

Anuario de Psicología Jurídica

ISSN: 1133-0740

ISSN: 2174-0542

Colegio Oficial de la Psicología de Madrid

Fernández-Fernández, Román; Navas, María Patricia; Sobral, Jorge
What is Known about the Intervention with Gender Abusers? A Meta-analysis on Intervention Effectiveness
Anuario de Psicología Jurídica, vol. 32, núm. 1, 2021, pp. 23-31
Colegio Oficial de la Psicología de Madrid

DOI: <https://doi.org/10.5093/apj2021a17>

Disponible en: <https://www.redalyc.org/articulo.oa?id=315070367003>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en [redalyc.org](https://www.redalyc.org)

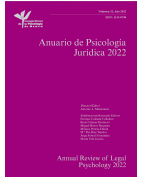
 redalyc.org

Sistema de Información Científica Redalyc
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto



Anuario de Psicología Jurídica 2022

<https://journals.copmadrid.org/apj>



What is Known about the Intervention with Gender Abusers? A Meta-analysis on Intervention Effectiveness

Román Fernández-Fernández, María Patricia Navas, and Jorge Sobral

Universidad de Santiago de Compostela, Spain

ARTICLE INFO

Article history:

Received 30 September 2020

Accepted 8 February 2021

Available online 12 April 2021

Keywords:

Effectiveness

Gender abuser

Gender violence

Intervention programs

Recidivism

Palabras clave:

Eficacia

Maltratador

Violencia de género

Programas de intervención

Reincidencia

ABSTRACT

In this research, a meta-analytic study is proposed with the aim of determining the state of the art of the efficacy of treatment programs for gender aggressors across different variables. Doing so is relevant in view of the inconsistency in the results of previous meta-analyses and the proliferation of new intervention programs for gender aggressors. The results of general meta-analysis show a positive, though not significant, effect of the intervention. The analysis of moderating variables shows a positive, significant, and close-to-moderate effect size for interventions lasting fewer than 16 weeks/sessions ($\delta_w = .39$) and follow-up periods greater than or equal to 12 months ($\delta_w = .38$), and a positive, significant and moderate size for interventions using cognitive-behavioral therapy ($\delta_w = .57$). However, along the same lines, as previous meta-analyses, no definite conclusions can be drawn about the effectiveness of this type of program.

¿Qué se sabe acerca de la intervención con maltratadores de género? Un metaanálisis sobre su efectividad

RESUMEN

En este artículo se propone un estudio metaanalítico con objeto de conocer la situación actual de la eficacia de los programas de tratamiento de agresores de género mediante diferentes variables. El estudio es obligado a tenor de la incongruencia de los resultados de metaanálisis anteriores y la proliferación de nuevos programas de intervención para agresores de género. Los resultados del metaanálisis general muestran un efecto positivo, aunque no significativo, de la intervención. El análisis de las variables moderadoras presenta un tamaño del efecto positivo, significativo y casi moderado para intervenciones de una duración inferior a 16 semanas/sesiones ($\delta_w = .39$) y períodos de seguimiento de 12 meses o más ($\delta_w = .38$) y un tamaño positivo, significativo y moderado para las intervenciones que utilizan terapia cognitivo-conductual ($\delta_w = .57$). No obstante, como en el caso de los metaanálisis previos, no pueden establecerse conclusiones definitivas sobre la efectividad de este tipo de programas.

Violence is inherent to life in society and can take various forms that show its complexity when viewed closely; in this sense, gender violence represents one of the faces of violence as a social issue (Boira & Marcuellos, 2014). The relevance of this casuistry is such that many authors label it as one of the greatest social and public health problems, both due to its magnitude and the severity of its consequences (Cáceres, 2011; Cuenca et al., 2015).

In line with this concern, some studies place the prevalence of this problem at around 30% of women (Ferrer-Pérez & Bosch-Fiol, 2016; Gracia & Lila, 2015; Organización Mundial de la Salud [World Health Organization] [OMS, 2013, 2016]) and the scientific community has echoed, as shown by the bibliometric analysis of research on gender

violence by Arias et al. (2016), registering more than 16,000 related scientific publications. In this sense, the need to provide an adequate response becomes imperative and, although historically it has been directed fundamentally towards the victims, there are actions that focus on aggressors, such as intervention programs with abusers; these programs arise as a consequence of the social recognition of gender violence as a problem of first magnitude (Mackay et al., 2015).

The first programs emerged in the United States and Canada between 1970 and 1980; then they spread throughout Europe from the 80s, e.g., in 1984 in Germany, in 1986 in Norway, or in 1990 in Spain (Feder & Wilson, 2005; Ferrer-Pérez et al., 2016; Subirana & Pueyo, 2013). This tradition of intervention has grown to such an

Cite this article as: Fernández-Fernández, R., Navas, M. P., Sobral, J. (2022). What is known about the intervention with gender abusers? A meta-analysis on intervention effectiveness. *Anuario de Psicología Jurídica*, 32, 23-31. <https://doi.org/10.5093/apj2021a17>

Correspondence: roman.ffdez@gmail.com (R. Fernández-Fernández).

extent that today it is estimated that there are over 2,500 programs in the United States and over 500 in Europe (Geldschläger et al., 2009) and, in parallel, research tradition on their characteristics has also suffered a considerable increase (Giné et al., 2015).

Along these lines, two have been the most widely used models in this area: the Duluth model and cognitive-behavioral therapy (CBT). The Duluth model is a program that was developed by Pence and Paymar (1983) in Duluth, Minnesota, that combines a feminist perspective with a psychoeducational approach. The basis on which it is based is that the cause of violence comes from patriarchal and male chauvinist learning that occurs in society, which urges and naturalizes the use of violence, giving rise to an asymmetry of power based on gender. Therefore, the efforts of this program are aimed at identifying people's behaviors to exercise power and control, promoting behavioral and attitudinal changes; for this it uses two tools, the Power and Control Wheel (negative behaviors and attitudes) and the Equality Wheel (positive vision of roles within relationships), which act as facilitators of the identification of causes of gender violence.

For its part, the CBT model was developed by Sonkin et al. (1985) and, although there are variations of it, they are collected under the same nomenclature. The common denominator of these programs is that their objective is focused on the modification of maladaptive behaviors that precede violence. Violence remains functional for the subject, since it acts by reducing bodily tension, achieving obedience from the victim or temporarily ending an uncomfortable situation (Sonkin et al., 1985). Since this is a learned behavior, alternative non-violent behaviors can be inserted in its place, mainly through techniques such as training in social skills or anger management techniques (Babcock et al., 2004).

On the other hand, meta-analyses have created a category of Other Types of Intervention (OTI) for all those programs that do not correspond to previous categories. Under this label, highly heterogeneous programs are collected, from mixtures of different interventions (for example, Duluth model and CBT), to psychodynamic orientation programs, anger control programs or mind body bridging (Arias et al., 2013). However, treatment labels are sometimes misleading, and membership in one category or another is increasingly diffuse since, for example, some CBTs implement the theme of values, or some programs that use the Duluth model contemplate aspects related to the learning of violence (Babcock et al., 2004).

Whatever the type of intervention, one of the aspects that has received the most attention is the concern for their effectiveness. In this sense, one of the main objectives of the programs is to reduce the probability of recidivism, which is therefore one of the variables used to measure the success of the interventions (Coulter & VandeWeerd, 2009). In this regard, there are a series of previous meta-analyses whose results show limited efficacy (Arias et al., 2013; Babcock et al., 2004; Feder & Wilson, 2005; Smedslund et al., 2011).

The meta-analysis carried out by Babcock et al. (2004) gathered a total of 22 studies that evaluated the effectiveness of the programs through different variables: (a) percentage of recidivism through official records (OR) or partner reports (PR); (b) type of study (experimental or quasi-experimental); (c) type of intervention (Duluth, CBT or OTI); (d) duration of the intervention (short or long); (e) follow-up period (greater or fewer than 12 months). Likewise, they also recorded the rate of individuals who successfully completed treatment but, due to the heterogeneity of criteria used across the studies, it was decided not to include this variable in the analysis. Their results show that there are no significant differences between the Duluth model and CBT, and that the effects of intervention programs on recidivism are small; in the best of cases, their results suggest that treated abusers would have an improvement of one third of a standard deviation compared to untreated ones in terms of recidivism ($d = 0.34$).

For their part, Feder and Wilson (2005) focused on the type of intervention, differentiating between (a) feminist-based psychoeducational approaches, (b) CBT, and (c) studies that combined both approaches. The results of their work were diverse depending on the source of information on recidivism (OR or PR) or the study design (experimental or quasi-experimental) but, in any case, the conclusions they reached were that the results left doubts on the effectiveness of this type of program. On the other hand, Smedslund et al. (2011) reviewed programs that used CBT, including six randomized experiments, but their conclusions were that the evidence was insufficient to say anything about the effectiveness of the intervention in reducing or eliminating violence.

Finally, the meta-analysis of Arias et al. (2013) collected a total of 22 articles, with which they computed 49 effect sizes and a cumulative sample of 18,941 abusers. In terms of method, the criteria used were very similar to the meta-analysis previously performed by Babcock et al. (2004), and they came to the conclusion that in general treatment has a positive, though not significant, effect and can sometimes even have significant negative effects; along these lines, they affirm that the evidence is not enough to reach robust conclusions.

The results of these meta-analytical studies reveal the limited evidence that exists regarding the effectiveness of this type of program and, therefore, highlight the challenge faced by professionals and researchers in this field (Lila et al., 2020). An effort to overcome these limitations is found in the study of moderating variables, such as the duration of the intervention program or the follow-up period. Regarding the first variable, given that gender violence is based on what Maruna (2004) calls toxic cognitions, that is, dysfunctional thoughts that are characterized by being internal, stable and global, brief interventions are related to a lower effectiveness; along these lines, long interventions facilitate the consolidation and generalization of acquired skills (Arias et al., 2013). On the other hand, the follow-up period has been related as one of the most relevant moderators in recidivism, since it usually occurs in the first two years and, in the case of gender-based violence, in the first six months (Gondolf, 2002; Redondo et al., 2001). However, the inclusion of these moderating variables in the studies has not been enough to reliably explain the results, so researchers consider various hypotheses in this regard, mainly related to possible methodological deficiencies or the characteristics of the intervention program itself.

As for the shortcomings of a more methodological nature, there is the scarcity of experimental designs and low reliability of measurements. On the one hand, although quasi-experimental designs can produce effect sizes comparable to experimental ones (Heinsman & Shadish, 1996), the latter allow greater confidence in conclusions in terms of causality relationships (Babcock et al., 2004); however, random assignment of subjects to conditions of treatment or no treatment implies both an ethical and a practical problem in this field of intervention. On the other hand, the measure about recidivism is usually carried out through OR, PR, or even self-reports of the abuser himself, but reliability of these measures as an index of recidivism has been questioned (Novo et al., 2012). Along these lines, the meta-analysis by Babcock et al. (2004) shows very different recidivism rates, being 21% according to ORs and 35% according to PRs.

Furthermore, there are the aspects more related to the intervention program itself, such as high dropout rates, lack of motivation and adherence to the treatment or lack of adaptation to aggressors' personal characteristics. The high dropout rate has been considered as one of the main reasons for the lack of evidence on effectiveness (Babcock et al., 2004; Crane et al., 2015), with a range that oscillates between 15% and 60%. This high rate may be conditioned to the lack of motivation and adherence of the subject, which can be partly explained by the component of mandatory by judicial order that usually accompanies this type of measures (Ferrer-Pérez & Bosch-Fiol, 2016), or the lack of adaptation to an individual's idiosyncratic characteristics.

In this regard, some authors attribute the low efficacy of interventions to the poor adaptation that exists regarding particular needs of individuals (Cavanaugh & Gelles, 2005; Huss & Ralston, 2008). As a consequence, a line of research has been derived regarding types of aggressors to test whether there really are differences. In this sense, the typology that has received the most attention is that of Holtzworth-Munroe & Stuart (1994), which, based on three dimensions (extent of violence, severity of violence, and psychopathological aspects of an abuser) generates three groups of aggressors (violent only in the family, borderline/dysphoric, and violent in general/antisocial), to which is added a fourth (low intensity antisocial) as a result of empirical validation carried out in a community context (Holtzworth-Munroe et al., 2000).

Either way, the issue of interventions with gender abusers has given rise to an in depth discussion on their efficacy, as a consequence of which the present meta-analytic review has been proposed.

Objectives

The main objective in this study is to perform a meta-analysis to learn the state-of-the-art of the efficacy of gender abusers' treatment programs through different variables, given the inconsistency in the results of previous meta-analyses and the proliferation of new intervention programs for gender abusers. For it, this study seeks to replicate previous meta-analytical studies on the effectiveness of intervention programs with gender abusers (Arias et al., 2013; Babcock et al., 2004; Feder & Wilson, 2005; Smedslund et al., 2011), as well as increasing knowledge about it by incorporating new studies. As moderating variables, follow-up period, type of intervention, and duration of the intervention are considered. In this way, we will proceed to check whether doubts persist about the limited efficacy of the interventions or whether, on the contrary, they are effective in reducing recidivism. Lastly, it is expected that this information may prove relevant in the fields of risk assessment and interventions with abusers in general, and specifically, for the technical teams employed in prisons.

Method

Study Search

In order to achieve the objectives proposed in this meta-analytical review, the recommendations of PRISMA statement (Urrútia & Bonfill, 2010) were followed, as well as other relevant articles (Perestelo-Pérez, 2013; see also Fernández-Ríos & Buelar-Casal, 2009; Sánchez-Meca, 2010). Along these lines, 4 different strategies were carried out to locate the studies:

- Search in electronic databases of international reference (PsycInfo, Scopus, Web of Science).
- Review of previous meta-analyses on the subject.
- Checking the references of selected publications ("ancestry approach").
- Contact with relevant researchers on the subject under review.

Regarding the selection of keywords, those used in previous meta-analyses were taken as a starting point. In addition, the bibliometric analysis of research on gender violence by Arias et al. (2016), as well as the thesaurus of PsycInfo and DeCS (*Descriptores en Ciencias de la Salud*), in order to design a search equation as comprehensive as possible, were used.

As a result of this revision, the Boolean search equation was configured, resulting from the combination of the following keywords: "intimate partner violence", "domestic violence", "family violence", "violencia de género", "violencia doméstica", "violencia familiar", "batterer*", "aggressor*", "maltratador*", "agresor*", "program*", "intervention*", "treatment*", "intervención*", "tratamiento*",

"assessment*", "evaluation*", "recidivism", "effectiveness", "evaluación*", "reincidencia", "efectividad".

It should be noted that we have chosen to use only articles published in double blind peer-reviewed journals. This criterion acts objectively as a filter to ensure that only high-quality research is published by determining the validity, methodology, and contribution of the study. On the other hand, no temporal or idiomatic criteria have been used to narrow the search in order to obtain as many results as possible. Likewise, researchers focused on the topic were contacted to prevent possible publication bias.

This strategy reported a total of 752 results (244 results in PsycInfo, 247 in Scopus, and 261 in WOS); after removing duplicate articles using the Endnote bibliographic manager, a total of 482 articles were obtained; 11 additional articles were identified from previous meta-analyses. Subsequently, the remaining results were analyzed, from which a total of 26 articles were selected to form part of this meta-analysis, of which 39 effect sizes could be computed, with a total sample of 19,773 abusers.

Inclusion and Exclusion Criteria

A total of 5 criteria were used for inclusion of studies:

- Samples made up of gender abusers.
- Adult population samples, that is, those over 18 years old.
- Provide information on the sample size, both the initial and corresponding to subjects who had successfully completed the treatment.
- Provide information on characteristics of the intervention program, that is, (a) follow-up time, (b) type of intervention, and (c) duration of intervention.
- Provide information on recidivism rate through RO, for example, judicial, police, or prison records.

On the other hand, all publications were excluded that (a) were not primary studies, (b) had a sample size less than 30, (c) had a follow-up period of fewer than 6 months, (d) did not report the recidivism rate relative only to crimes related to gender violence, and (e) did not provide sufficient data to be able to calculate effect size.

Coding

Once the article selection process was completed, the results were encoded based on (a) article reference, (b) sample size of subjects who successfully completed the intervention, (c) characteristics of the intervention program (type of intervention, duration, and follow-up period), (d) recidivism rate (%), and (e) effect size (δ). Encoding was carried out by an encoder who noted the levels of each of the categories created and described for each variable by researchers and marked exactly where these qualitative variables were referred to in the text. A second encoder reviewed all of the studies using the same encoding system. Thereafter, both registered categories for each variable and degree of agreement in encodings were checked with Cohen's kappa (k). After coding primary studies, Cohen's kappa was .80, with values between .72 and .89. Then, discrepancies were discussed among researchers using a third judge in those cases where it was not possible to reach an agreement. Finally, a full inter-judge agreement was obtained ($k = 1$). Figure 1 shows flow diagram of the study selection process.

All the studies that met inclusion criteria were divided into quasi-experimental ($k = 22$) and experimental ($k = 4$), whose characteristics can be observed in Table 1a and 1b, respectively. Experimental studies are those that randomly assign subjects to "treatment" or "no treatment" conditions; quasi-experimental studies, on the other hand, either do not use a control group or, if they do, it is not equivalent to the experimental group, for example, when comparing

Table 1a. Coding of Experimental Studies. Characteristics of Quasi-experimental Studies (I)

| Reference | N | Intervention | Duration | Follow-up | Recidivism | | δ | |
|----------------------------------|-------------------------|-----------------------------------|-----------------|--------------|-------------|-------------|-------------|-------------|
| | | | | | < 12 months | ≥ 12 months | < 12 months | ≥ 12 months |
| Babcock and Steiner (1999) | 106 | Duluth | 36 weeks | 24 months | - | 8% | - | .60 |
| Bowen et al. (2005) | 86 | Duluth | 24 weeks | 11 months | 21% | - | 0 | |
| Carbajosa et al. (2017) | 210 | Psychoeducational | 9 months | 24 months | - | 8.1% | - | .59 |
| Chen et al. (1989) | 120 | Handling aggressiveness | 8 weeks | 6 months | 5% | - | .83 | - |
| Coulter and VandeWeerd (2009) | EG ₁ : 1,424 | Duluth | 8-12 weeks | 1-10 years | - | 8.8% | - | .54 |
| | EG ₂ : 9,386 | Duluth | 26 weeks | 1-10 years | - | 8.3% | - | .58 |
| | EG ₃ : 1,712 | Specialized treatment | 26 weeks-1 year | 1-10 years | - | 8.6% | - | .56 |
| Dutton et al. (1997) | 156 | Handling aggressiveness | 16 weeks | 5.2 years | - | 18% | - | .11 |
| Farzan-Kashani and Murphy (2015) | 130 | CBT | 16 weeks | 8 years | - | 25.76% | - | -.16 |
| Graña et al. (2017) | EG ₁ : 19 | CBT | 23 weeks | 12 months | - | 5.3% | - | .81 |
| | EG ₂ : 74 | CBT | 23 weeks | 12 months | - | 5.4% | - | .81 |
| | EG ₃ : 173 | CBT | 23 weeks | 12 months | - | 4% | - | .94 |
| Haggard et al. (2015) | 249 | Duluth | 35 sessions | 4.6 años | - | 15% | - | .23 |
| Hamberger and Hastings (1988) | 32 | CBT | 15 weeks | 12 months | - | 9% | - | .53 |
| Herman et al. (2014) | 103 | Duluth | 24 weeks | 9 years | - | 28.5% | - | -.24 |
| Lauch et al. (2017) | 202 | Emergency + AMEND | 6 months | 12 months | - | 7.21% | - | .65 |
| Lila et al. (2020) | 206 | CBT | 30-32 sessions | 9 months | 9.22% | - | .52 | - |
| Lila et al. (2018) | 160 | CBT | 9 months | 6 months | 7.5% | - | .63 | - |
| Murphy et al. (1998) | 235 | Duluth | 22 sessions | 12-18 months | - | 15.7% | - | .20 |
| Pérez et al. (2012) | 598 | CBT | 25 weeks | 12 months | - | 4.6% | - | .87 |
| Saunders (1996) | EG ₁ : 61 | Duluth + CBT | 32 weeks | 24 months | - | 23.2% | - | -.08 |
| | EG ₂ : 68 | Psychodynamic processes | 32 weeks | 24 months | - | 20.3% | - | .02 |
| Scott et al. (2015) | 40 | Modelo RNR | 4 months | 12 months | - | 29.3% | - | -.26 |
| Tollefson and Gross (2006) | 102 | Duluth | 20 sessions | 7-58 months | - | 18% | - | .11 |
| Tollefson and Phillips (2015) | EG ₁ : 46 | Mind-body Bridging | 20 weeks | 428 days | - | 2.3% | - | 1.19 |
| | EG ₂ : 44 | Psychoeducational | 24 weeks | 428 days | - | 10.9% | - | .42 |
| Tollefson et al. (2009) | 57 | Mind-body Bridging | 8-10 sessions | 9-27 months | - | 9% | - | .53 |
| Zarling et al. (2017) | EG ₁ : 515 | Acceptance and Commitment Therapy | 24 weeks | 1 year | - | 3.6% | - | .99 |
| | EG ₂ : 1,921 | Duluth + CBT | 24 weeks | 1 year | - | 7% | - | .67 |

Note. EG = experimental group; RNR = risk, need, and responsiveness.

Table 1b. Coding of Experimental Studies. Characteristics of Experimental Studies (II)

| Reference | n | Intervention | Duration | Follow-up | Recidivism | | δ | |
|---------------------------|-----------------------|----------------------------------|-----------------------|-----------------|-------------|-------------|-------------|-------------|
| | | | | | < 12 months | ≥ 12 months | < 12 months | ≥ 12 months |
| Davis et al. (1998) | EG ₁ : 129 | Duluth | 8 weeks | 6 and 12 months | 7% | 10% | .7 | .64 |
| | EG ₂ : 61 | Duluth | 26 weeks | 6 and 12 months | 15% | 25% | .26 | .03 |
| | CG: 186 | - | - | 6 and 12 months | 22% | 26% | - | - |
| Dunford (2000) | EG1: 168 | CBT | 36 weeks + 6 sessions | 12 months | - | 4% | - | 0 |
| | EG2: 153 | CBT | 26 weeks + 6 sessions | 12 months | - | 3% | - | .13 |
| | EG3: 173 | Rigorous monitoring intervention | 12 sessions | 12 months | - | 6% | - | -.20 |
| | CG: 150 | - | - | 12 months | - | 4% | - | - |
| Palmer et al. (1992) | EG: 30 | Psychoeducational | 10 weeks | 12-24 months | - | 10% | - | .54 |
| | CG: 26 | - | - | 12-24 months | - | 31% | - | - |
| Taylor and Maxwell (2009) | EG: 317 | Duluth | 5 days | 6 and 12 months | 65.9% | 68.6% | -.005 | .03 |
| | CG: 312 | - | - | 6 and 12 months | 65.7% | 69.6% | - | - |

Note. EG = experimental group; CG = control group.

subjects under treatment conditions with subjects who drop out of the program. In the case of quasi-experimental studies, given that they do not have a control group to determine recidivism rate, 0.21 has been used as test value, since it has been consistently reported in the literature as base rate in terms of RO (Arias et al., 2013; Babcock et al., 2004; O'Leary et al., 1989; Rosenfeld, 1992).

In order to study moderating variables, those that have been outlined as relevant in previous meta-analyses have been selected. In addition, same criteria have been followed when coding, since in this way conclusions reached in this work can be easily compared with conclusions already obtained by previous studies.

First, the follow-up period was dichotomized according to whether programs had a follow-up period of fewer than 12 months ($k = 7$) or equal to or greater than 12 months ($k = 32$). On the other hand, the types of intervention were coded according to whether they used the Duluth model ($k = 14$), CBT ($k = 10$), or OTI ($k = 15$). Lastly, the duration of the intervention was coded according to whether the programs lasted fewer than 16 weeks/sessions ($k = 10$) or equal to or greater than 16 weeks/sessions ($k = 29$).

Finally, it is necessary to mention that previous meta-analyses used the source of information on recidivism as a moderating variable, which could be collected through ORs, PRs, or self-reports

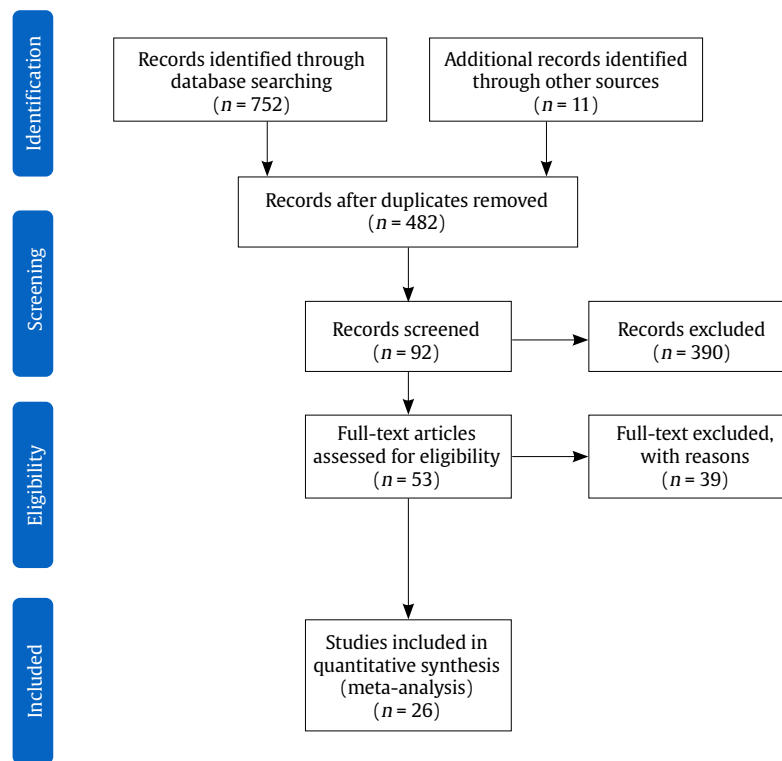


Figure 1. Flow Diagram of the Meta-analysis.

of the abuser. Although an abuser's self-reports have hardly been used due to their evident bias, PRs have been systematically coded; in this sense, given that the search for studies only reported a new study that used PRs, it was decided to reject it for the present study.

Data Analysis

The procedure used to carry out the present meta-analysis was the effect sizes of experiments and a set of workbooks designed by Suurmond et al. (2017), as well as a calculator of the probabilities of normal distribution and its inverses were used as tools for calculation purposes. To carry out the analysis, the rate of recidivism (dichotomous variable) was taken as an indicator of the effectiveness of the programs; this variable is generally expressed through proportions (percentages) and, therefore, the difference in proportions was used to calculate effect size. This difference implies a prior non-linear transformation of proportions, since the simple difference in proportions is not an adequate estimate of effect size (Arias et al., 2013). For this purpose, Hedges and Olkin's (1985) δ was used, a statistic derived from Kraemer and Andrews' (1982) procedure.

Kraemer and Andrews' (1982) procedure consists of a non-parametric approach for estimating the effect size through pre-posttest scores of experimental and control groups. However, if pretest scores are not available, there are derived statistics that allow effect size to be calculated from posttest scores, such as Hedges and Olkin's (1985) δ . Through this statistic, an estimate of pre-posttest effect size is obtained through the difference of the inverse function of the probability of the experimental group minus the probability of the control group, $\delta = \Phi^{-1}(\hat{p}_E) - \Phi^{-1}(\hat{p}_C)$. To complete the description of this formula, an example of effect size estimation is presented with the statistics obtained from the article by Haggard et al. (2015). Since the recidivism rate of the experimental group is 15%, the complementary outcome of this rate ($1 - .15 = .85$) was considered using the calculator of inverse function of the probability

for the estimation of $\Phi^{-1}(\hat{p}_E) = 1.0364$. For the estimation of $\Phi^{-1}(\hat{p}_C)$ in the case of quasi-experimental studies, since they do not have a control group to determine recidivism rate, 0.21 was used as the test value, since it has been consistently reported in the literature as the prime rate (Arias et al., 2013; Babcock et al., 2004; O'Leary et al., 1989; Rosenfeld, 1992). Given the complementary outcome of this test value is 0.79, the estimate of $\Phi^{-1}(\hat{p}_C) = .8064$. Therefore, $\delta = \Phi^{-1}(\hat{p}_E) - \Phi^{-1}(\hat{p}_C) = .23$. Then, an effect size of 0.20, 0.50, 0.80 is considered to be small, medium, and large, respectively.

Once all the effect sizes of experiments were computed, we calculated (a) average effect size weighted by sample size (δ_w); (b) weighted average observed variance ($S_{\delta_w}^2$); (c) observed standard deviation (SD_{δ_w}); (d) sampling error variance (S_e^2); (e) residual variance (S_{RES}^2); (f) percentage of observed variance explained by sampling error (% VE); and (g) confidence interval (90% CI_g). Assuming that said interval contains 0, it would imply the existence of heterogeneity (not significant effect) and, in this case, moderator variables should be studied. Likewise, heterogeneity was also evaluated using Q and I² statistics. Heterogeneity can be quantified as low, moderate, and high, with upper limits of 25%, 50%, and 75% for I², respectively. The results of the meta-analysis are presented below.

Results

Before carrying out the meta-analysis, a study of outliers was carried out, with the aim of avoiding contamination of results. For this purpose, $\pm 1.5 \times \text{IQR}$ (inter-quartile range) of the average effect size weighted by sample size (δ_w) was used as decision criterion. The results obtained did not yield any anomalous value, so the meta-analytical procedure was continued.

The present study consists of a total of 4 meta-analyses, the first of which corresponds to general meta-analysis (Table 2), and the subsequent ones to the analysis of the three previously coded

moderating variables: (a) follow-up period (Table 3); (b) type of intervention (Table 4); and (c) duration of intervention (Table 5).

Table 2. Results of the General Meta-analysis

| k | N | δ_w | $S_{\delta_w}^2$ | SD_{δ_w} | S_e^2 | S_{RES}^2 | % EV | 90% CI $_{\delta}$ |
|----|--------|------------|------------------|-----------------|---------|-------------|-------|--------------------|
| 39 | 19,773 | 0.54 | 0.049 | 0.22 | 0.008 | 0.041 | 16.32 | [-.18, .90] |

Note. k = number of studies; N = cumulative sample size; δ_w = average effect size weighted by sample size; $S_{\delta_w}^2$ = observed variance of δ_w ; SD_{δ_w} = observed standard deviation of δ_w ; S_e^2 = variance of the sampling error of δ_w ; S_{RES}^2 = residual variance; % EV = percentage of $S_{\delta_w}^2$ explained by sampling error; 90% CI $_{\delta}$ = confidence interval for δ_w .

Results for general meta-analysis illustrated in Table 2 contain the data of the 39 effect sizes that have been computed, with a cumulative sample of 19,773 subjects. When analyzing the data, a level of heterogeneity was observed to be elevated, $Q(38) = 238.59$, $p < .001$, which suggested the presence of moderating variables ($I^2 = .85$). This indicated the need to use an effects analysis model random. The use of this model showed a size of the average effect weighted by the sample size of .54, therefore positive, with an observed variance of 0.049, of which 16.32% is explained by sampling error. However, the effect found is not significant since the confidence interval contains the value 0 (90% confidence interval CI $_{\delta}$ [-.18, .90]; in this way, we proceed to search for moderator variables.

For the search for moderating variables, the sequence of analysis carried out by previous meta-analyses was respected, since it is covered by theoretical criteria (Rosenthal, 1995). In this sense, the scientific literature has indicated the duration of the follow-up period, as one of the most relevant moderators in recidivism (Gondolf, 2002; Redondo et al., 2001), which is why it has been the first moderating variable to be analyzed. Along these lines, results shown in Table 3 reflect that most of the studies take into account the importance of this criterion, including long follow-up periods ($k = 32$), while few studies use a period less than 12 months ($k = 7$). The average effect size weighted by the sample size is greater when the follow-up period exceeds one year ($\delta_w = 0.55$), but, nevertheless, the results are only significant for the follow-up period of fewer than 12 months (90% confidence interval CI $_{\delta}$ [.14, .90], showing an effect size close to

moderate ($\delta_w = 0.38$). The observed variance is greater in the group of fewer than 12 months ($S_{\delta_w}^2 = 0.104$) compared to that of more than 12 months ($S_{\delta_w}^2 = 0.045$), being explained by 25.96% and 15.55% by the sampling error, respectively.

The type of intervention implemented by the programs has also emerged as a relevant moderator variable. In this sense, in Table 4 it can be observed that the type of Duluth treatment ($k = 14$) has a positive, but not significant average effect ($\delta_w = .51$, 90% CI $_{\delta}$ [-.23, .79]), the CBT ($k = 10$) has a positive average effect, significant and moderate size ($\delta_w = .57$; 90% CI $_{\delta}$ [.04, 1.18]), and the interventions grouped under OTI ($k = 15$) show a positive average effect, but not significant ($\delta_w = .59$, 90% CI $_{\delta}$ [-.18, 1]). On the other hand, although OTIs are the type of program that gathers the most studies, they group a great heterogeneity of studies, so that the Duluth model and CBT remain the most widely implemented types of intervention.

Finally, Table 5 shows the results of the meta-analysis carried out with the moderating variable of the duration of the intervention. In this regard, it is interesting to mention that most of the interventions use long interventions according to the coding that has been carried out ($k = 29$), with a duration of 16 weeks or sessions onwards, while fewer interventions with a duration below this ($k = 10$). Regarding results, average effect size weighted by sample size is positive, but not significant, in the case of interventions that exceed 16 weeks/sessions ($\delta_w = .57$, 90% CI $_{\delta}$ [-.26, .88]), while it is positive, significant, and close to moderate in the case of interventions that do not exceed that duration ($\delta_w = .39$, 90% CI $_{\delta}$ [.09, .87]). The variance is 0.084 for shorter interventions, part of which is due to sampling error ($S_e^2 = 0.015$), while it is 0.039 for longer interventions, part of which is also due to sampling error ($S_e^2 = 0.007$).

Discussion

The results obtained in this meta-analysis show that, in general, intervention on gender abusers has a positive, but not significant effect. In this sense, conclusions follow the historical thread of those already reached by previous meta-analytic studies (Arias et al., 2013; Babcock et al., 2004; Feder & Wilson, 2005; Smedslund et al., 2011). However, it is necessary to specify that the results presented refer

Table 3. Results of the Meta-analysis of the Follow-up Period

| Period | k | N | δ_w | $S_{\delta_w}^2$ | SD_{δ_w} | S_e^2 | S_{RES}^2 | % EV | 90% CI $_{\delta}$ |
|-------------|----|-------|------------|------------------|-----------------|---------|-------------|-------|--------------------|
| < 12 months | 7 | 1079 | 0.38 | 0.104 | 0.32 | 0.027 | 0.077 | 25.96 | [.14, .90] |
| ≥ 12 months | 32 | 18694 | 0.55 | 0.045 | 0.21 | 0.007 | 0.038 | 15.55 | [-.21, .89] |

Note. k = number of studies; N = cumulative sample size; δ_w = average effect size weighted by sample size; $S_{\delta_w}^2$ = observed variance of δ_w ; SD_{δ_w} = observed standard deviation of δ_w ; S_e^2 = variance of the sampling error of δ_w ; S_{RES}^2 = residual variance; % EV = percentage of $S_{\delta_w}^2$ explained by sampling error; 90% CI $_{\delta}$ = confidence interval for δ_w .

Table 4. Results of the Meta-analysis of the Type of Intervention

| Type | k | N | δ_w | $S_{\delta_w}^2$ | SD_{δ_w} | S_e^2 | S_{RES}^2 | % EV | 90% CI $_{\delta}$ |
|--------|----|-------|------------|------------------|-----------------|---------|-------------|-------|--------------------|
| Duluth | 14 | 12705 | 0.51 | 0.029 | 0.17 | 0.005 | 0.024 | 17.24 | [-.23, .79] |
| CBT | 10 | 1713 | 0.57 | 0.139 | 0.37 | 0.025 | 0.114 | 17.99 | [.04, 1.18] |
| OTI | 15 | 5355 | 0.59 | 0.064 | 0.25 | 0.012 | 0.052 | 18.75 | [-.18, 1] |

Note. k = number of studies; N = cumulative sample size; δ_w = average effect size weighted by sample size; $S_{\delta_w}^2$ = observed variance of δ_w ; SD_{δ_w} = observed standard deviation of δ_w ; S_e^2 = variance of the sampling error of δ_w ; S_{RES}^2 = residual variance; % EV = percentage of $S_{\delta_w}^2$ explained by sampling error; 90% CI $_{\delta}$ = confidence interval for δ_w .

Table 5. Results of the Meta-analysis of the Duration of the Intervention

| Duration | k | N | δ_w | $S_{\delta_w}^2$ | SD_{δ_w} | S_e^2 | S_{RES}^2 | % EV | 90% CI $_{\delta}$ |
|----------------------|----|-------|------------|------------------|-----------------|---------|-------------|-------|--------------------|
| < 16 weeks/ sessions | 10 | 2728 | 0.39 | 0.084 | 0.29 | 0.015 | 0.069 | 17.86 | [.09, .87] |
| ≥ 16 weeks/ sessions | 29 | 17045 | 0.57 | 0.039 | 0.19 | 0.007 | 0.032 | 17.95 | [-.26, .88] |

Note. k = number of studies; N = cumulative sample size; δ_w = average effect size weighted by sample size; $S_{\delta_w}^2$ = observed variance of δ_w ; SD_{δ_w} = observed standard deviation of δ_w ; S_e^2 = variance of the sampling error of δ_w ; S_{RES}^2 = residual variance; % EV = percentage of $S_{\delta_w}^2$ explained by sampling error; 90% CI $_{\delta}$ = confidence interval for δ_w .

only to ORs, since in the search for studies only one additional study could be identified that reported recidivism through PRs.

Regarding the moderating variable of the follow-up period, the results were significant only for studies that used a period shorter than 12 months, showing an effect size in the reduction of recidivism close to moderate. In this sense, there is a certain correspondence with the scientific literature since, given that in gender violence recidivism tends to occur in the first six months (Gondolf, 2002; Redondo et al., 2001) and intervention with this group affects its reduction in this period; however, non-significant results for the period of more than 12 months do not allow us to extend this conclusion into the follow-up period. These results partially correspond to results found by Babcock et al. (2004), since in their study they found positive, significant, and small effect sizes for both follow-up periods.

For its part, the moderator of type of intervention turned out to be only significant for CBT, being positive and moderate size, unlike in previous meta-analyses such as Arias et al.'s (2013) or Babcock et al.'s (2004), who found significant results only for OTIs. A possible explanation for the effectiveness of CBT on reducing recidivism could be that the intervention focuses on behavioral aspects and the measure used (recidivism rate) records precisely these aspects; in this sense, the Duluth model has a greater emphasis on attitudinal and value-related aspects (Pence & Paymar, 1983).

Lastly, the duration of the intervention variable was not significant for long interventions (≥ 16 weeks/sessions), unlike previous meta-analyses, but it was significant for short interventions (< 16 weeks/sessions), being positive and close to moderate in size. These a priori results do not obey a logic and go in the opposite direction to findings of scientific literature, which establish that longer interventions are more effective because they affect toxic cognitions, which are characterized by their internality, stability, and globality (Maruna, 2004). However, as Boal and Mankowski (2014) state, the relationship between the duration of the intervention programs and the effectiveness in terms of recidivism is still not entirely clear.

In sum, the results of the present meta-analysis do not allow us to establish definitive conclusions about the effectiveness of the intervention with gender abusers. Following the conclusions reached by previous meta-analytical studies, results are not sufficient but, at the same time, the ineffectiveness of treatments cannot be deduced from them. In any case, the evidence collected in this paper should stimulate scientific interest in the analysis of other moderating variables that may influence recurrence of gender-based violence, as well as other characteristics of possible implementation in intervention programs that may become an improvement in its efficiency. In this regard, there is a rich and diversified line of research that ranges from aspects such as the consideration of particular needs of individuals (Cavanaugh & Gelles, 2005; Huss & Ralston, 2008) to aspects such as the problem of the high drop-out rate present among programs (Crane et al., 2015; Ferrer-Pérez & Bosch-Fiol, 2016).

Study Limitations

The results obtained in the present meta-analysis should be interpreted in light of the present limitations. On the one hand, the relative reliability of the measure used as an indicator of recidivism (OR) can be highlighted, since a significant rate of it may not be contained in these registries, which is known as hidden victimization or black crime. Furthermore, it has not been possible to consider PR, which, as has been highlighted in the scientific literature, show very different recidivism rates compared to OR.

On the other hand, it is worth noting both the heterogeneity of coded studies and the way in which they present the information, given that on many occasions key results were reported unclearly. Likewise, the scarcity of experimental designs as opposed to the number of quasi-experimental designs may affect the results

obtained, since for the latter a base rate consistently reported in the scientific literature has been used as the test value, but has not been properly contracted with an equivalent control group.

Finally, other moderating variables that could play an important role in explaining recidivism in gender violence have not been considered, such as a high drop-out rate, lack of motivation and adherence to treatment, or non-consideration of characteristics or personal needs of subjects. In this sense, it would be interesting for the new studies to focus on deepening these aspects, with the aim of ending the debate on the effectiveness of this type of programs and thus being able to face one of the greatest social problems of our time, gender violence.

Conclusions

The results of this meta-analytical review have significant implications in the criminological field. There is a literature corpus on the effectiveness of interventions, which shows significant effects on reduction of recidivism in official records compared to other measures for reduction of recidivism, such as protection orders (Herrera & Amor, 2017). However, in order to improve the effectiveness of these programs, other individual variables based on the risk-need-responsivity (RNR) model of Andrews et al. (2011) could also be considered for their evaluation. For example, taking into account the typology of batterers and the possible mental disorders they present can play an important role in intervention programs (Aguilar-Ruiz, 2018). It is also recommended that other relevant factors in gender violence be evaluated as a criterion for the analysis of the effectiveness of these programs, such as the cognitive competence involved in their re-education or the acquisition of skills and competences that reduce toxic cognitions and dysfunctional thoughts characteristic of sexual offenders (Arnoso et al., 2017).

Conflict of Interest

The authors of this article declare no conflict of interest.

References

- References marked with an asterisk indicate studies included in the metaanalysis.*
- Aguilar-Ruiz, R. (2018). Typologies of femicide perpetrators with a mental disorder in Spain. *Anuario de Psicología Jurídica*, 28(1), 39-48. <https://doi.org/10.5093/apj2018a4>
- Andrews, D. A., Bonta, J., & Wormith, J. S. (2011). The risk-need-responsivity (RNR) model: Does adding the good lives model contribute to effective crime prevention?. *Criminal Justice and Behavior*, 38(7), 735-755. <https://doi.org/10.1177/2F0093854811406356>
- Arias, E., Arce, R., & Vilariño, M. (2013). Batterer intervention programmes: A meta-analytic review of effectiveness. *Psychological Intervention*, 22(2), 153-160. <https://doi.org/10.5093/in2013a18>
- Arias, E., Velasco, J., & Novo, M. (2016). Análisis bibliométrico sobre la investigación en violencia de género. Fundamentos y nuevas tendencias. In A. A. Pueyo, F. Fariña, M. Novo, & D. Seijo (Eds.), *Colección de Psicología y Ley nº13. Avances en psicología jurídica y forense* (pp. 81-91). Sociedad Española de Psicología Jurídica y Forense. Madrid, Spain. <https://doi.org/10.13140/RG.2.1.2594.2807>
- Arnoso, A., Ibabe, I., Arnoso, M., & Elgorriaga, E. (2017). Sexism as predictor of intimate partner violence in a multicultural context. *Anuario de Psicología Jurídica*, 27(1), 9-20. <https://doi.org/10.1016/j.apj.2017.02.001>
- 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. <https://doi.org/10.1016/j.cpr.2002.07.001>
- *Babcock, J. C., & Steiner, R. (1999). 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(1), 46-59. <https://doi.org/10.1037/0893-3200.13.1.46>
- Boal, A. L., & Mankowski, E. S. (2014). The impact of legislative standards on batterer intervention program practices and characteristics. *American Journal of Community Psychology*, 53, 218-230. <https://doi.org/10.1007/s10464-014-9637-3>

- Boira, S., & Marcuellos, C. (2014, May). *Sociedad y violencia contra las mujeres: El caso español, una encrucijada técnica e ideológica*. Las paces de cada día. Zaragoza. <http://www.pensamientopenal.com.ar/system/files/2014/12/doctrina36363.pdf#page=180>
- *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(2), 189-209. <https://doi.org/10.1348/135532505X36778>
- Cáceres, J. (2011). Abuso y violencia en las relaciones de pareja. *Psicología Conductual*, 19(1), 91-116. https://www.researchgate.net/publication/236299341_Abuso_y_Violencia_en_las_relaciones_de_Pareja
- *Carbajosa, P., Catalá-Miñana, A., Lila, M., & Gracia, E. (2017). Differences in treatment adherence, program completion, and recidivism among batterer subtypes. *The European Journal of Psychology Applied to Legal Context*, 9(2), 93-101. <https://doi.org/10.1016/j.ejpal.2017.04.001>
- Cavanaugh, M. M., & Gelles, R. J. (2005). The utility of male domestic violence offender typologies. *Journal of Interpersonal Violence*, 20(2), 155-166. <https://doi.org/10.1177/0886260504268763>
- *Chen, H., Bersani, C., Myers, S. C., & Denton, R. (1989). Evaluating the effectiveness of a court sponsored abuser treatment program. *Journal of Family Violence*, 4(4), 309-322. <https://doi.org/10.1007/BF00978573>
- *Coulter, M., & VandeWeerd, C. (2009). Reducing domestic violence and other criminal recidivism: Effectiveness of a multilevel batterers intervention program. *Violence and Victims*, 24(2), 139-152. <https://doi.org/10.1891/0886-6708.24.2.139>
- Crane, C. A., Eckhardt, C. I., & Schlauch, R. C. (2015). Motivational enhancement mitigates the effects of problematic alcohol use on treatment compliance among partner violent offenders: Results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology*, 83(4), 689-695. <https://doi.org/10.1037/a0039345>
- Cuenca, M. L., Graña, J. L., & Redondo, N. (2015). Differences in prevalence of partner aggression as the Revised Conflict Tactics Scale: Individual and dyadic report. *Behavioral Psychology*, 23(1), 127-140. https://www.researchgate.net/publication/283473431_Differences_in_prevalence_of_partner_aggression_as_the_revised_conflict_tactics_scale_Individual_and_dyadic_report
- *Davis, R. C., Taylor, B. G., & Maxwell, C. D. (1998). Does batterer treatment reduce violence? A randomized experiment in Brooklyn. *Justice Quarterly*, 18(1), 171-201. <https://www.ncjrs.gov/pdffiles1/nij/grants/180772.pdf>
- *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(3), 468-476. <https://doi.org/10.1037/0022-006X.68.3.468>
- *Dutton, D. G., Bodnarchuk, M., Kropp, R., Hart, S. D., & Ogloff, J. R. P. (1997). Wife assault treatment and criminal recidivism: An 11-year follow-up. *International Journal of Offender Therapy and Comparative Criminology*, 41(1), 9-23. <https://doi.org/10.1177/0306624X9704100102>
- *Farzan-Kashani, J., & Murphy, C. M. (2015). Anger problems predict long-term criminal recidivism in partner violent men. *Journal of Interpersonal Violence*, 32(23), 3541-3555. <https://doi.org/10.1177/0886260515600164>
- 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. <https://doi.org/10.1007/s11292-005-1179-0>
- Fernández-Ríos, L., & Buela-Casal, G. (2009). Standards for the preparation and writing of psychology review articles. *International Journal of Clinical and Health Psychology*, 9(2), 329-344. <http://www.aepc.es/ijchp/ref-es-326.pdf>
- Ferrer-Pérez, V. A., & Bosch-Fiol, E. (2016). Batterer intervention programs in Spain: An analysis of their effectiveness. *International Journal of Offender Therapy and Comparative Criminology*, 62(4), 885-897. <https://doi.org/10.1177/0306624X16672455>
- Ferrer-Pérez, V. A., Ferreiro-Basurto, V., Navarro-Guzmán, C., & Bosch-Fiol, E. (2016). Programas de intervención con maltratadores en España: La perspectiva de los/as profesionales. *Psychosocial Intervention*, 25(3), 159-168. <https://doi.org/10.1016/j.psi.2016.06.001>
- Geldschläger, H., Ginés, D., & Ponce, A. (2009). *Informe de evaluación sobre programas de rehabilitación para hombres que maltratan desarrollados en el ámbito europeo*. Fundación IRES.
- Ginés, O., Geldschläger, H., Nax, D., & Ponce, A. (2015). European perpetrator programmes: A survey on day-to-day outcome measurement. *Studia Humanistyczne AGH*, 14(2), 33-52. <https://doi.org/10.7494/human.2015.14.2.33-52>
- Gondolf, E. W. (2002). *Batterer intervention systems: Issues, outcomes, and recommendations*. Sage
- Gracia, E., & Lila, M. (2015). *Attitudes towards violence against women in the EU*. Publication Office of the European Union.
- *Graña, J. L., Redondo, N., Muñoz-Rivas, M. J., & Cuenca, M. L. (2017). Short-term treatment response, attrition and recidivism in a partner violent men typology compared with a control group. *Behavioral Psychology*, 25(3), 465-482. http://behavioralpsycho.com/wp-content/uploads/Articulos2017/Diciembre2017/Completo02.Grana_25-3.pdf
- *Haggard, U., Freij, I., Danielsson, M., Wenander, D., & Langström, N. (2015). Effectiveness of the IDAP treatment program for male perpetrators of intimate partner violence: A controlled study of criminal recidivism. *Journal of Interpersonal Violence*, 32(7), 1027-1043. <https://doi.org/10.1177/0886260515586377>
- *Hamberger, L. K., & Hastings, J. E. (1988). Skills training for treatment of spouse abusers: An outcome study. *Journal of Family Violence*, 3(2), 121-130. <https://doi.org/10.1007/BF00994029>
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press. <https://doi.org/10.2307/1164953>
- Heinsman, D. T., & Shadish, W. R. (1996). Assignment methods in experimentation: When do nonrandomized experiments approximate answers from randomized experiments? *Psychological Methods*, 1(2), 154-169. <https://doi.org/10.1037/1082-989X.1.2.154>
- *Herman, K., Rotunda, R., Williamson, G., & Vodanovich, S. (2014). Outcomes from a duluth model batterer intervention program at completion and long term follow-up. *Journal of Offender Rehabilitation*, 53(1), 1-18. <https://doi.org/10.1080/10509674.2013.861316>
- Herrera, M. J., & Amor, P. J. (2017). Violation of protection orders by men previously convicted of intimate partner violence. *Anuario de Psicología Jurídica*, 27(1), 1-8. <https://doi.org/10.1016/j.apj.2016.05.003>
- 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(6), 1000-1019. <https://doi.org/10.1037/0022-006X.68.6.1000>
- Holtzworth-Munroe, A., & Stuart, G. L. (1994). Typologies of male batterers: Three subtypes and the differences among them. *Psychological Bulletin*, 116(3), 476-497. <https://doi.org/10.1037/0033-2909.116.3.476>
- Huss, M. T., & Ralston, A. (2008). Do batterer subtypes actually matter? Treatment completion, treatment response, and recidivism across a batterer typology. *Criminal Justice and Behavior*, 35(6), 710-724. <https://doi.org/10.1177/0093854808316218>
- Kraemer, H. C., & Andrews, G. (1982). A non-parametric technique for meta-analysis effect size calculation. *Psychological Bulletin*, 91(2), 404-412. <https://doi.org/10.1037/0033-2909.91.2.404>
- *Lauch, K. M., Hart, K. J., & Bresler, S. (2017). Predictors of treatment completion and recidivism among intimate partner violence offenders. *Journal of Aggression*, 26(5), 543-557. <https://doi.org/10.1080/10926771.2017.1299824>
- *Lila, M., Gracia, E., & Catalá-Miñana, A. (2018). Individualized motivational plans in batterer intervention programs: A randomized clinical trial. *Journal of Consulting and Clinical Psychology*, 86(4), 309-320. <https://doi.org/10.1037/ccp0000291>
- *Lila, M., Gracia, E., & Catalá-Miñana, A. (2020). More likely to dropout, but what if they don't? Partner violence offenders with alcohol abuse problems completing batterer intervention programs. *Journal of Interpersonal Violence*, 35(9-10), 1958-24. <https://doi.org/10.1177/0886260517699952>
- Mackay, E., Gibson, A., Lam, H., & Beecham, D. (2015). *Perpetrator interventions in Australia: Part one - Literature review: State of knowledge paper*. ANROWS
- Maruna, S. (2004). Desistance and explanatory style: A new direction in the psychology of reform. *Journal of Contemporary Criminal Justice*, 20(2), 184-200. <https://doi.org/10.1177/1043986204263778>
- *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(3), 263-284. <https://doi.org/10.1023/A:1022841022524>
- 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(2), 219-234. http://www.aepc.es/ijchp/articulos_pdf/ijchp-408.pdf
- O'Leary, K. D., Barling, K., 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(2), 263-269. <https://doi.org/10.1037/0022-006X.57.2.263>
- Organización Mundial de la Salud (OMS, 2013). *Estimaciones mundiales y regionales de la violencia contra la mujer: prevalencia y efectos de la violencia conyugal y de la violencia sexual no conyugal en la salud*. http://apps.who.int/iris/bitstream/10665/85243/1/WHO_RHR_HRP_13.06_spa.pdf?ua=1
- Organización Mundial de la Salud (OMS, 2016). *Violence against women. Intimate partner and sexual violence against women*. <http://www.who.int/mediacentre/factsheets/fs239/en>
- *Palmer, S. E., Brown, R. A., & Barrera, M. E. (1992). Group treatment program for abusive husbands: Long-term evaluation. *American Journal of Orthopsychiatry*, 62(2), 276-283. <https://doi.org/10.1037/h0079336>
- Pence, E., & Paymar, M. (1983). *Domestic violence information manual. The Duluth domestic abuse intervention project*. Springer. <http://www.duluth-model.org>
- Perestelo-Pérez, L. (2013). Standards on how to develop and report systematic reviews in psychology and health. *International Journal of Clinical and Health Psychology*, 13(1), 49-57. [https://doi.org/10.1016/S1697-2600\(13\)70007-3](https://doi.org/10.1016/S1697-2600(13)70007-3)
- *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]. Instituto de Ciencias Forenses y de la Seguridad (ICFS) y Secretaría General de Instituciones Penitenciarias. Madrid. <https://doi.org/10.13140/RG.2.1.2626.9282>
- 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(1), 47-62. <http://www.psychologyinspain.com/content/reprints/2001/6.pdf>
- Rosenfeld, B. D. (1992). Court ordered treatment of spouse abuse. *Clinical Psychology Review*, 12(2), 205-226. [https://doi.org/10.1016/0272-7358\(92\)90115-0](https://doi.org/10.1016/0272-7358(92)90115-0)
- Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, 118(2), 183-192. <https://pdfs.semanticscholar.org/3d31/3cf02c2c-dd8e19032f56268ac153b55fa58f.pdf>
- Sánchez-Meca, J. (2010). Cómo realizar una revisión sistemática y un meta-análisis. *Aula Abierta*, 38(2), 53-64. <http://www.um.es/metaanalysis/pdf/5030.pdf>
- *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(4), 393-414. <https://doi.org/10.1891/0886-6708.11.4.393>
- *Scott, K., Heslop, L., Kelly, T., & Wiggins, K. (2015). Intervening to prevent repeat offending among moderate- to high-risk domestic violence offenders: A second-responder program for men. *International Journal of Offender Therapy and Comparative Criminology*, 59(3), 273-294. <https://doi.org/10.1177/0306624X13513709>
- 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, CD006048. <https://doi.org/10.1002/14651858.CD006048.pub2>
- Sonkin, D. J., Martin, D., & Walker, L. (1985). *Male batterer: A treatment approach*. Springer.
- Subirana, M., & Pueyo, A. A. (2013). Retención proactiva y adherencia terapéutica en programas formativos para hombres maltratadores de la pareja. *Psychosocial Intervention*, 22(2), 95-104. <https://doi.org/10.5093/in2013a12>
- Suurmond, R., van Rhee, H., & Hak, T. (2017). Introduction, comparison and validation of meta-essentials: A free and simple tool for meta-analysis. *Research Synthesis Methods*, 8(4), 537-553. <https://doi.org/10.1002/jrsm.1260>
- *Taylor, B. G., & Maxwell, C. D. (2009). *The effect of a short-term batterer treatment program for detained arrestees: A randomized experiment in the Sacramento County, California jail*. Department of Justice, Crime and Violence Prevention Center. Sacramento, CA. <https://doi.org/10.1037/e640772009-001>
- *Tollefson, D. R., & Gross, E. R. (2006). Predicting recidivism following participation in a treatment program for batterers. *Journal of Social Service Research*, 32(4), 39-62. https://doi.org/10.1300/J079v32n04_03
- *Tollefson, D. R., & Phillips, I. (2015). A mind-body bridging treatment program for domestic violence offenders: Program overview and evaluation results. *Journal of Family Violence*, 30(6), 783-794. <https://doi.org/10.1007/s10896-015-9715-9>
- *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(1), 17-45. <https://doi.org/10.1080/10926770802610657>
- Urrútia, G., & Bonfill, X. (2010). Declaración PRISMA: Una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis. *Medicina Clínica*, 135(11), 507-511. <https://doi.org/10.1016/j.medcli.2010.01.015>
- *Zarling, A., Bannon, S., & Berta, M. (2017). *Evaluation of acceptance and commitment therapy for domestic violence offenders*. *Psychology of violence*. Advance publication online. <https://doi.org/10.1037/vio0000097>

