFS) y Secretaría General de Instituciones Penitenciarias. Retrieved from http://www.institucionpenitenciaria.es/web/export/sites/default/datos/descargables/descargas/VDG_EVALUACION_AUTONOMA_NIPO.pdf
Redondo, S., Sánchez-Meca, J., & Garrido, V. (1999). Tratamiento de los delincuentes y reincidencia: Una evaluación de la efectividad de los programas aplicados en Europa [Offender treatment and recidivism: An evaluation of the effectiveness of the programmes applied in Europe]. *Anuario de Psicología Jurídica, 5*,11-37.
Redondo, S., Sánchez-Meca, J., & Garrido, V. (2001). Treatment of offenders and recidivism: Assessment of the effectiveness of programmes applied in Europe. *Psychology in Spain, 5*, 47-62.
Redondo, S., Sánchez-Meca, J., & Garrido, V. (2002). Los programas psicológicos con delincuentes y su efectividad: La situación europea [Psychological programmes with offenders and their effectiveness: The European situation]. *Psicothema, 14*, 164-173.
Rosenfeld, B. D. (1992). Court ordered treatment of spouse abuse. *Clinical Psychology Review, 12*, 205-226.
Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. Journal of Educational Psychology, *74,* 166-169.
Rosenthal, R., & Rubin, D. B. (1994). The counternull value of an effect size: A new statistic. *Psychological Science, 5*, 329-334.
*Saunders, D. G. (1996). Feminist-cognitive-behavioral and process-psychodynamic treatments for men who batter: Interaction of abuser traits and treatment models. *Violence and Victims, 11*, 393-414.
Smedslund, G., Dalsbo, T. K., Steiro, A., Winsvold, A., & Clench-Aas, J. (2011). Cognitive behavioural therapy for men who physically abuse their female partner (Review). *The Cochrane Database of Systematic Reviews, 4*, Article No. CD006048. Retrieved from http://www.cochranelibrary.com
*Stith, S. M., Rosen, K. H., & McCollum, E. E. (2004). Treating intimate partner violence within intact couple relationships: Outcomes of multi-couple versus individual couple therapy. *Journal of Marital and Family Therapy, 30*, 305-318.
*Taylor, B. G., & Maxwell, C. D. (2009). *The effects of a short-term batterer treatment program for detained arrestees: A randomized experiment in the Sacramento County, California jail*. Sacramento, CA: Department of Justice, Crime and Violence Prevention Center. Retreieved from https://www.ncjrs.gov/pdffiles1/nij/grants/228275.pdf
*Tollefson, D. R., & Gross, E. (2006). Predicting recidivism following participation in a batterer treatment program. *Journal of Social Service Research, 32*, 39-62.
*Tollefson, D. R., Webb, K., Shumway, D., Block, S. H., & Nakamura, Y. (2009). A mind-body approach to domestic violence perpetrator treatment: Program overview and preliminary outcomes. *Journal of Aggression, Maltreatment &Trauma, 18*, 17-45.
Wormith, J. S., & Olver, M. E. (2002). Offender treatment attrition and its relationship with risk, responsivity and recidivism. *Criminal Justice and Behavior, 29*, 447-471.