



Methodological Challenges in Group-based Randomised Controlled Trials for Intimate Partner Violence Perpetrators: A Meta-summary

William Turner^a, Karen Morgan^b, Marianne Hester^a, Gene Feder^b, and Helen Cramer^b

^aSchool for Policy Studies, University of Bristol, UK; ^bBristol Medical School, Centre for Academic Primary Care, University of Bristol, UK

ARTICLE INFO

Received 30 November 2022
Accepted 10 April 2023

Keywords:

Group-based perpetrator programmes
Domestic violence
Partner abuse
Batterer intervention
Gender-based violence

ABSTRACT

Evidence for treatment effects of group-based Intimate Partner Violence (IPV) perpetrators programmes remains, at best, inconclusive. In the present review, systematic/meta-analytic reviews were used to identify randomised controlled trials and a meta-summary approach was employed to identify methodological challenges in the design and conduct of these trials. Of the fifteen studies identified, seven were comparative effectiveness trials. A range of methodological challenges were also identified by the trialists; source of outcome data, treatment modality, attrition and sample characteristics were the most frequently mentioned. Although there are only a few randomised controlled trials compared to non randomised studies, the findings of both highlight the need to invest in the development of innovative and/or combined IPV treatment programmes to address co-occurring issues such as substance use and trauma. The summary of methodological challenges will provide the first step in the development of methods guidance for researchers working in this area.

Desafíos metodológicos en los ensayos controlados aleatorizados con grupos de intervención para agresores de pareja: un meta-resumen

RESUMEN

La evidencia del efecto del tratamiento de los programas de intervención grupal para agresores de pareja (IPV) siguen siendo, en el mejor de los casos, no concluyentes. En la presente revisión se emplearon revisiones sistemáticas/meta-analíticas para identificar ensayos controlados aleatorizados y se empleó un enfoque de meta-resumen para identificar los desafíos metodológicos en el diseño y la realización de estos ensayos. De los quince estudios identificados, siete fueron ensayos de comparación de la efectividad. Los autores de los ensayos también identificaron una serie de desafíos metodológicos: la fuente de la que se obtienen los datos relativos a los resultados, la modalidad de tratamiento, la mortalidad de la muestra y sus características fueron los mencionados con más frecuencia. Aunque son todavía escasos los ensayos controlados aleatorizados en comparación con los estudios no aleatorizados, los resultados de ambos tipos de estudios han destacado la importancia de invertir en el desarrollo de programas de tratamiento de IPV innovadores y/o combinados para tratar problemáticas conjuntas tales como consumo de sustancias y trauma. El resumen de los desafíos metodológicos proporcionará el primer paso para el desarrollo de guías metodológicas para los investigadores que trabajan en este área.

Palabras clave:

Programas grupales para perpetradores
Violencia doméstica
Abuso de pareja
Intervención para agresores
Violencia de género

Thirty years after the United Nations Assembly declared the need to eliminate violence against women, the prevalence of violence against this population remains high, with more than a quarter of women aged 15 to 49 experiencing physical and/or sexual violence performed by their intimate partner (World Health Organization, 2021). Intimate partner violence (IPV) has a severe impact on the mental health of victim/survivors as well as an economic and social impact, which can exert an overload on health systems (Peterson et al., 2018; Sardinha et al., 2022; World Health Organization, 2013). In addition to its impact on the development of mental disorders such as post-traumatic stress disorder (PTSD) and depression,

IPV is associated with chronic pain, headache, gynecological or gastrointestinal problems in victim/survivors (Baker et al., 2021; Campbell et al., 2018; Daugherty et al., 2021; Gibson et al., 2019).

The implementation of IPV interventions is fundamental to reducing the victim/survivor's risk of suffering future violence. IPV perpetrator programmes were first developed over 40 years ago with the realisation that providing services to survivors of domestic violence and abuse without addressing the behaviour of perpetrators, was an inadequate response to this damaging violation of human rights (Hamel & Nicholls, 2006; Healey et al., 1999). Group treatment was favoured instead of individual therapy/counselling as it was

Cite this article as: Turner, W., Morgan, K., Hester, M., Feder, G., & Cramer, H. (2023). Methodological challenges in group-based randomised controlled trials for intimate partner violence perpetrators: A meta-summary. *Psychosocial Intervention*, 32(2), 123-139. <https://doi.org/10.5093/pi2023a9>

Correspondence: Helen.Cramer@bristol.ac.uk (H. Cramer).

ISSN: 1132-0559/© 2023 Colegio Oficial de la Psicología de Madrid. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

believed that it offered abusers: more opportunities to expand their social networks to include others who are supportive of being non-abusive (Barner & Carney, 2011; Krishna et al., 2011; National Research Council et al., 1996); provided opportunities for peer learning (Yalom & Leszcz, 2020); and be potentially more cost effective by being able to cover a wide range of different treatment approaches (Karakurt et al., 2019). Group programmes do not work for all perpetrators and the individual approach is necessary with high harm individuals, who are difficult to engage and unlikely to attend groups (Tsantefski et al., 2021). Eckhardt et al. (2013) note that the social and historical analysis of IPV has been instrumental in formulating the philosophy, structure, and goals of perpetrator intervention. However, while such programmes in the United Kingdom at least share the goals of reducing levels of violence and promote victim/survivor safety, there is considerable variability in intervention methods and approaches (Murphy et al., 2020; Phillips et al., 2013). Many programmes operate an open group format, with one or two facilitators managing group discussion and activities. Participation may either be on a 'voluntary' basis or mandatory; the latter may be required by criminal courts as part of a pre-trial diversion programme, ordered by judges as part of a sentence, or may be imposed by probation agencies (Hamberger & Hastings, 1993). Even when IPV programmes are referred to as 'voluntary', there are likely to be pressures to participate from partners and ex-partners or social services.

Historically, most IPV male perpetrator programmes were grounded in a feminist analysis of abuse, informing psychoeducational models or the Duluth model (Barner & Carney, 2011; Phillips et al., 2013). Programmes operating on the Duluth model (Pence et al., 1993) are based on an understanding of the patriarchal nature of society and institutions that support and perpetuate male power, control and privilege within the domestic sphere (Dobash & Dobash, 1979), experienced as a sense of entitlement by male perpetrators of IPV. Such programmes assume that to reduce violence it is necessary to expose patriarchal or misogynistic behaviour and encourage perpetrators to take responsibility for adopting coercive and abusive tactics in relationships. In this way, more egalitarian attitudes and behaviours can be promoted (Eckhardt et al., 2013). Cognitive behavioural therapy (CBT) is another explicit model for perpetrator programmes. CBT approaches aim to change a person's behaviour by developing a therapeutic relationship, exposure and dissolution of distorted cognitions, behaviour change strategies, working on core beliefs and trying to prevent relapse and further abuse (Nesset et al., 2019). Currently, the work with IPV perpetrators is quite complex, and most group programmes use a mixture of interventions that include feminist approaches and psychotherapeutic approaches such as CBT (Lilley-Walker et al., 2018). Most programmes focus on IPV involving partners or ex-partners rather than the broader focus on domestic violence and abuse involving wider family violence (Eckhardt et al., 2013; Westmarland & Kelly, 2013).

There has been a proliferation of empirical studies evaluating IPV group perpetrator programmes over the last four decades. Reviews have attempted to summarise the evidence for these programmes, which remains uncertain (Cluss & Bodea, 2011; Eckhardt et al., 2013; Nesset et al., 2019). Generally, these reviews report uncertainty about programme effectiveness due to heterogeneity of populations and interventions, limited or poorly measured outcomes, short follow-up or more biased research designs with longer follow-up, and insufficient attention to the context of the interventions (Akoensi et al. 2013; Arias et al., 2013; Gondolf, 2011).

Conflicting results reported by evaluations of IPV perpetrator programmes fuel fierce debate about appropriate intervention models and outcomes. Mixed findings from studies leave practitioners and service providers with conflicting guidance on how best to deliver programmes (Akoensi et al., 2013). Additionally, such findings are challenging to funders of perpetrator programmes in criminal justice, local authority or health care sectors. For example, the UK National Institute of Care Excellence guidelines do not recommend

the commissioning of perpetrator programmes outside the context of further evaluation (National Research Council [NICE, 2013]).

In the light of uncertainty about IPV group perpetrator programme effectiveness, at least in part driven by problems in research design, treatment modalities and outcome measurement, in this review we aim to articulate the methodological challenges in the conduct of randomised controlled trials (RCTs) of IPV perpetrator programmes, to inform more robust evaluation. Previous meta-analysis has been performed to investigate the effectiveness of intervention programs for perpetrators (Karakurt et al., 2019). However, the authors included individual and group interventions without considering structural differences between such interventions. Furthermore, they only considered pre- and post-intervention results and did not differentiate between varied experimental conditions. Other recent systematic reviews have focused on different aspects of IPV interventions: a) effectiveness in different cultural contexts (Satyen et al., 2022); b) effectiveness for male perpetrators or male victims in health settings (Tarzia et al., 2020); c) comparison of experimental group versus group without any intervention (Cheng et al., 2021); d) risk-need-responsivity framework (Travers et al., 2021); e) motivational interviewing (Pinto E Silva et al., 2022; Santirso et al., 2020); and f) substance abuse (Stephens-Lewis et al., 2021).

As a research team developing and trialing our own group IPV programme for perpetrators considered to be in the mild to moderate risk category, we were particularly keen to gather and critically reflect on the key methodological challenges. While other reviews have tended to focus on the mixed or relatively weak evidence from RCTs or other designs (group format and individual programmes) with some consideration as to how methodological challenges might have made drawing conclusions so difficult, our review concentrates solely on the methodological challenges. Assessing these within RCTs is an opportunity to address these limitations and thereby improve the testing and the effectiveness of interventions. Focusing specifically on IPV programmes, our objectives were to identify and describe:

- 1) randomised controlled trials measuring the effects of IPV group male perpetrator programmes.
- 2) methodological challenges (contextual or procedural), as identified by the trialists in the included studies, that may explain the heterogeneity of the findings and why interpretation is difficult.
- 3) articulate common challenges involved in conducting evaluations of IPV group perpetrator programmes.

Method

Eligibility Criteria

To be included, studies had to meet the following criteria: (1) individual or cluster-randomised controlled trials (RCTs); (2) the study population comprised male perpetrators of IPV, where participants attended the treatment programme voluntarily or were court-ordered to participate (primary studies where the focus was on women who abuse their partner or spouse were excluded and trials having a mixed population of men and women perpetrators); (3) group format IPV programmes of any duration delivered in any setting employing any psychotherapeutic approach (for example based on Duluth, CBT or strengths-based or a combination of these); (4) any outcomes.

Search Methods for Identification of RCTs

RCTs were identified through published systematic reviews. Literature searches were conducted to identify systematic and meta-analytic reviews on the effectiveness of group interventions with IPV perpetrators.

Eight online databases, including Web of Science, Google Scholar, Scopus, PsycINFO, Cochrane Library, ERIC, PubMed, and DARE were

searched using Boolean operators and the following search and Mesh terms in the title and the abstract: “systematic review” OR “meta-analysis” AND “domestic violence” OR “batterer” OR “abuse” OR “intimate partner violence” OR “partner-violent men” OR “intervention program” OR “group” OR “offend*” AND “intervention” OR “treatment” OR “evaluation” OR “assessment” OR “effect*”.

These searches initially covered the period from January 2000 to March 2020; top-up searches were run again in February 2022 and March 2023. Reference lists of selected reviews were searched. The search was limited to systematic reviews and meta-analyses.

Selection of Studies

After mapping the studies (number and type of studies, i.e., experimental or quasi-experimental) included in the systematic and meta-analytic reviews, titles and abstracts of potentially relevant trials were screened by the first author and HC. We then retrieved the full text of all potentially relevant studies. Full-text articles were assessed for inclusion by WT, HC and GF independently against the eligibility criteria, with discrepancies resolved through discussion. Please see [Figure 1](#) for a flow diagram showing the selection process for RCTs from the systematic reviews and meta-analytic evaluations.

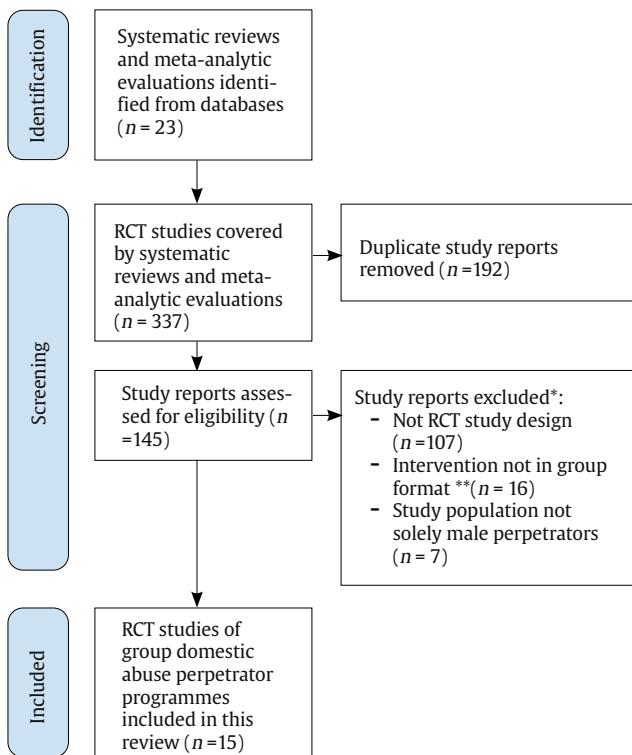


Figure 1. Flow Diagram Showing Identification of Randomised Controlled Trial Studies of Group-based Intimate Partner Violence Perpetrator Programmes from Systematic Reviews and Meta-analytic Evaluations.

Note. *Only 1 reason for exclusion was counted for each study. If there were more reasons than the fact that the study design was not a RCT then ‘not RCT study design’ was given as the primary reason. If the study did have a RCT design, then ‘intervention not in a group format’ was given as the main reason in preference to the study population being mixed male and females, even where both reasons were true.

**Intervention not in a group format is a shorthand exclusion label meaning the IPV perpetrator programme intervention was excluded if delivered in a one-to-one format. Please, also refer to the full inclusion and exclusion criteria which states: ‘To be included, studies had to meet the following criteria: (3) group format IPV programmes of any duration delivered in any setting employing any psychotherapeutic approach (for example based on Duluth, CBT or strengths-based or a combination of these). An obvious exception of an intervention which did not take a psychotherapeutic approach nor was delivered in a group format is [Lee et al. \(2008\)](#), where the medication fluoxetine was tested with male IPV perpetrators.’

Data Extraction and Management

The following data were extracted:

A) Study characteristics. Year of publication, country and setting where the study was conducted.

B) Participants. Number of participants assigned to intervention and control condition(s), referral status (e.g., voluntary and/or court-mandated), previous history of violent behaviour and treatment, mental health status, current substance abuse, additional problems/disorders.

C) Intervention. Content, components, duration/time, profession of person delivering the programme (or intervention), gender and number of therapist(s) or group leader(s), support to (ex)partner, risk management, the degree of mandatory delivery, attrition, adherence, type of comparison group (no intervention, other intervention).

D) Type of outcome measure(s). Recidivism or abuse outcomes (e.g., physical violence, aggression), attitudinal outcomes, interpersonal outcomes (e.g., managing anger) and psychological outcomes (e.g., self-esteem).

E) Source of outcome data. Routinely collected clinical or administrative data; self-reports, partner report, or other forms for gathering outcome data.

F) Length of follow-up time. Months and/or years.

G) Author narratives. For all included studies we also extracted: a) text providing the authors’ own assessment of the findings and b) text from the discussion and conclusion sections of each trial report (and, where available, its linked publications) that referred to *methodological challenges* in conducting IPV male group perpetrator programme evaluations. Methodological challenges were defined as: Any text within the discussion/conclusion section of the paper that *provides authors’ observations/reflections/insights* in relation to methodological issues implicated in the conduct of the trial. This includes any reference to the *context* for IPV group perpetrator programmes and any narrative that *provides authors’ observations/reflections/insights* related to the *process* of conducting such programmes (i.e., what factors should be considered in the conduct of IPV evaluation programmes) and that attempted to offer any explanations for the positive (or negative) findings.

H) Risk of bias. We used the Cochrane Collaboration Risk of Bias tool (RoB2) to assess the methodological quality of all eligible randomised controlled trials ([Higgins et al., 2011](#)). The RoB2 is comprised of six domains and assesses each study’s risk of bias as ‘high risk’, ‘low risk’, or ‘some concerns’.

Data Analysis

All data collection and analysis were conducted by one reviewer (WT) and checked for accuracy by members of the author team (HC, MH, KM & GF), with disagreements resolved through discussion. Qualitative meta-summary was conducted using Microsoft Word and Excel, starting with identification of eligible text from the relevant discussion section of the included studies. We extracted text that referred to methodological challenges of evaluating IPV group perpetrator programme effects. We separated text referring to methodological challenges from all other text in the papers and edited the statements on methodological challenges to ensure that it was presented in a way that was accessible to readers while preserving their underlying content and meaning ([Sandelowski & Barroso, 2003](#)). Methodological challenges were then grouped, abstracted, and considered separately.

Pre-selected methodological characteristics were used to group similar findings together. For example: context of study (e.g., setting, community or prison); sample characteristics (e.g., participants’ referral route [voluntary and court-mandated]); mental health status and substance abuse history; intervention (modality, implementation,

Table 1. Characteristics of Included Studies on Group-based IPV Male Perpetrator Programmes ($N = 15$)

Study, year, country, setting	Population (N)	Design	Intervention	Duration	Outcome Measures	Length of Follow-up	Results – Primary Outcomes
Palmer et al. (1992) Canada Counselling agency	56 men convicted of wife abuse, placed on probation and court-mandated to attend treatment.	Randomised controlled trial. Treatment ($n = 30$). Control (probation only) ($n = 26$).	Psychoeducational	10 weekly, 1.5 hours.	1. Recidivism (police records) 2. Basic Personality Inventory (BPI) (12 dimensions).	1-2 years	According to police records, there was evidence of physical abuse or serious threats to partners by 20% of all the men in the follow-up period, 12 to 24 months after treatment. Recidivism was significantly higher for the controls than for the men in treatment: 31% ($N = 26$) and 10% ($N = 30$), respectively ($p < .05$). There was no significant association between recidivism and the level of attendance at treatment (p. 280).
Saunders (1996) USA Family counselling agency	218 men referred by a deferred prosecution program (17%) or probation department following prosecution (59%). The others were referred by social service agencies, attorneys, friends, family members, or themselves (mean age = 32.4 years., $SD = 8.3$).	Randomised comparative effectiveness trial. FCBT ($n = 91$). PPT ($n = 87$).	Feminist-Cognitive-Behaviour Treatment (FCBT). Process-Psychodynamic Treatment (PPT).	20 weekly 2.5 hours.	1. Recidivism (Women's and men's report of violence-Conflict Tactics Scale [CTS]; arrest records). 2. Women's measures (fear, conflict resolution, general changes). 3. Men's measures (packet of self-report measures, incl. Mil-lon Clinical Multiaxial Inventory, relationship satisfaction, beliefs about woman abuse, self-esteem, General hostility, traditional values of women's role; democratic decision making; level of conflict; anger towards partner; jealousy; depression; adjustments for social desirability).	18-54 months after treatment.	There were no differences reported between treatments in rates of physical abuse or the women's fear levels, general perceptions of change in their partners, or ways of resolving conflicts (p. 409).
Dunford 2000 USA Family Advocacy Centre, San Diego Naval Base	861 married US Navy couples in which active-duty husbands were substantiated as having physically assaulted their wives (mean age = 27 years).	Randomised controlled trial. 4 groups: 1. Men's group ($n = 168$). 2. Conjoint group ($n = 153$). 3. Rigorous monitoring ($n = 173$). 4. Control ($n = 150$).	Cognitive-behavioural in both men's and conjoint groups.	1.5 hours weekly for 6 months and then monthly for another 6 months, for a total 1-year treatment period.	1. A self-reported episodic measure assessed the number of incidents or episodes in which a victim or perpetrator reported being abused across three different levels of abuse. 2. Modified Conflict Tactics Scale (MCTS). 3. Official police and court records for all respondents (both victims and perpetrators) living within the boundaries of San Diego County. 4. Date of the first instance in which a repeat case of spouse assault occurred as indicated by both official arrest records and victim reports of new physical injuries.	1-year follow-up after the first 6 months of treatment.	No statistically significant differences ($p < .05$) were found between the four experimental groups for the prevalence of continued abuse using victim reports of episodic spouse abuse (p. 471). No statistically significant differences between the groups were found for male perpetrator reports of the abuse of their wives (p. 473). Comparisons of data for the episodic measures, for both victim and perpetrator reports of new abuse produced no evidence that membership in any of the three experimental treatment groups was any more effective in reducing continued abuse than was membership in the control group (p. 474).
Taylor, Davis & Maxwell (2001) USA Community Setting serving perpetrators of domestic violence	376 male criminal court defendants charged with assaulting their intimate female partners (mean age = 33 years).	Randomised controlled trial: 3 treatment length groups: 1. Long ($n = 129$). 2. Brief ($n = 61$). 3. Control (community service, 40 hours) ($n = 186$).	Duluth model (Alternatives to Violence [ATV] program) (English or Spanish).	ATV: 40 hours (long = 26 weeks, brief = 8 weeks).	1. Criminal Justice records 2. Victim self-report surveys. Revised Conflict Tactics Scale used to assess frequency and severity of violence.	6 & 12 months after intake.	In the ATV treatment group, the failure rate (i.e., cases in which new violence occurred) is significantly smaller (about 50%), based on officially recorded incidents, at six months ($M = .38$ for the controls, and $M = .16$ for the treatment group; $f = 8.09$; $p < .01$) and again at 12 months ($M = .55$ for controls, and $M = .28$ for treatment group; $f = 6.82$; $p < .01$). In the 12-month follow-up using the experimental model, we found results nearly identical to those of the six-month model. This later model showed again that treatment was associated with a significantly lower prevalence of officially recorded failures ($b = -.61$, $p < .01$). At 12 months, however, the difference in incident rates between the two experimental groups had diminished (from $\text{Exp}(B) = .41$ to $\text{Exp}(B) = .55$). This suggests that the early effect of treatment may lessen with time (p. 191).

Table 1. Characteristics of Included Studies on Group-based IPV Male Perpetrator Programmes (*N* = 15) (continued)

Study, year, country, setting	Population (<i>N</i>)	Design	Intervention	Duration	Outcome Measures	Length of Follow-up	Results – Primary Outcomes
Feder & Dugan (2002) USA Community settings (x5) serving perpetrators of domestic violence	447 men convicted of misdemeanor domestic violence (mean age = 35 years).	Randomised controlled trial. Experimental group (one-year probation and court-mandated counselling) (<i>n</i> = 174). Control group (one-year probation only conditions) (<i>n</i> = 230).	Duluth Model	26 weekly sessions (time of weekly session not specified).	1. Abbreviated version of the Inventory of Beliefs About Wife Beating Scale. 2. A shortened six-item Attitudes Toward Women Scale. 3. Criminalization of domestic violence. 4. Attitudes about the partner's responsibility for the instant offense. 5. Self-reported likelihood to hit their partners again Conflict Tactics Scale (CTS2) 7. Official records of rearrest	6 months & 12 months (police records only).	The Inventory of Beliefs About Wife Beating showed that approximately 46% of the men viewed wife beating as acceptable behaviour in various situations. There were no significant differences between the experimental and control groups at Time 1, Time 2, or over time (p. 362). Men gave neutral answers on the Attitudes Toward Women Scale and did not show differences between groups or demonstrate changed attitudes over time. Analyses of variance (ANOVAs) of the differences of the five scales of the CTS2 over time suggested that there were no significant differences over time. These results indicate that the offenders' self-reported abusive behaviour did not change over time (p. 365). Crime records indicated that 24% of the men in both the experimental and control groups were rearrested on one or more occasions during their one-year probation. At a bivariate level, there are no significant differences between the experimental and control groups (p. 366).
Labriola, Rempel & Davis (2005) USA Community settings (x2) serving perpetrators of domestic violence	420 men arraigned on a IPV misdemeanor, convicted of a violation, and sentenced to a conditional discharge with a one-year protection order in favour of the victim.	Randomised controlled trial. Four conditions: 1. Batterer program + monthly monitoring (<i>n</i> = 102). 2. Batterer program + graduated monitoring (<i>n</i> = 100). 3. Only monthly monitoring (<i>n</i> = 109). 4. Only graduated monitoring (<i>n</i> = 109).	Duluth Model In total 202 were in a batterer program compared to 218 in control group being monitored.	75 minutes, over 26 weekly meetings.	1. Official re-arrest records 2. Victim reports	One-year post-sentence period.	Impact of batterer programs on re-arrest: Batterer programs did not produce a reduction in re-arrests. Impact of judicial monitoring schedule on re-arrest: Neither form of monitoring proved more effective than the other (p. vii). Survival analysis: None of the interventions under examination outperformed any other in delaying the onset of recidivism; there were no significant differences in the average number of crime-free days prior to first re-arrest. Predictors of recidivism: The strongest predictors of future recidivism were prior criminal history, younger age, lack of a "stake-in-conformity" (e.g., stemming from employment or living with the intimate partner), and more serious current arrest charges. Prevalence of victim reports of re-abuse: 46% of the women interviewed reported experiencing at least one incident of re-abuse in the year after sentencing: 15% reported physical abuse, 18% reported threats, and 44% reported other forms of abuse.

Table 1. Characteristics of Included Studies on Group-based IPV Male Perpetrator Programmes ($N = 15$) (continued)

Study, year, country, setting	Population (N)	Design	Intervention	Duration	Outcome Measures	Length of Follow-up	Results – Primary Outcomes
Easton et al. (2007) USA outpatient substance abuse treatment facility	85 alcohol-dependent males who were arrested for domestic violence.	Randomised comparative effectiveness trial. SADV ($n = 40$) TFS ($n = 38$)	Cognitive behavioural Substance Abuse Domestic Violence (SADV) Twelve-Step Facilitation (TSF)	1.5 hours weekly over 12 weeks	1. Addiction Severity Index (ASI). 2. (Weekly assessments) included breathalysers, onsite urine toxicology screens (Roche Diagnostic's Testcup5 with adulteration checks), as well as self-reported alcohol and drug use (marijuana, cocaine, opioids, benzodiazepines, phencyclidine, and other drug use). 3. Revised Conflict Tactic Scale for Couples (CTS2). 4. Collateral Reports of Violent Behaviour (CTS2).	weekly, monthly, at 12 weeks post-treatment, and at a six-month.	SADV reported 90.2 days abstinent ($SD = 13.7$) while the TSF group reported a mean of 79.8 ($SD = 23.1$) days abstinent across the 12 weeks of treatment [$F = 5.4, p < 0.02$] (p. 29). No significant difference between groups during the 12 weeks of treatment on breathalyser tests [$F = 0.12, p < 0.73$] nor was there any significant difference between groups that tested positive for drug use during the 12 weeks of treatment [$F = 1.7, p < 0.2$] (p. 29). SADV condition reported decreases in physical violence within the past 30 days from pre-treatment (42.1%) to post-treatment (10.0%), while patients in the TSF condition reported a change from 22.2% at pre-treatment to 6.9% at post-treatment [$\chi^2 = 7.0, p < 0.03$] (p. 29). The six-month follow-up outcomes, there was no significant difference between participants in the SADV versus the TSF condition on the ASI medical [$F = 0.19, p < 0.67$], legal [$F = 0.03, p < 0.86$], other drugs [$F = 0.08, p < 0.77$], alcohol [$F = 0.04, p < 0.84$], employment [$F = 2.1, p < 0.15$], family [$F = 0.001, p < 0.97$], or psychological [2.1, $p < 0.15$] ASI composite scores (p. 30).
Taylor & Maxwell (2009) USA County Sheriff's Department's Jail	629 men arrested for a misdemeanor or felony IPV offense with no criminal gang affiliations and no prior history of serving a sentence in prison for one year or greater (mean age = 32.8 years).	Randomised controlled trial. Treatment: Batterer treatment wing of the jail ($n = 317$). No-treatment control group: In another wing of the jail. ($n = 312$).	Duluth Model	Classes held daily and were about three hours in length. Average length of incarceration was about 5 days spent in this special wing of the jail.	Recidivism/Violence included: 1. Time to failure (i.e., the amount of time between the conclusion of the treatment/control condition and any new repeat offenses); 2. Prevalence of failure (i.e., the proportion of batterers that committed new offenses); 3. Incidence/frequency of failure; 4. Severity of failure (related to the use of controlling behaviour, psychological abuse, threats of physical assault, or actual physical and sexual assault); 5. Self-reported violence: Modified version of the Conflict Tactics Scale; 6. Drug and Alcohol use; 7. Official police arrest data.	6 months (victims & perpetrators). Official police arrest data on recidivism were collected and analysed for up to one year post arrest.	Arrestee/6-month follow-up measures of IPV recidivism: There was no difference between the treatment and control group in 6-month prevalence, frequency, and time-to-failure for batterer reported acts of physical abuse (p. 5). Arrestee/batterer based 6-month follow-up measures of alcohol and drug use: Two statistically significant differences were observed in batterer self-reported alcohol and marijuana use. None of the other drug measures (related to other drugs, frequency, or dependency) were statistically significant (p. 5). Victim-based accounts of IPV recidivism results: No difference between the treatment and control group in 6-month prevalence, frequency, and time-to-failure for victim reported acts of controlling behaviour by the perpetrator. Also, there was no difference between the treatment and control group in 6-month prevalence, frequency, and time-to-failure of victim self-reported acts of any IPV. Finally, there was no difference between the treatment and control group in 6-month prevalence, frequency, and time-to-failure for victim self-reported acts of physical abuse (p. 5). Police IPV arrest results: The treatment and control groups both had over 65 percent of the men in the sample re-arrested for a new IPV offense at six months post intervention. The treatment and the control groups both had over 68 percent of the men in the sample re-arrested for a new IPV offense at twelve months post intervention (p. 5).

Table 1. Characteristics of Included Studies on Group-based IPV Male Perpetrator Programmes ($N = 15$) (continued)

Study, year, country, setting	Population (N)	Design	Intervention	Duration	Outcome Measures	Length of Follow-up	Results – Primary Outcomes
Alexander et al. (2010) USA Community setting serving victims and perpetrators of IPV	528 male perpetrators (96.1% court-ordered) (mean age = 34.18 years).	Randomised comparative effectiveness trial. SOCMI ($n = 247$) CBTGR ($n = 281$)	Stages of change motivational interviewing (SOCMI). Cognitive behavioural therapy gender re-education (CBTGR).	26 weeks (session length not specified).	Perpetrator data: 1. Conflict Tactics Scales-Revised (CTS2; an eight-item Psychological Aggression and 12-item Physical Assault subscales) and 2. The University of Rhode Island Change Assessment (URICA; a 32-item scale with subscales for precontemplation, contemplation, action, and maintenance). Victim/partner data: 1. The CTS2 Psychological Aggression and Physical Assault subscales and 2. The Danger Assessment Scale (DAS).	6 and 12 months post intake.	No differences found between the two treatment conditions with respect to men's self-reports of violence at the end of treatment, but did emerge in partners' reports of physical aggression at follow-up. Significantly fewer partners of men assigned to the SOCMI treatment condition as opposed to the CBTGR condition reported having experienced physical aggression at follow-up (p. 582).
Palmstierna et al. (2012) Norway Specialised outpatient mental health service	26 male perpetrators voluntary seeking therapy (mean age = 35 years).	Randomised controlled trial. Treatment ($n = 15$). Waiting list ($n = 11$).	Cognitive behaviour therapy (manualized).	2 hours weekly sessions over 15 weeks.	CTS extended version (incl. 'verbal aggression' and 'physical aggression/ violence', and 'material violence').	15 weeks after treatment. <i>Note.</i> Waiting list group was reassessed after 4 months on the waiting list.	Only 15 of the 26 previously violent men reported violence of any kind. Only four of the 21 men who had reported a history of physical violence still reported such behaviour, whereas six of the 19 men who had reported verbal aggression still reported such behaviour. A reduction in violence towards property was also reported. All these changes were significant (p. 363).
Tollefson and Phillips, 2015 USA Domestic violence treatment agency in the Salt Lake City, Utah	90 men who had been ordered by the court to complete a domestic violence offender treatment program.	Randomised controlled trial. Comparison program group ($n = 46$). MBB group ($n = 44$).	The comparison program incorporated an eclectic mix of interventions. Mind-Body Bridging Domestic Violence Program.	20 sessions weekly, 1 hour. 16 sessions weekly, 1 hour.	1. Recidivism (police record) 2. SF-36 Health Survey (SF-36)	45 days – 875 days.	According to police records, 6.7% of participants relapsed during the follow-up period. There was no difference in recidivism between the MBB group and the comparison group.
Taft et al. (2016) USA Department of Veteran Affairs hospitals (x2)	135 male veterans or service members (clinician-referred, self-referred, and court-referred) (mean age = 37.85 years).	Randomised controlled trial. SAH-M ($n = 67$) ETAU ($n = 68$)	Strength at Home Men's Program (SAH-m). Enhanced treatment as usual (ETAU).	2-hour weekly group sessions over 12 weeks.	1. Mini-International Neuropsychiatric Interview (MINI). 2. Clinician- Administered PTSD scale (CAPS). 3. Physical and psychological Violence: Conflict Tactics Scales (CTS2) (Physical Assault and Psychological Aggression subscales). 4. MMEA: Psychological IVP	3 and 6 months after baseline (perpetrators and partner assessments).	Primary analyses using hierarchical linear modelling indicated significant time-by-condition effects such that SAH-M participants compared with ETAU participants evidenced greater reduction in physical and psychological aggression ($B = -0.135$ [$SE = 0.061$], $p = .29$; $B = -0.304$ [$SE = -0.135$], $p = .026$, respectively). Additional analyses of a measure that disaggregated forms of psychological IPV showed that SAH-M, relative to ETAU, reduced controlling behaviours involving isolation and monitoring of the partner. ($B = -0.072$ [$SE = 0.027$], $p = .0105$) (p. 1168).

Table 1. Characteristics of Included Studies on Group-based IPV Male Perpetrator Programmes ($N = 15$) (continued)

Study, year, country, setting	Population (N)	Design	Intervention	Duration	Outcome Measures	Length of Follow-up	Results – Primary Outcomes
Lila et al. (2018) Spain community based batterer intervention program (BIP)	160 males convicted of IPV and court mandated to BIP (mean age = 40.66 years).	Randomised comparative effectiveness trial. SBIP + IMP ($n = 80$). SBIP ($n = 80$)	Standard batterer intervention program (SBIP) + Individualized motivational plan (SBIP + IMP). Standard batterer intervention program (SBIP). SBIP consisted of 70 hours of a cognitive-behavioural intervention. IMP: (a) individual motivational interviews; (b) three group sessions; (c) therapists' reinforcement of goals in all group sessions; (d) use of retention techniques.	SBIP: 35 weekly group sessions, each lasting 2 hr. IMP: (a) five individual motivational interviews (approx. 1 hour each); (b) three group sessions; (c) therapists follow-up.	Recidivism data: 1. Official databases 2. Self-reported recidivism 3. Therapists' assessment of recidivism risk. Proximal outcome: 1. Treatment compliance (dropout and intervention dose). 2. Stage of change.	9 months (end of the intervention program); and at 15 months (official recidivism data).	Findings indicated that the SBIP plus IMP participants received significantly more intervention dose ($R^2 = .08$), finished the intervention in a more advanced stage of change (ITT, $R^2 = .17$; PP, $R^2 = .22$), reported less physical violence after treatment (ITT, odds ratio = .63; PP, odds ratio = .34), and had a higher reduction in recidivism risk (ITT, $R^2 = .64$; PP, $R^2 = .56$) than SBIP participants (p. 309).
Murphy et al. (2020) USA Community-based domestic violence agency	42 male perpetrators (mean age = 34.38 years).	Randomised comparative effectiveness trial. ICBT ($n = 21$) GCBT ($n = 21$)	Individual Cognitive behavioural therapy (ICBT). Group Cognitive Behavioural Approach (GCBT).	ICBT: 1 hour, 20 weekly sessions. GCBT: 2 hours, 20 weekly sessions.	A. Partner-abusive behaviour: 1. CTS2 2. Multidimensional Measure of Emotional Abuse. B. Relationship adjustment (Dyadic Adjustment Scale (DAS)). C. Participant communication difficulties (27-item Spouse Verbal Problem Checklist (VPC)) D. Criminal justice outcomes (publicly available electronic state database)	3, 6, 9 and 12 months after baseline (perpetrator and partners).	Participant self-reports revealed significant reductions in abusive behaviour and injuries across conditions with no differential benefits between conditions. Victim partner reports revealed more favourable outcomes for group treatment (GCBT), including a statistically significant difference in psychological aggression, and differences exceeding a medium effect size for physical assault, emotional abuse, and partner relationship adjustment. Treatment competence ratings suggest that flexible, individualized administration of CBT creates challenges in session agenda setting, homework implementation, and formal aspects of relationship skills training. Although caution is needed in generalizing findings from this small-scale trial, the results suggest that the mutual support and positive social influence available in group intervention may be particularly helpful for IPV perpetrators (p. 2847).
Nesset et al. (2020) Norway University Hospital, Forensic Department and Research Centre	144 men voluntary referred by general practitioners.	Randomised comparative effectiveness trial. CBGT ($n = 67$) MBSR ($n = 58$)	Cognitive behavioural group therapy (CBGT). Mindfulness-based stress reduction group therapy (MBSR).	CBGT: 2 individual sessions followed by 15 two-hour group therapy sessions (total 30 h). MBSR: 1 individual session before and 1 session after 8 group sessions of MBSR group therapy (16 h).	Violent behavior at 12 months' follow-up. Norwegian version of the revised Conflict Tactics Scales (CTS2).	3, 6, 9 and 12 months after baseline. (perpetrators and partners).	Both the intervention and the comparator group showed substantial reductions in violent behaviour during 12 months follow-up, and no time-by-condition differences between the CBGT-group and the MBSR-group could be found. This finding was consistent across all dimensions of the primary outcome as measured by the CTS2, based on a combination of the highest reported level of violence from the participant or his partner over 1 year (p. 7).

duration & evaluation); and outcomes (type, duration of outcome measurement).

We abstracted findings (e.g., authors' narrative) to summarise the content of each group of methodological challenges. For each study characteristic, we reworked our lists of methodological challenges until we developed a new list of abstracted statements that captured the overall meaning of the original statements. This was done by eliminating redundancies, refining statements to ensure they were inclusive of the authors' interpretations, preserving contradictions and ensuring clarity and accessibility.

Lastly, we provided a narrative summary of the abstracted methodological issues to identify whether any specific methodological challenges (pertaining to study/intervention characteristics) were implicated in positive (or negative) programme effects.

Frequency of methodological challenges were also used to extract more meaning from the narrative summaries by numerically describing the magnitude of the abstracted findings (Sandelowski & Barroso, 2003). Frequency of methodological challenges were calculated by dividing the number of studies contributing data on specific methodological challenges by the total number of studies.

Results

We identified 23 systematic reviews/ meta-analytic evaluations of trials on the effects of male perpetrator programmes (Akoensi et al., 2013; Aos et al., 2006; Arce et al., 2020; Arias et al., 2013; Babcock et al., 2004; Cheng et al., 2021; Feder et al., 2008; Feder & Wilson, 2005; Ferrer-Perez & Bosch-Fiol, 2018; Karakurt et al., 2019; Miller et al., 2013; Nettet et al., 2019; NICE, 2013; O'Connor et al., 2021; Pinto E Silva et al., 2022; Santirso et al., 2020; Satyen et al., 2022; Smedslund et al., 2011; Stephens-Lewis et al., 2021; Tarzia et al., 2020; Travers et al., 2021; Vigurs et al., 2016; Wilson et al., 2021). These 23 systematic reviews and meta-analyses were used to identify primary studies evaluating IPV perpetrator programmes, which is the main focus of this paper. After removing duplicates, we identified 15 randomised controlled trials from the 23 review papers. We additionally identified quasi-experimental and pre/post-test design studies which informed our thinking on methodological challenges in RCTs but are not included here.

Description of Randomised Controlled Trials

The 15 RCTs on the effects of male perpetrator programmes span the period 1992 – 2020 (see Table 1). Eleven were conducted in the United States (Alexander et al., 2010; Dunford, 2000; Easton et al., 2007; Feder & Dugan, 2002; Labriola et al., 2005; Murphy et al., 2020; Saunders, 1996; Taft et al., 2016; Taylor et al., 2001; Taylor & Maxwell, 2009; Tollefson & Phillips, 2015), one in Canada (Palmer et al., 1992), one in Spain (Lila et al., 2018) and the remaining two in Norway (Nettet et al., 2020; Palmstierna et al., 2012).

Setting

Eleven RCTs were conducted in community settings (e.g., family advocacy centres and/or agencies serving perpetrators and victims/survivors of domestic violence); two in outpatient hospital facilities (e.g., substance misuse treatment facility in the Easton et al. (2007) study and specialised mental health services in the Palmstierna et al (2012) study). The sheriff department jail was the setting for the Taylor & Maxwell (2009) study while the Nettet et al (2020) study was conducted under the auspices of a hospital-affiliated forensic research department.

Population

In total, the trials had a population of 4,216 men; in most cases men were court-mandated to treatment. Only a small minority of men were voluntary referrals to the IPV programme (Nettet et al., 2020; Palmstierna et al., 2012).

Intervention Model

The Duluth and the cognitive-behavioural model were the two dominant theoretical paradigms that informed the interventions tested in the studies. The study by Palmer et al. (1992) tested the effects of a brief psychoeducational programme with limited information provided about its content and the study by Tollefson and Phillips (2015) tested a programme using mind-body bridging techniques to prevent explosive states of mind.

Research Design

Five studies compared Duluth-informed interventions with no treatment (Feder & Dugan, 2002; Labriola et al., 2005; Taft et al., 2016; Taylor et al., 2001; Taylor & Maxwell, 2009) and three compared CBT with a control (Dunford, 2000; Palmstierna et al., 2012; Taft et al., 2016). The trial by Palmer et al. (1996) compared the effects of a psychoeducational programme to a control group. The study by Tollefson and Phillips (2015) compared a mind-body programme with a standard psychoeducational IPV perpetrator programme. Six were comparative effectiveness trials involving group-based CBT compared with either psychodynamic treatment (Saunders, 1996), twelve-step facilitation (TSF) (Easton et al., 2007), Mindfulness-based stress reduction (MBSR) (Nettet et al., 2020), individually delivered CBT (Murphy et al., 2020) or enhanced CBT treatment incorporating motivational interviewing (Alexander et al., 2010; Lila et al., 2018). None of the studies were cluster-randomised trials.

Duration

Group-based sessions lasted between 60 minutes and 3 hours and were in most cases delivered weekly. The exception is the Taylor & Maxwell (2009) study where sessions were held daily (prison setting). The duration of interventions ranged from 8 weeks (Taylor et al., 2001) to a year (Dunford, 2000).

Length of Follow-up

There were variations in the length of follow-up. The longest follow-up period of 54 months was noted for the Saunders (1996) study. Except for the Easton et al. (2007) study which incorporated weekly and monthly follow-ups, most studies included 6 and 12 months follow-up.

Outcome Measures

Table 1 lists the range of measures used to assess outcomes across studies. There is considerable heterogeneity in the type and number of measures researchers have employed to assess outcomes with a particular focus on recidivism data (based on preparators' and/or victim/survivors' self-reports and official/police data), psychological profile data (based on men's, victim/survivors' self-reports and therapists' assessments) and in some instances other proximal outcomes (e.g., treatment compliance (Lila et al., 2018)).

Overview of Findings

We did not perform a meta-analysis due to heterogeneity between studies in terms of setting, outcome and follow-up. To facilitate

Table 2. Frequency of Methodological Challenges in Randomised Controlled Trials ($n = 15$) of Group-based Programs for IPV Perpetrators

Study	Research Design	Context	Procedural					
			Sample	Intervention		Measurement		
				Treatment Modality	Fidelity Monitoring	Source of Outcome Data	Duration of Follow-up	Attrition
Palmer et al. (1992)		X	X			X		
Saunders (1996)			X	X		X		
Dunford (2000)	X		X	X				
Taylor et al. (2001)	X		X	X		X	X	X
Feder and Dugan (2002)	X		X	X		X		X
Labriola et al. (2005)		X	X	X		X		
Easton et al. (2007)			X	X	X	X	X	
Taylor and Maxwell (2009)						X	X	X
Alexander et al. (2010)			X	X	X	X	X	X
Palmstierna et al. (2012)						X	X	
Taft et al. (2016)							X	X
Tollefson and Phillips (2013)		X		X		X		
Lila et al. (2018)				X		X	X	X
Murphy et al. (2020)			X	X				X
Nesset et al. (2020)				X		X		X
Frequency of Methodological Challenges	3/15	3/15	9/15	11/15	2/15	12/15	7/15	8/15

further appreciation of the results of the meta-summary approach relating to methodological challenges, we considered the authors' own assessments of their primary statistically 'significant' findings which are summarised below.

We identified five studies (Lila et al., 2018; Palmer et al., 1992; Palmstierna et al., 2012; Taft et al., 2016; Taylor et al., 2001) reporting a significant reduction in official (police recorded) and/

or men's self-reported rates of physical violence towards partners. Testing a brief psychoeducational programme, Palmer et al. (1992) reported significantly higher recidivism rates for controls than men in treatment (31% vs. 10% respectively). Taylor et al. (2001) tested the effects of a Duluth-informed programme and noted a 50% reduction in new violence at 6-months ($M = 0.38$ for controls and $M = 0.10$ for treatment) and at 12 months ($M = 0.55$ for controls and $M = 0.28$

Table 3. Randomised Controlled Trials that Reported How They Managed Study or Intervention Attrition

Study	How to Overcome Attrition	Details
Alexander et al. (2010)	Statistical analysis	They did not report how they performed analyses for attrition.
Feder and Dugan (2002)	Statistical analysis	The authors had difficulty tracking victims, while a high percentage of defendants' non-response was due to defendants' refusal to be interviewed. Therefore, the authors performed attrition analyses to test for differential response rates.
Lila et al. (2018)	Statistical analysis	They performed univariate analyses to check for sociodemographic and behavioural differences between completers and non-completers. Significant variables (income and risk of recidivism) were added to the regression model as covariates/confounders.
Murphy et al. (2020)	Statistical analysis	14% of individuals assigned to the group condition dropped out between baseline assessment and treatment. Follow-up was limited to those who attended at least one treatment session. Criminal justice outcomes were gathered on the entire intent-to-treat sample, and those who dropped out did not differ significantly from those exposed to treatment.
Nesset et al. (2020)	Contacting participants and statistical analysis	Participants received incentives and reminders to participate in the survey. In addition, the authors used a linear mixed model because it is less sensitive to missing data.
Taft et al. (2016)	Contacting participants, flexible session schedules and statistical analysis	The therapist would call participants when they missed any sessions and sessions were scheduled in the evening to avoid conflict with work schedule. All analyses were conducted on the intention-to-treat sample.
Taylor et al. (2001)	Statistical analysis	Authors performed statistical analysis to verify if there were differences between participants and nonparticipants and victim's ethnicity. Furthermore, they used missing information as a covariate in regression models.
Taylor and Maxwell (2009)	Statistical analysis	They did not find any statistical difference between completers and non-completers using sociodemographic variables.
Tollefson and Phillips (2015)	Statistical analysis	Attrition (programme completion or not) was used as an outcome variable to identify its factors associated.

Unique ID	Study ID	Experimental	Comparator	Outcome	Weight	D1	D2	D3	D4	D5	Overall
V1	Palmstierna, 2012	CBT	Waiting list	CTS	1	!	+	+	-	-	-
V2	Lila, 2018	Individualized motivational plan + BIP	Batterer intervention program (BIP)	Recidivism	1	+	+	+	+	!	+
V3	Taft, 2016	Strength at home men's program	TAU	CTS	NA	+	+	+	-	-	-
V4	Taylor, 2001	Batterer treatment	Community service	Recidivism	NA	+	+	+	+	!	!
V5	Palmer, 1992	Psychoeducation	Control group	Recidivism	1	+	+	+	+	!	!
V6	Saunders, 1996	Feminist CBT	Process- psychodynamic	CTS	NA	+	-	+	-	-	-
V7	Alexander, 2010	SOCMI	CBTGR	CTS	1	+	+	+	+	+	+
V8	Dunford, 2000	CBT	Control group	CTS	NA	+	+	+	-	!	-
V9	Easton, 2007	CBT	Twelve step	CTS	NA	+	+	+	+	!	!
V10	Feder and Dugan, 2002	Duluth model	Control group	CTS	NA	+	+	+	-	-	-
V11	Labriola, 2005	Duluth model	Control group	Recidivism	NA	+	-	+	+	!	-
V12	Murphy, 2017	Individual CBT	Group CBT	CTS	NA	!	+	+	-	-	-
V13	Nesset, 2020	Mindfulness-based stress reduction	Group CBT	CTS	NA	+	+	+	-	-	-
V14	Taylor and Maxwell, 2019	Duluth-informed intervention	Control group	Recidivism	NA	+	+	+	+	-	-
V15	Tollefson and Phillips, 2015	Mind-Body Bridging	Comparison program	Recidivism	NA	+	!	+	+	-	-




	Low risk	D1 Randomisation process	D4 Measurement of the outcome
	Some concerns	D2 Deviations from the intended interventions	D5 Selection of the reported result
	High risk	D3 Missing outcome data	

Figure 2. Risk of Bias Assessment for Randomised Controlled Trials.

Note. CBT = Cognitive Behavioural Therapy; CTS = Conflict Tactics Scale; BIP = Batterer Intervention Programme; SOCMI = Stages-Of-Change Motivational Interviewing; TAU = Treatment as usual; CBTGR = Cognitive Behavioral Therapy Gender Reeducation.

for treatment) based on official records. [Palmstierna et al. \(2012\)](#), testing a 15-week CBT-informed programme, reported that most of the men had reduced their self-reported violent behaviour. [Taft et al. \(2016\)](#), testing a 12-week trauma-informed intervention, noted greater reductions in self-reported physical and psychological IPV use particularly related to controlling behaviour involving isolation and monitoring. Finally, [Lila et al. \(2018\)](#) showed that adding motivational strategies to a standard intervention (i.e., cognitive-behavioural) programme can produce greater reductions in self-reported violence and recidivism risk of IPV relative to the standard programme alone.

Among the ten studies reporting negative results, only four compared the control and intervention groups ([Dunford, 2000](#); [Feder & Dugan, 2002](#); [Labriola et al., 2005](#); [Taylor & Maxwell, 2009](#)), whilst the other six studies evaluated the effectiveness of two or more distinct types of intervention to reduce IPV ([Alexander et al., 2010](#); [Easton et al., 2007](#); [Murphy et al., 2020](#); [Nesset et al., 2020](#); [Saunders, 1996](#); [Tollefson & Phillips, 2015](#)). Six studies used some type of CBT approach and one of them compared CBT versus psychodynamic therapy ([Saunders, 1996](#)). In addition, session lengths in hours were shorter among RCTs that reported negative results ($M = 29.72$) compared to those that were successful in preventing IPV ($M = 35.8$).

Methodological Challenges

The methodological challenges articulated in discussion and/or conclusion sections of the 15 RCTs fell into three broad

categories: contextual, research design specific, and procedural. The last category was further subdivided into three topic areas: sample, intervention (modality, duration, implementation) and measurement (type and duration of outcome measurement). These challenges are summarised below.

Frequency of Methodological Challenges

All 15 RCTs identified at least one challenge involved when conducting IPV group perpetrator programmes ([Table 2](#)). The main challenges reported, as measured using frequency of methodological challenges, were related to source of outcome measure used (12/15), treatment modality (11/15), sample (9/15), duration of follow-up (7/15), and challenges due to study or intervention attrition (8/15). Challenges relating to the implementation of the research design (3/15) and contextual challenges were the least discussed (3/15).

Furthermore, among the 15 RCTs, only eight reported how they tackled study attrition (numbers of participants giving feedback and filling in questionnaires) or intervention attrition (numbers of participants continuing to attend the intervention or treatment, [Table 3](#)) ([Alexander et al., 2010](#); [Feder & Dugan, 2002](#); [Lila et al., 2018](#); [Nesset et al., 2019](#); [Taft et al., 2016](#); [Taylor et al., 2001](#); [Taylor & Maxwell, 2009](#); [Tollefson & Phillips, 2015](#)). Among them, seven used different statistical techniques to control for loss to follow-up (study attrition) and two studies used statistical methods and contacted participants via phone calls or reminders to avoid loss to follow-up ([Nesset et al.,](#)

2020) or to avoid loss to intervention (Taft et al., 2016). Study attrition and intervention attrition can be related such as when a participant may be unwilling or unable to continue participation in the trial because they become frustrated with the intervention in some way and/or other aspects of their life makes involvement difficult. For example, a male participant's partner might leave the relationship and he may blame this on the intervention, stop attending the groups and refuse to complete any further follow-up questionnaires.

In relation to risk of bias, ten studies were classified as at high risk of bias, two at low risk, and three studies as having some concerns (Figure 2). There was a slight tendency for the studies with the lower risk of bias rating to identify and report more methodological challenges (Table 2). The mean number of challenges reported for lower risk of bias studies (Alexander et al., 2010; Easton et al., 2007; Lila et al., 2018; Palmer et al., 1992; Taylor et al., 2001) was 5 compared to a mean of 3.4 for higher risk studies (Dunford, 2000; Feder & Dugan, 2002; Labriola et al., 2005; Murphy et al., 2020; Nessel et al., 2020; Palmstierna et al., 2012; Saunders, 1996; Taft et al., 2016; Taylor & Maxwell, 2009; Tollefson & Phillips, 2015).

Contextual

There is acknowledgment that these intervention programmes operate within specific judicial systems which have an ethical, as well as legal, responsibility to facilitate their smooth implementation. Three studies highlight context-related challenges. Palmer et al. (1992) refers to the *legal system's responsibility* to address attrition and ensure treatment attendance by pursuing group participants who fail to attend with whatever sanctions are available in achieving this aim (Ford & Regoli, 1993). Labriola et al. (2005) note that research sites should ideally be engaged in rigorous monitoring practices (rather than simply checking-in) facilitated by clear and frequent communication of court policies and expectations from perpetrators. Tollefson and Phillips (2015) recognise that it might be hard to generalise the findings to other areas as usual care for convicted domestic abuse offenders in their area (Utah) was a minimum of a 16-week IPV perpetrator programme.

Research Design Specific

Dunford (2000) suggests that the risks of conducting RCTs to assess interventions are likely to be fewer than the consequences of doing so. This observation underscores the need to use rigorous designs to evaluate interventions for domestic violence. Aspects of randomised trial design were difficult to implement in two studies. Taylor et al. (2001) needed to make substantial concessions to court officials to gain their cooperation, allowing judges to override assignment to the control condition. Feder and Dugan (2002) highlighted random assignment as essential for reducing bias but acknowledged this could be problematic; some of the agencies involved in this study were resistant to the experimental design and would not help in contacting the victim/survivors thereby directly contributing to low victim/survivor response rates.

Procedural

Individuals' Needs. Among included studies, six of them refer to 'sample' as a contributing factor in explaining treatment effects (Alexander et al., 2010; Easton et al., 2007; Feder & Forde, 2000; Palmer et al., 1992; Saunders, 1996; Taylor et al., 2001). There was explicit reference to the need to identify and tailor the content of interventions to match types of participants to improve outcomes (Dunford, 2000; Labriola et al., 2005; Murphy et al., 2020). Two studies explicitly question the assumption of the 'one size fits all' approach (Dunford, 2000; Saunders, 1996). Study authors have articulated the heterogeneity of

trial participants especially with reference to: a) their mental health status (Palmer et al., 1992); b) their childhood traumas (Saunders, 1996); c) motivation to change their behaviour even if they have been mandated to treatment (Alexander et al., 2010; Taylor et al., 2001); d) their stakes in conforming with court decisions (Feder & Forde, 2000); and e) their substance abuse histories (Easton et al., 2007).

Intervention

Intervention or programme challenges featured prominently in the discussion section of the included studies. Challenges were organised by whether they referred to: a) treatment modality; b) duration and length of treatment; and c) treatment implementation. Each of these are examined below.

Treatment Modality

Taylor et al. (2001) were the first to explicitly question the dominance of Duluth-informed treatment and proposed that treatments informed by a client-centred approach that seek to engage clients as much as possible rather than confront participants (as some Duluth programmes may do) may be most useful. A similar observation was shared by Feder and Dugan (2002) advocating the need for new and innovative IPV perpetrator programmes to reduce IPV to help victim/survivors. Similarly, Labriola et al. (2005) encouraged the development of intervention programmes with different conceptual bases and curricula considering the typology of participants and adopting a more direct behavior modification approach. Tollefson and Phillips (2015) had concerns about cross contamination of the interventions as the same facilitator delivered both the mind-body bridging programme and the usual care programme, and a small dose of mind-body techniques were delivered in the usual care programme. The trial by Easton et al. (2007) that targeted substance (alcohol and illicit drugs) abuse and co-occurring IPV and the studies by Alexander et al. (2010) and Lila et al. (2018), which incorporated stages-of-change motivational interviewing, provide good attempts to explicitly target participants' characteristics. However, the combination of several intervention strategies makes it difficult to disentangle the extent to which the interventions' effects are attributable to each of these strategies (Lila et al., 2018; Murphy et al., 2020; Nessel et al., 2020).

Treatment Implementation

We found little reference to treatment implementation problems in the included trials. Easton et al. (2007) note that the findings may be confounded by therapist effects (only one therapist delivered either the SADV or TSF conditions) and Alexander et al. (2010) highlighted the low number of audiotaped sessions which precluded analysis of the differential impact of therapist adherence on treatment modality.

Measurement

We found extensive reference to various measurement problems in the included trials.

Source of Outcome Measurement

Palmer et al. (1992), Saunders (1996) and Taylor and Maxwell (2009) questioned the reliance on police reports as a measure of recidivism and highlighted the importance of cooperation of partners or ex-partners in assessing treatment outcomes, an observation echoed by other researchers (Alexander et al., 2010; Easton et al., 2007; Feder

& Forde, 2000; Lila et al., 2018; Nessel et al., 2020; Taft et al., 2016; Taylor et al., 2001). While victim/survivor reports are more credible in assessing continuation of abuse (Lila et al., 2018; Nessel et al., 2020; Palmstierna et al., 2012), low victim/survivor response rates (below 25%) were common and particularly problematic due to potential bias from getting responses from a small sample of victim/survivors (Alexander et al., 2010; Feder & Dugan, 2002; Labriola et al., 2005; Taylor et al., 2001). Easton et al. (2007) highlighted the importance of independently assessing outcomes with reliability scores being provided. That said, Easton et al. (2007) only managed to collect collateral data from female partners of 55% of the male participants; and Lila et al. (2018) did not collect any as Spanish legislation prevents the services referring IPV perpetrators to intervention programmes from providing information that would allow access to their victims.

Duration of Follow-up

The short duration of follow-up was noted in six trials (Alexander et al., 2010; Easton et al., 2007; Lila et al., 2018; Palmstierna et al., 2012; Taft et al., 2016; Taylor & Maxwell, 2009). Davis et al. (2000) suggested that both short- and long-term outcomes need to be included; short-term to assess transient programme effects and long-term outcomes to determine whether the intervention leads to permanent changes.

Attrition

There was extensive reference to the effects of treatment attrition on data quality (Feder & Dugan, 2002; Nessel et al., 2020; Palmer et al., 1992) and the initial randomisation (Murphy et al., 2020; Saunders, 1996). Taylor et al. (2001) highlighted the need for researchers to find ways of minimising study attrition while maximising interview response rates when interviewing victim/survivors about continuing abusive behavior from their partners.

We found little explicit reference to the effects of treatment duration and length of treatment in the included trials. Taylor and Maxwell (2009) refer to the short duration of the prison-based Duluth-informed intervention. In a re-analysis of the Brooklyn experiment, Maxwell et al. (2020) concluded that 'there is little support that longer programs are more effective than shorter ones' (p. 493). This finding runs contrary to the earlier summary observation that session length was shorter among the randomised controlled trials that reported negative results compared to those that were successful in preventing IPV. This assertion also goes against the more recent conclusions of Arce et al. (2020) who in their meta-analysis found that longer programmes (those lasting over 16 weeks) produced more effective outcomes.

Discussion

Through extensive searching and mapping of studies reported in systematic reviews, we identified 15 RCTs examining the effects of group-based IPV perpetrator programmes. Seven studies were comparative effectiveness trials. With two exceptions, Duluth or cognitive-behavioural models informed the group-based interventions which included mostly court-referred participants. A range of measures were employed, including psychological and treatment-related, with recidivism as the primary outcome substantiated by official/police records, self-reports from perpetrators and/or their victim/survivors followed over a period of 15 weeks to 12 months in the majority of cases.

Five studies reported statistically significant findings relating to violence reduction either substantiated by police records or self-reported. The findings relate to the effects of a psychoeducational programme for men convicted of domestic assault (Palmer et al., 1992), a Duluth-informed programme for men charged with domestic abuse

(Taylor et al., 2001) and two CBT-informed programmes for voluntary/clinician-referred men (Palmstierna et al., 2012; Taft et al., 2016).

The use of the qualitative meta-summary approach facilitated the emergence of important insights on several methodological (i.e., contextual, research design and procedural) challenges which may contribute to the heterogeneity of findings from trials and the persistent uncertainty about effects of perpetrator programmes. Evidenced by the frequency of methodological challenges, major challenges include the source of outcome measurement (whether self-report and/or victim/survivor-collaborated), treatment modality, sample (i.e., participants' characteristics), and duration of follow-up. Challenges identified with less frequency relate to the context of the trial and implementation of the research design, although identification of these problems by trialists in fewer studies does not mean they are irrelevant.

There was agreement on the importance of corroborating men's self-reports of violence with both official (police, e.g. report and/or arrest data) and victim/survivor data (Nessel et al., 2020; Taylor et al., 2001). Police arrest data (where behaviour is deemed a criminal offence) as a measure on its own may exclude information on repeat perpetration/recidivism where behaviour, such as non-physical IPV, is not considered a criminal offence. Police reports may therefore provide a more useful measure of recidivism but are dependent on quality and depth of detail in police recording. Comparing recidivism data across trials from different countries can also create problems as arrest and criminal offences related to IPV behaviour may be defined and framed differently across countries. Additional sources for substantiating treatment effects can be particularly problematic due to generally low victim/survivor response rate (Alexander et al., 2010; Feder & Dugan, 2002; Labriola et al., 2005; Taylor & Maxwell, 2009). These observations echo findings from quasi-experimental research on the effects of IPV perpetrator programmes (Babcock & Steiner, 1999; Bloomfield & Dixon, n.d.; Coulter & VandeWeerd, 2009).

Generally, we noted investigators questioning the 'one size fits all' intervention approach (Dunford, 2000; Saunders, 1996). Treatment effects appear to be greater when interventions consider perpetrators' motivations for treatment attendance and adherence (Labriola et al., 2005; Lila et al., 2018; Taylor et al., 2001). Perpetrators are not a homogenous group and their typology (e.g., age, employment status, offence type, educational level, mental health status and/or substance abuse histories) is an important methodological consideration that can mitigate treatment effects (Labriola et al., 2005; Saunders, 1996) and restrict the generalisability of findings (Taft et al., 2016).

Our analysis indicates that participants, especially in court-mandated programme evaluations, need to be sufficiently integrated into a coordinated judicial and/or community response. The observation is not unique to the trialists considered in this review (Chen et al., 1989; Harrell, 1991). While some investigators question whether systemic factors are as important as participants' characteristics in influencing outcomes (Tollefson & Gross, 2006), the importance of agency collaboration, information sharing and the creation of strong links of the agencies with the local communities in which they operate cannot be underestimated (Kelly & Westmarland, 2015).

Implications for Research

RCTs of IPV perpetrator programmes are an important vehicle to increase the uptake and application of knowledge by clinical and policy decision-makers (Feder & Boruch, 2000). Our analysis highlighted consensus among trialists about the need to evaluate treatment effects through rigorous research designs, namely RCTs. While most quasi-experimental studies in the area acknowledge lack of randomisation as a threat to internal validity (Hendricks et al., 2006; Morrel et al., 2003; Tollefson & Gross, 2006), the conduct of RCTs is not without its difficulties.

Recent systematic reviews, partly because they employ different inclusion criteria, continue to provide conflicting findings relating to the effects of different interventions models (Arce et al., 2020; Karakurt et al., 2019). In particular, those examining the effects of the dominant cognitive-behavioral model (Nesset et al., 2019) note insufficient evidence to confirm its positive effects on men's IPV reduction. However, promising effects from some of the included RCTs are trauma-informed interventions (Karakurt et al., 2019; Taft et al., 2016), more integrated interventions such as psychological therapies in conjunction with substance abuse treatment (Tarzia et al., 2020) and CBT (Palmstierna et al., 2012). Nevertheless, the generalisability of findings remains limited, pointing to the need for more evidence to evaluate the effectiveness of group perpetrator interventions. Future studies should aim to test well-articulated and innovative treatment approaches (Romero-Martínez et al., 2022) specifically tailored to address the complex social, psychological and substance misuse needs of IPV perpetrators. When possible, investigators should consider how to circumvent difficulties arising from random allocation relating to a) ethical concerns (e.g., victim/survivor safety); b) differential retention of participants (due to group allocation) and c) keeping to a minimum judicial interference (in cases of perpetrators mandated to treatment) which can negatively influence the conduct of the study. With regard to ethical concerns it has, for example, been noted that attending an IPV programme may give some victim/survivors false hope about the future relationship, increasing their risk of harm from abuse (Feder & Dugan, 2002).

Furthermore, intervention attrition is an important barrier found in these programmes which potentially impacts the findings. A meta-analysis showed that dropout rates were similar regardless of the type of programme target audience, i.e. sex offenders or perpetrators of domestic abuse (Olver et al., 2011). The factors most associated with intervention attrition were age, criminal history, personality, psychological concerns, and motivation for treatment. Those who did not complete treatment were higher-risk offenders, and treatment abandonment predicted recidivism (Olver et al., 2011). The authors suggest that providers should build a strong therapeutic bond with clients and develop strategies to avoid intervention attrition. In this way, clients can be more motivated to participate in the programmes (McMurrin, 2003; Olver et al., 2011). However, there may be particular problems mixing low/moderate risk of harm and high harm perpetrators and Olver and colleagues' recommendation is probably too optimistic.

Other ideas for tackling intervention attrition are to separate out early non engagement and no shows at the first sessions from participants who engage initially in IPV perpetrator programmes but then drop out later. Richards et al. (2021) found that early no shows and non-engagers were more likely to have mental health problems whereas dropouts were more like to have substance use problems. Donovan and Griffiths' (2015) study focused on the pre-IPV programme phase and suggested that motivational interviewing should be used more in the early stages with a bigger role for social work practitioners and health practitioners supporting better initial engagement. Early screening of participants to identify factors that might make them more likely to not engage early on or drop out later may be helpful. In a meta-analysis of intervention attrition variables Jewell and Wormith (2010) recommend the early screening approach to facilitate the provision of better support having found younger, unemployed perpetrators who had not been court mandated to attend to be at greater risk of drop out. Another advocate for the use of screening, Priester et al. (2019) found that adverse childhood experiences, especially experiences of household dysfunction such as a loss of a parent, having a household member in prison or having substance use or mental health problems, predicted IPV perpetrator programme intervention attrition. Addressing the problems of later programme attrition, Richards et al. (2021) suggested that monitoring employment throughout an IPV perpetrator programme would be useful as this may be a red flag and precursor to dropping out. Other

suggestions for tackling attrition from Richards and colleagues are to invest in intensive case management and where probation staff are involved, for their role to be focused on longer term supporting goals such as programme completion and engagement, rather than just early signs of engagement.

Future trials should therefore consider pre-treatment screening (Dunford, 2000, Jewell & Wormith, 2010; Priester et al., 2019); careful monitoring of perpetrator behaviour during treatment and signposting men with previous abusive histories and severe psychological disorders to more intensive interventions (Alexander et al., 2010; Richards et al., 2021); and to identify and a priori match participants (in terms of motivation/needs) to interventions to improve attendance, engagement and outcomes (Alexander et al., 2010; Dunford, 2000; Jewell & Wormith, 2010; Saunders, 1996). As these are complex interventions with multiple factors affecting, for example, intervention engagement, attrition and outcomes, embedding routine and more comprehensive data collection and evaluation in all IPV perpetrator programmes would help support the ongoing improvement of understanding and practice. The standards against which the treatment is being evaluated would need to be explicitly defined (Davis et al., 2000) and both short (up to 6 months) and long term (preferably up to or longer than two year) follow-ups need to be included. Outcomes need to be assessed, preferably by independent assessors and reliability scores provided when independent raters are used (Easton et al., 2007). Careful consideration and planning are required to overcome low victim/survivor response rate during these evaluations. All outcome measures are open to different forms of bias. Perpetrator accounts might be unreliable and too invested in their own change narrative (Dobash & Dobash, 1998; Henning & Holdford, 2006); partners might be reluctant to report abuse because, amongst other reasons, of fear of retaliation, shame, or concern about professional intervention (Boethius & Åkerström, 2020; Overstreet & Quinn, 2013); and police are more likely to record and make arrests in relation to incidents of physical abuse than coercively controlling behaviour (McPhee et al., 2022). We do not think that a combined outcome reduces the bias in any of these measures or that alternative outcome measures are preferable, such as risk scores (Turner et al., 2019). We recommend reporting all three outcomes as they are of potential value to different audiences such as people experiencing abuse, service commissioners and practitioners (Westmarland & Kelly, 2013).

The current review aimed to identify and collate all published RCTs on group IPV male perpetrator programmes. As primary studies were identified through systematic reviews, due to their various inclusion/exclusion criteria, it is possible we may have missed some primary studies. We might not have exhaustively covered all possible database search terms and we did not, for example, use the term "perpetrator" although we did use the terms "batterer" and "partner-violent men". To maximize retrieval of systematic reviews and meta-analyses, our search utilised multiple complementary methods in addition to database searching. We searched broadly using eight electronic databases. Previous systematic reviews and/or meta-analyses used between two and six databases except for one article that used 13 databases to assess cultural differences in IPV interventions (Satyen et al., 2022). Although our review used multiple databases it does not fulfil the search criteria for a systematic review (Uman, 2011). This does not weaken the validity of our methodological meta-summary based on the systematic reviews (and their primary trials) that we identified in our search. We then used a rigorous process for identifying, extracting, and analysing statements referring to methodological challenges from the discussion section of these documents. Importantly, we were interested in any reference to methodological challenges from trialists assessing the effectiveness of IPV male perpetrator interventions. It is important to note, however, that the methodological challenges summarised here were extracted from the Discussion section of the reports only, it is possible

that additional reference to methodological challenges may have been identified should we have included the whole report (Gondolf, 2015). Our review contributes by pointing out the limitations and methodological challenges of previous randomised controlled trials, which we hope will be helpful to the development of more rigorous study designs and further development of feasibility and effectiveness studies of IPV interventions.

Conclusion

This review provides a systematic summary of existing methodological challenges, as identified by trialists, in the conduct of randomised controlled trials on IPV male perpetrator programmes. The present review will serve as a useful resource for authors wishing to conduct IPV group perpetrator programmes, as well as researchers wishing to conduct empirical research on IPV group programme effectiveness. It is also a necessary first step to developing a cohesive methods guidance document that addresses relevant issues and areas of uncertainty when planning studies on the effects of IPV group perpetrator interventions. Accordingly, the results of this study were used to inform the design of a randomised controlled trial of a community-based intervention for male perpetrators of mild to moderate risk in the Southwest of England and South Wales (<http://www.bristol.ac.uk/primaryhealthcare/researchthemes/reprovide/>). The results of this REPROVIDE study are due at the end of 2024.

Acknowledgements

The authors would like to express their sincere thanks to Francisco Diego Rabelo da Ponte for all his help with this paper. We would also like to thank the editors for their support and anonymous reviewers for their comments and suggestions.

Conflict of Interest

The authors of this article declare no conflict of interest.

References

- References marked with an asterisk* indicate studies included in the meta-summary
- Akoensi, T. D., Koehler, J. A., Lösel, F., & Humphreys, D. K. (2013). Domestic violence perpetrator programs in Europe, Part II. *International Journal of Offender Therapy and Comparative Criminology*, 57(10), 1206–1225. <https://doi.org/10.1177/0306624X12468110>
- *Alexander, P. C., Morris, E., Tracy, A., & Frye, A. (2010). Stages of change and the group treatment of batterers: A randomized clinical trial. *Violence and Victims*, 25(5), 571–587. <https://doi.org/10.1891/0886-6708.25.5.571>
- Aos, S., Miller, M. G., & Drake, E. (2006). *Evidence-based adult corrections programs: What works and what does not*. Washington State Institute for Public Policy. https://www.wsipp.wa.gov/ReportFile/924/Wsipp_Evidence-Based-Adult-Corrections-Programs-What-Works-and-What-Does-Not_Preliminary-Report.pdf
- Arce, R., Arias, E., Novo, M., & Fariña, F. (2020). Are interventions with batterers effective? A meta-analytical review. *Psychosocial Intervention*, 29(3), 153–164. <https://doi.org/10.5093/pi2020a11>
- Arias, E., Arce, R., & Vilariño, M. (2013). Batterer intervention programmes: A meta-analytic review of effectiveness. *Psychosocial Intervention*, 22(2), 153–160. <https://doi.org/10.5093/in2013a18>
- 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(8), 1023–1053. <https://doi.org/10.1016/j.cpr.2002.07.001>
- 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(1), 46–59. <https://doi.org/10.1037/0893-3200.13.1.46>
- Baker, D. E., Hill, M., Chamberlain, K., Hurd, L., Karlsson, M., Zielinski, M., Calvert, M., & Bridges, A. J. (2021). Interpersonal vs. non-interpersonal cumulative traumas and psychiatric symptoms in treatment-seeking incarcerated women. *Journal of Trauma & Dissociation: The Official Journal of the International Society for the Study of Dissociation*, 22(3), 249–264. <https://doi.org/10.1080/15299732.2020.1760172>
- Barner, J. R., & Carney, M. M. (2011). Interventions for intimate partner violence: A historical review. *Journal of Family Violence*, 26(3), 235–244. <https://doi.org/10.1007/s10896-011-9359-3>
- Bloomfield, S., & Dixon, L. (n.d.). *An outcome evaluation of the Integrated Domestic Abuse Programme (IDAP) and Community Domestic Violence Programme (CDVP)*. Retrieved March 11, 2023, from <https://www.bl.uk/britishlibrary/~media/bl/global/social-welfare/pdfs/non-secure/o/u/t/outcome-evaluation-of-the-integrated-domestic-abuse-programme-idap-and-community-domestic-violence-programme-cdvp.pdf>
- Boethius, S., & Åkerström, M. (2020). Revealing hidden realities: Disclosing domestic abuse to informal others. *Nordic Journal of Criminology*, 21(2), 186–202. <https://doi.org/10.1080/2578983X.2020.1787725>
- Campbell, J. C., Anderson, J. C., McFadgion, A., Gill, J., Zink, E., Patch, M., Callwood, G., & Campbell, D. (2018). The effects of intimate partner violence and probable traumatic brain injury on central nervous system symptoms. *Journal of Women's Health*, 27(6), 761–767. <https://doi.org/10.1089/jwh.2016.6311>
- Cheng, S.-Y., Davis, M., Jonson-Reid, M., & Yaeger, L. (2021). Compared to what? A meta-analysis of batterer intervention studies using nontreated controls or comparisons. *Trauma, Violence & Abuse*, 22(3), 496–511. <https://doi.org/10.1177/1524838019865927>
- Chen, H.-T., 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>
- Cluss, P., & Bodea, A. (2011). The effectiveness of batterer intervention programs. *A Literature Review & Recommendations for next Steps*. University of Pittsburg. <https://southwestpasaysnomore.org/wp-content/uploads/2015/03/BatterersInterventionProgramEffectivenessReport-without-articles.pdf>
- 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>
- Daugherty, J. C., Pérez-García, M., Hidalgo-Ruzzante, N., & Bueso-Izquierdo, N. (2021). Perceived executive functioning among female survivors of intimate partner violence. *Journal of Aggression, Maltreatment & Trauma*, 30(1), 25–42. <https://doi.org/10.1080/10926771.2020.1783734>
- Davis, R. C., Victim Services Research, Taylor, B. G., & Maxwell, C. D. (2000). Does batterer treatment reduce violence? A randomized experiment in Brooklyn - executive summary included. *PsycEXTRA Dataset*. <https://doi.org/10.1037/e300332003-001>
- Dobash, R. E., & Dobash, R. (1979). *Violence against wives: A case against the patriarchy*. Free Press.
- Dobash, R. E., & Dobash, R. P. (1998). Violent men and violent contexts. In *Rethinking violence against women*, (pp. 141–168). SAGE Publications. <https://doi.org/10.4135/9781452243306.n6>
- Donovan, C., & Griffiths, S. (2015). Domestic violence and voluntary perpetrator programmes: Engaging men in the pre-commencement phase. *British Journal of Social Work*, 45(4), 1155–1171. <https://doi.org/10.1093/bjsw/bct182>
- *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>
- *Easton, C. J., Mandel, D. L., Hunkele, K. A., Nich, C., Rounsaville, B. J., & Carroll, K. M. (2007). A cognitive behavioral therapy for alcohol-dependent domestic violence offenders: An integrated Substance Abuse–Domestic Violence Treatment Approach (SADV). *The American Journal on Addictions / American Academy of Psychiatrists in Alcoholism and Addictions*, 16(1), 24–31. <https://doi.org/10.1080/10550490601077809>
- 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(2), 196–231. <https://doi.org/10.1891/1946-6560.4.2.196>
- Feder, G., Hester, M., Williamson, E., & Dunne, E. (2008). Behavioral interventions to reduce intimate partner violence against women. In J. A. Trafton & W. P. Gordon (Eds.), *Best practices in the behavioural management of health from preconception to adolescence*, (Vol. 3, pp. 522–567). Institute for Disease Management/Institute for Brain Potential.
- Feder, L., & Boruch, R. F. (2000). The need for experiments in criminal justice settings. *Crime & Delinquency*, 46(3), 291–294. <https://doi.org/10.1177/001128700046003001>
- *Feder, L., & Dugan, L. (2002). A test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward experiment. *Justice Quarterly*, 19(2), 343–375. <https://doi.org/10.1080/0741882020095271>
- Feder, L., & Forde, D. R. (2000). A test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward experiment, executive summary. *PsycEXTRA Dataset*. <https://doi.org/10.1037/e526812006-001>

- Feder, L., & Wilson, D. B. (2005). A meta-analytic review of court-mandated batterer intervention programs: Can courts affect abusers' behavior? *Journal of Experimental Criminology*, 1(2), 239–262. <https://doi.org/10.1007/s11292-005-1179-0>
- Ferrer-Perez, V. A., & Bosch-Fiol, E. (2018). 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>
- Ford, D. A., & Regoli, M. J. (1993). The criminal prosecution of wife assaulters: Process, problems, and effects. In N. Z. Hilton (Ed.), *Legal responses to wife assault: Current trends and evaluation* (pp. 127–164). Sage.
- Gibson, C. J., Huang, A. J., McCaw, B., Subak, L. L., Thom, D. H., & Van Den Eeden, S. K. (2019). Associations of intimate partner violence, sexual assault, and posttraumatic stress disorder with menopause symptoms among midlife and older women. *Journal of the American Medical Association Internal Medicine*, 179(1), 80–87. <https://doi.org/10.1001/jamainternmed.2018.5233>
- Gondolf, E. W. (2011). The weak evidence for batterer program alternatives. *Aggression and Violent Behavior*, 16(4), 347–353. <https://doi.org/10.1016/j.avb.2011.04.011>
- Gondolf, E. W. (2015). *Gender-based perspectives on batterer programs: program leaders on history, approach, research, and development*. Lexington Books.
- Hamberger, L. K., & Hastings, J. E. (1993). Court-mandated treatment of men who assault their partner: Issues, controversies, and outcomes. In N. Z. Hilton (Ed.), *Legal responses to wife assault* (pp. 96–121). Sage.
- Harrell, A. (1991). *Evaluation of court-ordered treatment for domestic violence offenders: Final report*. Urban Institute.
- Healey, K., Smith, C., & O'Sullivan, C. S. (1999). *Batterer intervention: Program approaches and criminal justice strategies*. DIANE Publishing.
- Hendricks, B., Werner, T., Shipway, L., & Turinetti, G. J. (2006). Recidivism among spousal abusers: Predictions and program evaluation. *Journal of Interpersonal Violence*, 21(6), 703–716. <https://doi.org/10.1177/0886260506287310>
- Henning, K., & Holdford, R. (2006). Minimization, denial, and victim blaming by batterers. *Criminal Justice and Behavior*, 33(1), 110–130. <https://doi.org/10.1177/0093854805282322>
- Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savovic, J., Schulz, K. F., Weeks, L., Sterne, J. A. C., Cochrane Bias Methods Group, & Cochrane Statistical Methods Group. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *British Medical Journal*, 343, Article d5928. <https://doi.org/10.1136/bmj.d5928>
- Hamel, J. L., & Nicholls, T. L. (2006). *Family interventions in domestic violence: A handbook of gender-inclusive theory and treatment*. Springer Publishing Company.
- Jewell, L. M., & Wormith, J. S. (2010). Variables associated with attrition from domestic violence treatment programs targeting male batterers: A meta-analysis. *Criminal Justice and Behavior*, 37(10), 1086–1113. <https://doi.org/10.1177/0093854810376815>
- Karakurt, G., Koç, E., Çetinsaya, E. E., Ayluçtarhan, Z., & Bolen, S. (2019). Meta-analysis and systematic review for the treatment of perpetrators of intimate partner violence. *Neuroscience and Biobehavioral Reviews*, 105, 220–230. <https://doi.org/10.1016/j.neubiorev.2019.08.006>
- Kelly, L., & Westmarland, N. (2015). *Domestic violence perpetrator programmes: Steps towards change. Project Mirabal final report*. <https://repository.londonmet.ac.uk/1458/1/ProjectMirabalfinalreport.pdf>
- Krishna, M., Jauhari, A., Lepping, P., Turner, J., Crossley, D., & Krishnamoorthy, A. (2011). Is group psychotherapy effective in older adults with depression? A systematic review. *International Journal of Geriatric Psychiatry*, 26(4), 331–340. <https://doi.org/10.1002/gps.2546>
- *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*. Center for Court Innovation.
- Lee, R., Kavoussi, R. J., & Coccaro, E. F. (2008). Placebo-controlled, randomized trial of fluoxetine in the treatment of aggression in male intimate partner abusers. *International Clinical Psychopharmacology*, 23(6), 337–341. <https://doi.org/10.1097/YIC.0b013e32830fbbd2>
- *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>
- Lilley-Walker, S. J., Hester, M., & Turner, W. (2018). Evaluation of European domestic violence perpetrator programmes: Toward a model for designing and reporting evaluations related to perpetrator treatment interventions. *International Journal of Offender Therapy and Comparative Criminology*, 62(4), 868–884. <https://doi.org/10.1177/0306624x16673853>
- Maxwell, C. D., Davis, R. C., & Taylor, B. G. (2010). The impact of length of domestic violence treatment on the patterns of subsequent intimate partner violence. *Journal of Experimental Criminology*, 6, 475–497. <https://doi.org/10.1007/s11292-010-9106-4>
- McMurran, M. (2003). *Motivating offenders to change: A guide to enhancing engagement in therapy*. John Wiley & Sons.
- McPhee, D., Hester, M., Bates, L., Lilley-Walker, S. J., & Patsios, D. (2022). Criminal justice responses to domestic violence and abuse in England: An analysis of case attrition and inequalities using police data. *Policing and Society*, 32(8), 963–980. <https://doi.org/10.1080/10439463.2021.2003358>
- Miller, M. G., Drake, E., & Nafziger, M. (2013). *What works to reduce recidivism by domestic violence offenders?* (Document No. 13-01-1201). Washington State Institute for Public Policy.
- Morrel, T. M., Elliott, J. D., Murphy, C. M., & Taft, C. T. (2003). Cognitive behavioral and supportive group treatments for partner-violent men. *Behavior Therapy*, 34(1), 77–95. [https://doi.org/10.1016/S0005-7894\(03\)80023-0](https://doi.org/10.1016/S0005-7894(03)80023-0)
- *Murphy, C. M., Eckhardt, C. I., Clifford, J. M., LaMotte, A. D., & Meis, L. A. (2020). Individual versus group cognitive-behavioral therapy for partner-violent men: A preliminary randomized trial. *Journal of Interpersonal Violence*, 35(15–16), 2846–2868. <https://doi.org/10.1177/0886260517705666>
- National Research Council. (1996). *Understanding violence against women*. Division of Behavioral and Social Sciences and Education, Commission on Behavioral and Social Sciences and Education, & Panel on Research on Violence Against Women. National Academy Press.
- *Nesset, M. B., Lara-Cabrera, M. L., Bjørngaard, J. H., Whittington, R., & Palmstierna, T. (2020). Cognitive behavioural group therapy versus mindfulness-based stress reduction group therapy for intimate partner violence: A randomized controlled trial. *BMC Psychiatry*, 20, Article 178. <https://doi.org/10.1186/s12888-020-02582-4>
- Nesset, M. B., Lara-Cabrera, M. L., Dalsbø, T. K., Pedersen, S. A., Bjørngaard, J. H., & Palmstierna, T. (2019). Cognitive behavioural group therapy for male perpetrators of intimate partner violence: A systematic review. *BMC Psychiatry*, 19, Article 11. <https://doi.org/10.1186/s12888-019-2010-1>
- National Institute for Health and Care Excellence, British Columbia Centre of Excellence for Women's Health. (2013). Review of interventions to identify, prevent, reduce and respond to domestic violence. National Institute for Health and Care Excellence. <https://www.nice.org.uk/guidance/ph50/resources/review-of-interventions-to-identify-prevent-reduce-and-respond-to-domestic-violence2>
- O'Connor, A., Morris, H., Panayiotidis, A., Cooke, V., & Skouteris, H. (2021). Rapid review of men's behavior change programs. *Trauma, Violence & Abuse*, 22(5), 1068–1085. <https://doi.org/10.1177/1524838020906527>
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2011). A meta-analysis of predictors of offender treatment attrition and its relationship to recidivism. *Journal of Consulting and Clinical Psychology*, 79(1), 6–21. <https://doi.org/10.1037/a0022200>
- Overstreet, N. M., & Quinn, D. M. (2013). The intimate partner violence stigmatization model and barriers to help-seeking. *Basic and Applied Social Psychology*, 35(1), 109–122. <https://doi.org/10.1080/01973533.2012.746599>
- *Palmer, S. E., Brown, R. A., & Barrera, M. E. (1992). Group treatment program for abusive husbands: Long-term evaluation. *The American Journal of Orthopsychiatry*, 62(2), 276–283. <https://doi.org/10.1037/h0079336>
- *Palmstierna, T., Haugan, G., Jarwson, S., Rasmussen, K., & Nøttestad, J. A. (2012). Cognitive-behaviour group therapy for men voluntarily seeking help for intimate partner violence. *Nordic Journal of Psychiatry*, 66(5), 360–365. <https://doi.org/10.3109/08039488.2012.665080>
- Pence, E., Paymar, M., & Ritmeester, T. (1993). *Education groups for men who batter: The Duluth model*. Springer Publishing Company.
- Peterson, C., Kearns, M. C., McIntosh, W. L., Estefan, L. F., Nicolaidis, C., McCollister, K. E., Gordon, A., & Florence, C. (2018). Lifetime economic burden of intimate partner violence among U.S. Adults. *American Journal of Preventive Medicine*, 55(4), 433–444. <https://doi.org/10.1016/j.amepre.2018.04.049>
- Phillips, R., Kelly, L., & Westmarland, N. (2013). *Domestic violence perpetrator programmes: An historical overview*. London Metropolitan University and Durham University. <https://www.dur.ac.uk/resources/criva/anhistoricaloverviewbriefingnote.pdf>
- Pinto E Silva, T., Cunha, O., & Caridade, S. (2022). Motivational interview techniques and the effectiveness of intervention programs with perpetrators of intimate partner violence: A systematic review. *Trauma, Violence & Abuse*, 0(0), 1–20. <https://doi.org/10.1177/1524838022111472>
- Postmus, J. L., Stylianou, A. M., & McMahon, S. (2016). The abusive behavior inventory—revised. *Journal of Interpersonal Violence*, 31(17), 2867–2888. <https://doi.org/10.1177/0886260515581882>
- Priester, M. A., Kulkarni, S., Mennicke, A., & Bell, B. A. (2019). Factors associated with batterer intervention program attrition. *Violence and Victims*, 34(2), 296–311. <https://doi.org/10.1891/0886-6708.VV-D-17-00067>
- Richards, T. N., Jennings, W. G., & Murphy, C. (2021). Risk and protective factors for batterer intervention treatment program attrition: How completers are distinct from dropouts and no-shows. *Journal of Interpersonal Violence*, 36(15–16), 7351–7370. <https://doi.org/10.1177/0886260519834096>
- Romero-Martínez, Á., Santirso, F., Lila, M., Comes-Fayos, J., & Moya-Albiol, L. (2022). Cognitive flexibility and reaction time improvements after cognitive training designed for men perpetrators of intimate partner violence: Results of a pilot randomized controlled trial. *Journal of Family Violence*, 37(3), 461–473. <https://doi.org/10.1007/s10896-021-00304-2>
- Sandelowski, M., & Barroso, J. (2003). Creating metasummaries of qualitative findings. *Nursing Research*, 52(4), 226–233. <https://doi.org/10.1097/00006199-200307000-00004>
- Santirso, F. A., Gilchrist, G., Lila, M., & Gracia, E. (2020). Motivational strategies in interventions for intimate partner violence offenders: A

- systematic review and meta-analysis of randomized controlled trials. *Psychosocial Intervention*, 29(3), 175–190. <https://doi.org/10.5093/pi2020a13>
- Sardinha, L., Maheu-Giroux, M., Stöckl, H., Meyer, S. R., & García-Moreno, C. (2022). Global, regional, and national prevalence estimates of physical or sexual, or both, intimate partner violence against women in 2018. *The Lancet*, 399(10327), 803–813. [https://doi.org/10.1016/s0140-6736\(21\)02664-7](https://doi.org/10.1016/s0140-6736(21)02664-7)
- Satyen, L., Hansen, A., Green, J. L., & Zark, L. (2022). The effectiveness of culturally specific male domestic violence offender intervention programs on behavior changes and mental health: A systematic review. *International Journal of Environmental Research and Public Health*, 19(22), Article 1580. <https://doi.org/10.3390/ijerph192215180>
- *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>
- Shepard, M. F., & Campbell, J. A. (1992). The Abusive Behavior Inventory: A measure of psychological and physical abuse. *Journal of Interpersonal Violence*, 7(3), 291–305. <https://doi.org/10.1177/088626092007003001>
- Smedslund, G., Dalsbø, T. K., Steiro, A. K., Winsvold, A., & Clench-Aas, J. (2011). Cognitive behavioural therapy for men who physically abuse their female partner. *Cochrane Database of Systematic Reviews*, 2, Article CD006048. <https://doi.org/10.1002/14651858.CD006048.pub2>
- Stephens-Lewis, D., Johnson, A., Huntley, A., Gilchrist, E., McMurrin, M., Henderson, J., Feder, G., Howard, L. M., & Gilchrist, G. (2021). Interventions to reduce intimate partner violence perpetration by men who use substances: A systematic review and meta-analysis of efficacy. *Trauma, Violence & Abuse*, 22(5), 1262–1278. <https://doi.org/10.1177/1524838019882357>
- *Taft, C. T., Macdonald, A., Creech, S. K., Monson, C. M., & Murphy, C. M. (2016). A randomized controlled clinical trial of the strength at home men's program for partner violence in military veterans. *The Journal of Clinical Psychiatry*, 77(9), 1168–1175. <https://doi.org/10.4088/JCP.15m10020>
- Tarzia, L., Forsdike, K., Feder, G., Hegarty, K. (2020) Interventions in health settings for male perpetrators or victims of intimate partner violence. *Trauma Violence Abuse*, 21(1), 123–137. <https://doi.org/10.1177/1524838017744772>
- *Taylor, B. G., Davis, R. C., & Maxwell, C. D. (2001). The effects of a group batterer treatment program: A randomized experiment in Brooklyn. *Justice Quarterly: JQ/Academy of Criminal Justice Sciences*, 18(1), 171–201. <https://doi.org/10.1080/07418820100094861>
- *Taylor, B. G., & Maxwell, C. (2009). *The effects of a short-term batterer treatment program for detained arrestees: A randomized experiment in the Sacramento County, California jail*. Department of Justice, Crime and Prevention Center.
- 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>
- Travers, A., McDonagh, T., Cunningham, T., Armour, C., & Hansen, M. (2021). The effectiveness of interventions to prevent recidivism in perpetrators of intimate partner violence: A systematic review and meta-analysis. *Clinical Psychology Review*, 84, Article 101974. <https://doi.org/10.1016/j.cpr.2021.101974>
- Tsantefski, M., Young, A., Wilde, T., & O'Leary, P. (2021). High-risk cases at the intersection of domestic/family violence and child protection: Learning from practice. *Journal of Family Violence*, 36(8), 941–952. <https://doi.org/10.1007/s10896-021-00255-8>
- Turner, E., Medina, J., & Brown, G. (2019). Dashing hopes? The predictive accuracy of domestic abuse risk assessment by police. *The British Journal of Criminology*, 59(5), 1013–1034. <https://doi.org/10.1093/bjc/azy074>
- Uman, L.S. (2011). Systematic reviews and meta-analyses. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 20(1), 57–59.
- Vigurs, C. A., Schucan Bird, K., Quy, K., & Gough, D. (2016). *The impact of domestic violence perpetrator programmes on victim and criminal justice outcomes: A systematic review of reviews of research evidence*. College of Policing.
- Westmarland, N., & Kelly, L. (2013). Why extending measurements of “success” in domestic violence perpetrator programmes matters for social work. *British Journal of Social Work*, 43(6), 1092–1110. <https://doi.org/10.1093/bjsw/bcs049>
- Wilson, D. B., Feder, L., & Olaghere, A. (2021). Court-mandated interventions for individuals convicted of domestic violence: An updated Campbell systematic review. *Campbell Systematic Reviews*, 17(1), Article e1151. <https://doi.org/10.1002/cl2.1151>
- World Health Organization. (2013). *Responding to intimate partner violence and sexual violence against women: WHO clinical and policy guidelines*.
- World Health Organization. (2021). *Violence against women prevalence estimates, 2018: Global, regional and national prevalence estimates for intimate partner violence against women and global and regional prevalence estimates for non-partner sexual violence against women*.
- Yalom, I. D., & Leszcz, M. (2020). *The theory and practice of group psychotherapy*. Hachette UK.

