

BMJ Open

Searching for the mechanisms of change: A protocol for a realist review of batterer treatment programs

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2015-010173
Article Type:	Protocol
Date Submitted by the Author:	02-Oct-2015
Complete List of Authors:	Velonis, Alisa; St. Michael's Hospital, Centre for Research on Inner City Health; University of Illinois at Chicago, School of Public Health, Community Health Sciences Cheff, Rebecca; St. Michael's Hospital, Centre for Research on Inner City Health Finn, Debbie; St. Michael's Hospital, Centre for Research on Inner City Health Davloor, Whitney; St. Michael's Hospital, Centre for Research on Inner City Health O'Campo, Patricia; St. Michael's Hospital, Centre for Research on Inner City Health; University of Toronto, Dalla Lana School of Public Health
Primary Subject Heading:	Public health
Secondary Subject Heading:	Research methods, Mental health
Keywords:	domestic violence treatment, realist review, evidence-based practices

SCHOLARONE™
Manuscripts

Only

**SEARCHING FOR THE MECHANISMS OF CHANGE: A PROTOCOL FOR A REALIST
REVIEW OF BATTERER TREATMENT PROGRAMS**

Alisa J. Velonis, MPH, PhD; ^{1,2}
Rebecca Cheff, MPH; ¹
Debbie Finn, RN BScN MN; ¹
Whitney Davloor, MPP; ¹
Patricia O’Campo, PhD^{1,3}

Corresponding Author:
Alisa J. Velonis, MPH, PhD
University of Illinois at Chicago
School of Public Health, Division of Community Health Sciences
1603 W. Taylor St
Chicago, IL 60612 U.S.A.
avelonis@uic.edu
312-355-1217

Institutional Affiliations:
¹ St. Michael’s Hospital, Centre for Research on Inner City Health, Toronto, Ontario, Canada
² University of Illinois at Chicago, School of Public Health, Division of Community Health Sciences, Chicago, Illinois, U.S.A.
³ University of Toronto, Dalla Lana School of Public Health, Toronto, Ontario, Canada

Key Words:
Domestic Violence Treatment, Batterer Treatment, Realist Review Protocol,

Word Count: 3,984

ABSTRACT

Introduction: While some research suggests that the most common community-based psycho-educational approaches to reducing the violent behaviour used by perpetrators of intimate partner violence to have little or no impact, others report contrary findings, leaving judicial officials and policy makers in search of evidence based solutions uncertain about recommending and funding these programs.

Traditional approaches to program evaluation and systematic reviews of batterer intervention tend to focus predominately on whether the programs “worked” (e.g. reduced recidivism) often at the exclusion of understanding for whom they may or may not have worked, under what circumstances, and why.

Methods and Analysis: We are undertaking a realist review of the batterer treatment program literature with the aim of addressing this gap. Keeping with the goals of realist review, our primary aims are to identify the theory that underlies these programs, highlight the mechanisms that would lead participants to reducing their violence, and finally explain why these programs help some individuals reduce their use of violence and under what conditions they are effective or not effective. We begin by describing the process of perpetrator treatment and proposing a possible theoretical model of behaviour change that will be tested by our review. We then describe the criteria for inclusion of an evaluation into the review, the search strategy we will use to identify the studies, and the plan for data extraction and analysis.

Ethics and Dissemination. The results of this review will be written up using the RAMESES Guidelines for Realist Synthesis and disseminated through peer-reviewed

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

publications aimed at the practitioner community as well as presented at
community forums and at violence against women conferences. Ethics approval
was not needed.

For peer review only

INTRODUCTION/BACKGROUND

Since the emergence of early batterer treatment or intervention programs (BIPs) in the 1970s and 1980s, discussions about their efficacy have proliferated within research, professional, and policy circles. In attempting to answer the general question, “Do these programs work?” a number of sub-surface debates have emerged, highlighting points of contention about the nature of intimate partner violence (IPV), about the multiple levels and types of influences that may contribute to abusive behavior, about what “success” means in terms of batterer treatment, and about the “right” approach to both achieving and evaluating this success. These differences in perspective lead to conflicting conclusions about the effectiveness of these programs, making the job of navigating the literature that surrounds batterer intervention or treatment programs challenging.

Yet, as part of the official response to domestic violence across North America, judicial statutes frequently require individuals who are convicted of crimes against intimate partners to attend treatment or educational programs (the content and format of which can vary widely) as a condition to receiving a deferred sentence, probation, or parole. This, among other factors, has spurred a proliferation of program evaluations and systematic reviews, most sharing a defined goal of determining whether or not the programs that currently exist can be proven to directly reduce subsequent violence and criminal behaviors. While there has been significant debate over what theoretical approach(s) should be used to guide these programs (e.g. feminist theory, family systems theory, cognitive behavioral theory), the nature of these discussions tends to be as political (e.g., pro- or anti-

feminist in rhetoric) as it is scientific ¹⁻⁴. Few systematic reviews have attempted to examine the underlying programmatic theory and understand how, why, and in what contexts these programs work – or do not work.

Focusing on Theory Rather Than on Programs

Drawing upon the work of Sayer ⁵ and other realist philosophers, Pawson ⁶ describes interventions as “complex process that are inserted into complex structures” (pg 79). Anyone who has tried to implement community development or health promotion programs recognizes the truth in this statement: budgets get cut, referrals increase or decrease, participants resent being there, staff feel overworked and underappreciated, and the list goes on. Yet even under the best of conditions, behavior change is difficult, and programs often use multiple theoretical and methodological approaches to help clients move along the path to improvement.

This is certainly true of batterer intervention efforts. Regardless of the therapeutic, philosophical, or political framework of the program, when implemented in community settings (i.e., outside of self-contained institutions such as the military or correctional facilities), programs aimed at curbing batterers’ violence are impacted by a variety of factors, including the characteristics and experiences of participants and staff, the mission of the lead organization, and the social and political climate of the larger community. In turn, these factors influence how programs run, how closely they adhere to the program design, how the strategies are received by participants, and more.

Yet, a great deal of the research describing the “effectiveness” of batterer intervention programs has been designed to minimize the influence of these real-world contextual factors, generally by controlling for many of the very forces that could explain program success or failure (e.g., cultural backgrounds, income, substance use). Often, experimental and quasi-experimental evaluations are held up as the gold standard; these models endeavor to link the intervention – and ONLY the intervention – to a narrowly defined outcome, usually recidivism or re-offense⁷⁻¹⁰. A result has been the proliferation of a vast body of literature that shows mixed evidence of program success with little explanation of why. It was our frustration with this lack of explanation (and the frustration that program administrators and domestic violence advocates expressed to us about not knowing what to do to improve treatment) that led us to conduct this realist synthesis of the literature.

What follows is a description of a protocol we have developed for undertaking this review. As with other realist reviews, the purpose of this synthesis is explanatory: to articulate underlying program theories and use evidence to determine their usefulness and relevance for batterer treatment⁶. In writing and sharing this protocol, we set forth three goals. First, to clearly delineate the theory or theories that provide the framework for the most common batterer intervention programs with the aim of identifying both strengths and gaps; second, to illustrate how this approach to understanding and evaluating programs can add value by shifting focus from whether programs work to how and why they [should] work; and three, to provide transparency to the forthcoming review, enabling readers to

know the specific steps that will be followed throughout the review process to ensure both rigor and scientific validity.

This protocol paper is organized into several sections, following the steps of realist synthesis laid out by Pawson and others ¹¹ (Figure 1). The core of any realist synthesis or evaluation is the description of the programmatic theory that underpins the activity being evaluated (Step 1). As will become evident, we started this process in conjunction with the development of this protocol, laying out what we identified as the relevant middle-range theory that forms the basis for the most common approaches to perpetrator treatment; this is likely to evolve and be refined over the course of the review process. Next, we describe the remaining steps (2 – 5) that will be taken between before the end of 2015 to conduct this review. Finally, we conclude with a discussion of why we believe this approach will contribute to our understanding of batterer treatment programs and how it can inform not only program implementation but also larger policy.

FIGURE 1 about here

METHODS/DESIGN

Step 1a. Establishing the Scope of our Work

The impetus for this review was a series of conversations with program facilitators and judicial personnel who had grown frustrated with the lack of conclusive information about what they could do, from their respective positions, to stop the ongoing perpetration of partner violence. As they saw it, the evidence was

insufficient to declare batterer interventions useless, yet these programs clearly did not work in all circumstances. Using this to inform our approach, we framed our research question as *for whom, and under what conditions, will batterer intervention programs help men who have been identified as perpetrators of partner violence reduce their violent behaviors and why?*

Before refining our scope further and articulating the theories that we believed to be at work, we needed to understand what the bigger picture of batterer response looked like in North America. After selecting a handful of evaluations, systematic reviews, and reports from both the scientific and grey literature a sample of these includes: ^{7 8 10 12 13-19}, we learned that the majority of individuals who attend batterer intervention in the U.S. and Canada undergo a two-part process: first, they enter the criminal justice system and are adjudicated for an offense against a partner, and second, they are mandated to attend an educational/therapeutic “treatment” program – what we refer to as a BIP – as part of their sentence or agreement with the court. Because the vast majority of participants in BIPs first have contact with the criminal justice system (and because our interest is in explaining how and why these programs work), we have chosen to view this as part of the *context* that surrounds the BIP rather than as a separate batterer intervention in and of itself.

In the language of realist synthesis, the key elements of programmatic theory that need to be considered include the *context* in which programs operate, the *outcomes* or goals of the program, and the *mechanisms* that trigger the outcome ^{11 20}. After reviewing program descriptions of BIPs, we differentiated two sets of

outcomes that were generally discussed: proximal outcomes, or the changes that happen within the participants as a result of the intervention (attitudes, skills, and intentions), and distal outcomes, or the ways in which those initial changes manifest in terms of recidivism and re-assault (see Figure 2).

Figure 2 About Here

The majority of batterer intervention evaluations are concerned primarily or exclusively with distal outcomes, especially recidivism and/or re-assault, which we found problematic for several reasons. First, the definition and measurement of *recidivism* differs from study to study, and often is not limited to assaults against a partner but can include a conviction for any violent crime, an arrest for a violent crime (regardless of conviction) or even any subsequent arrest. Secondly, because many – if not most – acts of abuse do not result in law enforcement intervention, these studies may underestimate the reoccurrence of these behaviours. However, other measures of re-assault are equally unreliable, in that unless evaluators have been able to contact the perpetrators’ current partners, these data rely on participants’ self-reported descriptions of their behaviours. Even when victim reports are included, the numbers are often small and subject to self-report bias.

More important, however, is the recognition that numerous factors external to both the BIP and the criminal justice system likely influence recidivism or re-assault rates, including factors related to the perpetrator and his family as well as those found in the larger community and political context. While these influences

may also affect the proximal outcomes (attitudes, skills, and intentions) or whether the perpetrator enters or completes the program, we decided that because the distal outcomes are at least one step removed from the BIP itself, it would be especially difficult to determine the role that these contextual factors play. For example, does substance use (a contextual factor) prevent clients from attaining the proximal outcomes, which leads to a failure to reduce subsequent violence, or does it make it less likely that improvement in attitudes, skills, and intentions will lead to a reduction of violence? For these reasons, we chose to limit our review to the relationship between BIPs and these proximal outcomes, which are addressed in more detail below.

Step 1b. Developing Our Hypothesized Theories of Change: CMOs

A key step in realist synthesis is the development of a working hypothesis describing the theory or theories that underpin the intervention: *why* should certain program activities (or strategies) result in the desired outcomes? What is it that happens or shifts within the participants as a result of those activities that cause those outcomes to occur? In other words, *what are the mechanisms that lead to change?* The process of identifying these candidate theories tends to be iterative²¹, and we reviewed formal and informal descriptions of programs as well as qualitative research conducted with both BIP facilitators and participants to learn why certain strategies were used and how those who are closely connected with program believe they operated examples include^{13 15-17 19 22}.

Because of the lack of agreement in the scientific community about whether partner violence is primarily a psychological, cognitive, developmental, or social

problem, the literature is replete with a variety of approaches to intervention. In spite of this heterogeneity in the research arena, the vast majority of programs that are being implemented across the U.S. and Canada employ what can generally be described as a *feminist-informed cognitive-behavioral therapy (CBT)* model^{1 2 23 24}. Although differences clearly exist across programs (e.g., some are tailored to specific ethnic communities, some include motivational components, some are less feminist-oriented than others), numerous commonalities also exist, including strategies intended to help perpetrators understand and control their violent behaviors and to change attitudes or beliefs that may support the use of violence.

The theories that we selected as a starting point in our effort to explain for whom, under what circumstances, and why BIPs would be effective are grounded in two middle-range theories, Cognitive Behavioral Theory (the basis for CBT) and Social Cognitive Theory. Feminist theory informs our model by providing one framework for understanding partner violence. Finally, the transtheoretical model offers an organizational structure for the implementation of the intervention.

Cognitive Behavioral Theory: A core tenant of cognitive behavioral theory (and the therapeutic approaches that have emerged) is the notion that distorted patterns of thinking lay at the root of many of our mental health problems. By identifying, challenging, and ultimately “re-learning” these thought processes, we can change how we react to external and internal stimuli, ultimately shifting our behaviors, if not our internal perceptions²⁵. Although the evidence is mixed, many

¹ Although this approach is often referred to as “the Duluth Model” because of the influential community-based program that emerged from that city in the early 1990s, we are referring to it as a *feminist-informed CBT model* because it is likely that very few current programs are true replications of the Pence and Paymar’s Duluth program.

researchers and treatment providers argue that domestically violent men hold certain beliefs that are linked to abusing women, whether a sense of entitlement or rigid gender roles or more specific thoughts about themselves or their partners; by working with men to identify these perceptions and thoughts and develop strategies to replace them with less destructive ones, men will also be able to change their behavioral patterns^{23 24}.

Social Cognitive Theory (SCT), which developed from Social Learning Theory, suggests that our personal characteristics, our environment, and our behaviors are intricately related, with each factor influencing and being influenced by the others. Frequently used as a basis for promoting health behavior changes, SCT postulates that effectively changing complex behaviors requires that we know what needs to change and how to do this (goals), that we believe that the positive outcomes of changing it (or the negative consequences of not changing it) outweigh the ease of the status quo (outcome expectations), and that we believe we can adopt and maintain that change (self-efficacy). Furthermore, it emphasizes the strong role that others in our lives play, recognizing that we learn from watching others and from receiving both positive and negative feedback^{26 27}. Thus, in the case of partner violent men, SCT promotes strategies that a) expose perpetrators to men who are not violent and can model the use of communication and conflict resolutions skills (especially men who were violent and have successfully changed), b) illustrate the consequences associated with using violence, including the cost of violating social norms, and c) offer the opportunity to learn and practice these skills¹⁷.

Feminist theory serves less as a programmatic theory and more as a common approach to explaining male violence against women. Although individual BIPs vary in how central of a theme this is, most acknowledge that men’s abusive behaviours are rooted (at least in part) in a social structure that grants men power over women and that defines “masculinity” as always maintaining control. Some program models directly confront men about sexist beliefs and behaviors, directly challenging all perceived expressions of male entitlement. Others are less confrontational, but most programs include at least some discussion about power and control ¹⁷.

The Transtheoretical/“Readiness to Change” Model generally provides underlying structure for most programs. At the heart of this framework is the recognition that for most of us, change is a process and can only happen when we are ready to do so, often after moving along a psychological continuum from not believing that we need to change, to planning to change to actively changing, and then maintaining that change ²⁸. Although participants often move along this continuum at different rates (with some never moving very far at all), program strategies are often chosen with the intention of helping to move men along this pathway (for example, through activities that encourage men to identify the benefits of using non-controlling communication over violence).

Figure 3 about here

Step 2 – Search for Evidence

Using a strategy designed in conjunction with a medical reference librarian, we are selecting search terms that have been identified in previous literature searches (examples include variations on words such as “batterer,” “perpetrator,” “intervention” “evaluation”, etc.). The disciplinary and interdisciplinary databases in our scope include (but are not limited to): Medline, EBM Reviews (including Cochrane Database of Systematic Reviews), Embase, PsycINFO, CINAHL, Criminal Justice Abstracts, Social Sciences Abstracts, International Bibliography of the Social Sciences (IBSS), Applied Social Sciences Index and Abstracts (ASSIA), ProQuest Criminal Justice, ProQuest Dissertations & Theses Full Text, ProQuest Social Services Abstracts, and Sociological Abstracts. The search will be limited to English language articles published from 1995 to present.

To answer our research questions and stay within our scope, our evidence search will be guided by a number of inclusion and exclusion criteria based on study design, program focus, outcomes measured and study participants. Quantitative and qualitative evaluations, regardless of study design, will be included if they a) assess programs that are offered in a community-based setting (rather than in institutionalized prison or military settings), b) include a facilitated group treatment/education component, c) run at least 8 weeks or 16 hours in duration, d) include primarily male, court-mandated partner violence perpetrators and e) measure at least one distal (partner-violence related re-assault/recidivism) and one proximal (skills, attitudes, intentions) outcome. Because these populations likely have unique needs that traditional programs are not intended to address, studies with treatment populations composed predominately of voluntary participants,

non-IPV offenders, and/or women and adolescence will be excluded, as will articles that are evaluative reviews or program descriptions. Articles that focus solely on reducing program attrition or increasing completion, that measure generic violent crime as outcomes, or that explore the impact of a participant's stage of change on program success will also be excluded. Articles will be assessed for inclusion or exclusion at this stage based on a review of titles, keywords, and /or abstracts; when necessary, the full text will be reviewed.

Step 3 – Study Appraisal and Data Extraction

After the first round of review, identify which proximal outcomes – attitudes/beliefs, skills, and desire to cease using violence against a partner – are measured in each remaining article. The review team will categorize articles according to these proximal outcomes (articles can be in more than one outcome category). Examining the literature affiliated with each proximal outcome separately, all articles will be reviewed in full text form by two separate team members, who will extract the following information: a) the program strategies that are described, b) how the proximal outcome is measured, c) the context that surrounds participation (e.g., do participants have to be in substance use treatment before attending the BIP), d) what – if any – mechanisms are described as leading to the change anticipated, and e) the study design/fit for purpose. We expect additional articles will be excluded after this in-depth review process, and disagreements over inclusion/exclusion or the information being extracted will be discussed as a team.

Next, the research team will attempt to develop a description of the “context-mechanism-outcome” configuration for each evaluation that is included in the review (which may consist of a family of articles); in other words, we will ask how the opportunities, resources and/or constraints provided by the program strategies (context) result in a change of the reasoning, beliefs and norms of the participants (mechanism), which consequently resulted in changes in proximal outcomes (outcomes). When gaps in this information become apparent, as we anticipate they will be, they will be recorded for use in the development of recommendations for future implementation and evaluation research. The team will also make note of study design, particularly if it seems to bias the results (e.g. if only successful program participants were included in the evaluation) so that we can appraise the strength and fit for purpose of the evidence provided.

Step 4: Analysis and Synthesis

The focus of realist review is to refine our understanding of how and why interventions succeed in achieving the intended outcomes. Having started with a hypothesized theoretical model for BIPs, this stage will involve examining the evidence gathered and determining whether it supports or contradicts our proposed programmatic theory. Again, using the proximal outcomes as our organizing framework, we will look carefully at each evaluation or family of evaluations to assess how the data that were extracted from the studies inform our understanding of how batterer intervention works. We will be particularly cognizant of the ways in which different contextual factors – when addressed – appear to influence the mechanisms that lead to these proximal outcomes. As

described by other realist review teams, this will be an iterative process that will benefit from the interdisciplinary perspectives and expertise of our research team. Ultimately, we anticipate using the synthesis process to refine our original theoretical model in light of our review findings.

Step 5: Presentation and dissemination

The findings from this process will be presented in at least two formats: through at least one peer-reviewed article that conforms to the RAMESES publication standards put forth by Wong et al.²⁹ that is intended to inform implementation scientists and others in academic settings; and through targeted outreach and conversations intended to reach decision-makers and practitioners: program directors, policy makers, and community coalitions charged with overseeing coordinated responses to partner violence. This will include presentations at domestic violence and batterer treatment coalitions and conferences and the preparation of plain-language reports and briefs for dissemination through national and international networks.

DISCUSSION

This review was conceived after conversations with batterer intervention treatment providers, domestic violence advocates, and justice system employees revealed frustration over the current understanding of batterer intervention programs. Results from evaluations and systematic reviews were found to be both conflicting and confusing, and without substantial guidance about how approaches to batterer treatment could be improved. In response, our team decided to complete this realist synthesis of perpetrator treatment program evaluations with

the aim of clarifying how and why BIPs work for some men and the role that certain contextual factors may play in that success (or lack of success).

This is the first realist review of the perpetrator treatment evaluation literature that we know of, and we believe it will offer key insights into to the debate over how communities can respond to partner violence. By focusing on the mechanisms that lead participants to change (rather than only looking at whether or not change was achieved), we believe we will provide much-needed insight into promising (and not-so-promising) theoretically informed strategies that can lead to the types of changes that will allow men to reduce their use of violent and/or controlling behaviours.

One of the key contributions of this review, in relation to the majority of BIP program evaluations and systematic reviews that have been done, is that we will focus on the impact that programs have on *proximal* outcomes, rather than on re-offending or recidivism. We believe this is important for several reasons. To truly understand *why* programs are or are not successful (and for whom and under what conditions), we need to gain a clear picture of the processes that lead to these outcomes. In the case of batterer treatment, it is unlikely that participants are magically transformed into nonviolent partners simply because they attended a program; rather, the program promotes certain outcomes within participants that then lead to these more distal behavioural changes. Identifying what these proximal outcomes are and how programs can achieve them is a key part of understanding this process.

This is not to say that we think the elimination of violent and controlling behaviours on the part of perpetrators should not be the ultimate goal of community responses; it most definitely should. We anticipate that one of the conclusions that may be drawn by this review is that significantly more work needs to be done to show how the achievement of these proximal outcomes ultimately can lead to the cessation of violence. However, as we look more closely at the literature surrounding BIPs, it has also become apparent that factors unrelated to the programs themselves also contribute to the likeliness that men will cease to engage in violence against partners and family members. Influences at both the individual and interpersonal level as well as the community and social level are also at play, and these need to be identified and accounted for within the larger coordinated response.

Authors' Contributions: AV conceptualized the study, reviewed the overview literature on batterer intervention and completed the initial program process and theoretical models. AV and PO led the design and drafting of the review protocol, assisted by RC, DF, and WD. AV and RC wrote the first draft of this manuscript, which was critically reviewed by PO, WD, and DF. All authors have approved submission of the manuscript to BMJ Open.

Acknowledgements: St. Michael's Hospital Information Specialist Carolyn Ziegler designed and conducted the search strategy. Special thanks to Dr. Maritt Kirst who provided methodological advice during the early stages of protocol development.

Competing Interests: The authors declare no competing interests.

Funding Statement: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

For peer review only

References

1. Dutton DG, Corvo K. Transforming a flawed policy: A call to revive psychology and science in domestic violence research and practice. *Aggression and Violent Behavior* 2006;**11**:457 – 83.

2. Gondolf EW. Theoretical and research support for the Duluth Model: A reply to Dutton and Corvo. *Aggression and Violent Behavior* 2007;**12**:664-57.

3. Dixon L, Archer J, Graham-Kevan N. Perpetrator programmes for partner violence: Are they based on ideology or evidence? *Legal and Criminological Psychology* 2012;**17**:196-215.

4. Gelles RJ. The politics of research: The use, abuse, and misuse of social science data—The cases of intimate partner violence. *Family Court Review* 2007;**45**(1):42-51.

5. Sayer A. Critical realism and the limits to critical social science. *Journal for the Theory of Social Behaviour* 1997;**27**(4):473-88.

6. Pawson R. *Evidence-Based Policy*. London: SAGE Publications Ltd, 2006.

7. Miller M, Drake E, Nafziger M. What works to reduce recidivism by domestic violence offenders. Olympia, WA: Washington State Institute for Public Policy, 2013.

8. Jackson S, Feder L, Forde DR, et al. Batterer Intervention Programs: Where do we go from here. Washington, D.C.: U.S. Department of Justice, Office of Justice Programs, 2003.

9. Stover CS, Meadows AL, Kaufman J. Interventions for Intimate Partner Violence: Review and Implications for Evidence-Based Practice. *Prof Psychol-Res Pr* 2009;**40**(3):223-33.
10. Babcock JC, Green CE, Robie C. Does batterers' treatment work? A meta-analytic review of domestic violence treatment. *Clinical Psychology Review* 2004;**23**(8):1023-53.
11. Pawson R, Greenhalgh T, Harvey G, et al. Realist Review - A New Method of Systematic Review Designed for Complex Policy Interventions. *Journal of Health Services Research & Policy* 2005;**10**(Suppl 1):21-34.
12. Sartin R, Hansen D, Huss M. Domestic violence treatment response and recidivism: A review and implications for the study of family violence. *Aggression and Violent Behavior* 2006;**11**(5):425-40.
13. Saunders D. Group Interventions for Men Who Batter: A Summary of Program Descriptions and Research. *Violence and Victims* 2008;**23**(2):156-72.
14. Carter LS. Batterer Intervention: Doing the Work and Measuring the Progress. A Report on the DEcember 2009 Dexperts Roundtable. San Francisco: Family Violence Prevention Fund & National Institute of Justice, 2010.
15. Paymar M, Garnes G. Countering Confusion about the Duluth Model. Duluth, MN: Domestic Abuse Intervention Programs, 2007.
16. Price BJ, Rosenbaum A. Batterer Intervention Programs: A Report From the Field. *Violence and Victims* 2009;**24**(6):757-70.

17. Mankowski ES, Haaken J, Silvergleid CS. Collateral Damage: An Analysis of the Achievements and Unintended Consequences of Batterer Intervention Programs and Discourse. *Journal of Family Violence*;17(2):167-84.
18. Gondolf EW. How batterer program participants avoid reassault. *Violence Against Women* [HWWilson - SSA] 2000;6(11):1204.
19. Rosenbaum A, Leisring PA. Group intervention programs for batterers. *Journal of Aggression, Maltreatment and Trauma* 2001;5(2):57-71.
20. Connelly JB. Evaluating complex public health interventions: theory, methods and scope of realist enquiry. *J Eval Clin Pract* 2007;13(6):935-41.
21. Wong G, Greenhalgh T, Pawson R. Internet-based medical education: a realist review of what works, for whom and in what circumstances. *Bmc Medical Education* 2010;10.
22. Silvergleid CS, Mankowski ES. How Batterer Intervention Programs Work: Participant and Facilitator Accounts of Processes of Change. *Journal of Interpersonal Violence* 2006;21(1):139-59.
23. Diefenbeck C. Group therapy for male batterers: Comparison of cognitive-behavioral & object relations approaches. *Journal of Psychological Nursing & Mental Health* 2003;41(10):18-25.
24. Gilchrist E. The Cognition of Domestic Abusers: Explanations, Evidence and Treatment. In: Gannon T, Ward T, Beech A, et al., eds. *Aggressive Offenders' Cognition: Theory, Research, and Practice*: John Wiley & Sons, 2007:247-66.

25. Hupp SD, Reitman D, Jewell JD. Cognitive-Behavioral Theory. In: Herson M, Gross A, eds. Handbook of Clinical Psychology. New Jersey: John Wiley & Sons, Inc, 2008:263-90.
26. Bandura A. Health Promotion from the Perspective of Social Cognitive Theory. Psychology and Health 1998;**13**:623-49.
27. Baranowski T, Perry CL, Parcel GS. How individuals, environments, and health behavior interact. Health behavior and health education San Francisco, CA: Jossey-Bass 2002:165-84.
28. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. American journal of health promotion 1997;**12**(1):38-48.
29. Wong G, Greenhalgh T, Westhorp G, et al. RAMESES publication standards: realist syntheses. BMC medicine 2013;**11**(1):21.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

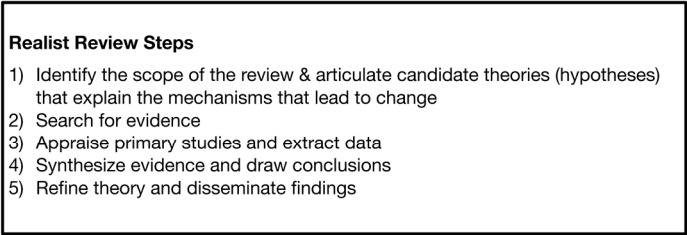


Figure 1: Key Steps in Realist Review

Figure1
279x361mm (300 x 300 DPI)

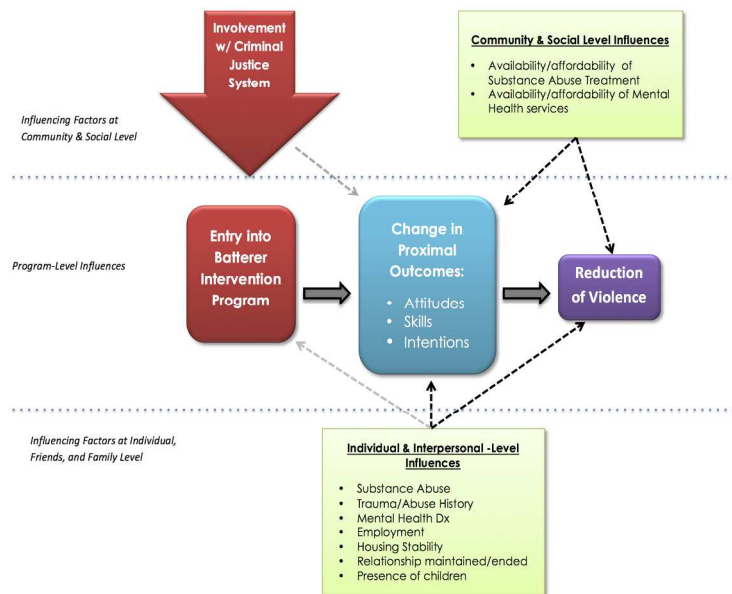


Figure 1: General Batter Intervention Program Process and Influences

Figure2
279x361mm (300 x 300 DPI)

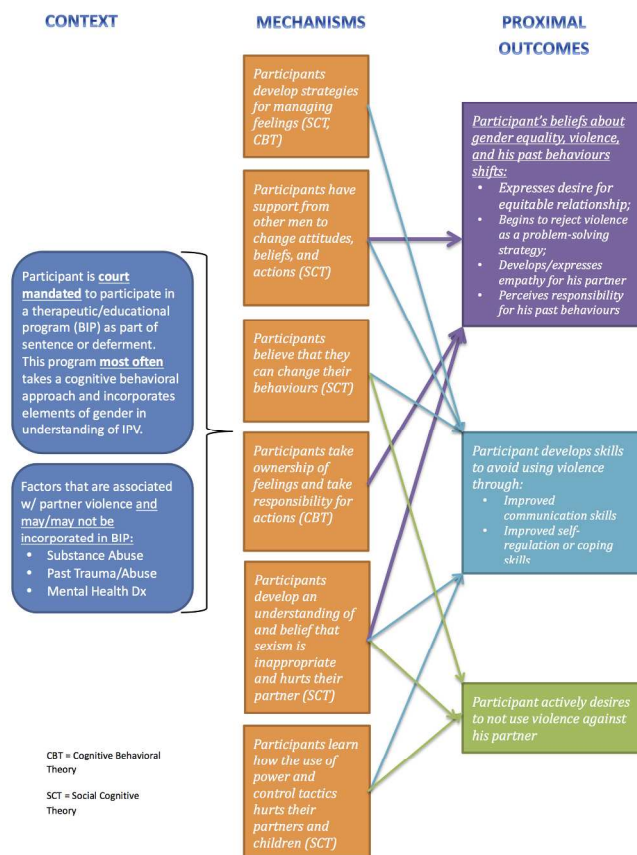


Figure 1: Candidate BIP Program Theory

Figure3
215x279mm (300 x 300 DPI)

BMJ Open

Searching for the mechanisms of change: A protocol for a realist review of batterer treatment programs

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2015-010173.R1
Article Type:	Protocol
Date Submitted by the Author:	15-Jan-2016
Complete List of Authors:	Velonis, Alisa; St. Michael's Hospital, Centre for Research on Inner City Health; University of Illinois at Chicago, School of Public Health, Community Health Sciences Cheff, Rebecca; St. Michael's Hospital, Centre for Research on Inner City Health Finn, Debbie; St. Michael's Hospital, Centre for Research on Inner City Health Davloor, Whitney; St. Michael's Hospital, Centre for Research on Inner City Health O'Campo, Patricia; St. Michael's Hospital, Centre for Research on Inner City Health; University of Toronto, Dalla Lana School of Public Health
Primary Subject Heading:	Public health
Secondary Subject Heading:	Research methods, Mental health
Keywords:	domestic violence treatment, realist review, evidence-based practices

SCHOLARONE™
Manuscripts

Only

SEARCHING FOR THE MECHANISMS OF CHANGE: A PROTOCOL FOR A REALIST REVIEW OF BATTERER TREATMENT PROGRAMS

Alisa J. Velonis, MPH, PhD;^{1,2}
Rebecca Cheff, MPH;¹
Debbie Finn, RN BScN MN;¹
Whitney Davloor, MPP;¹
Patricia O'Campo, PhD^{1,3}

Corresponding Author:

Alisa J. Velonis, MPH, PhD
University of Illinois at Chicago
School of Public Health, Division of Community Health Sciences
1603 W. Taylor St
Chicago, IL 60612 U.S.A.
avelonis@uic.edu
312-355-1217

Institutional Affiliations:

- ¹ St. Michael's Hospital, Centre for Research on Inner City Health, Toronto, Ontario, Canada
- ² University of Illinois at Chicago, School of Public Health, Division of Community Health Sciences, Chicago, Illinois, U.S.A.
- ³ University of Toronto, Dalla Lana School of Public Health, Toronto, Ontario, Canada

Key Words:

Domestic Violence Treatment, Batterer Treatment, Realist Review Protocol,

Word Count: 5,550

ABSTRACT

Introduction: The conflicting results reported by evaluations of typical batterer intervention programs leaves many judicial officials and policy makers uncertain about the best way to respond to domestic violence and whether to recommend and fund these programs. Traditional evaluations and systematic reviews tend to focus predominately on whether the programs “worked” (e.g. reduced recidivism) often at the exclusion of understanding for whom they may or may not have worked, under what circumstances, and why.

Methods and Analysis: We are undertaking a realist review of the batterer treatment program literature with the aim of addressing this gap. Keeping with the goals of realist review, our primary aims are to identify the theory that underlies these programs, highlight the mechanisms that trigger changes in participant behavior, and finally explain why these programs help some individuals reduce their use of violence and under what conditions they are effective or not effective. We begin by describing the process of perpetrator treatment and proposing an initial theoretical model of behavior change that will be tested by our review. We then describe the criteria for inclusion of an evaluation into the review, the search strategy we will use to identify the studies, and the plan for data extraction and analysis.

Ethics and Dissemination. The results of this review will be written up using the RAMESES Guidelines for Realist Synthesis and disseminated through peer-reviewed publications aimed at the practitioner community as well as presented at community forums and at violence against women conferences. Ethics approval was not needed.

Strengths and Limitations of this study

- Realist syntheses are based on the development of a solid initial theory describing how, why, for whom, and under what conditions program strategies generate key outcomes.
- We present our initial theory for a realist synthesis of batterer intervention program evaluations and our explanation for the importance of conducting such a review.
- Our initial theory draws from existing theory about batterer intervention programs and presents a hypothesis about how the primary strategies of education, skills-building, and group process might generate immediate outcomes including participants' desire to develop alternatives to violence, use of non-violent communication and violence avoidance skills, and development of empathy for partners, and shifts among others.
- We present a search strategy and approach to analysis that are consistent with realist principles.

INTRODUCTION/BACKGROUND

Since the emergence of early batterer treatment or intervention programs (BIPs) in the 1970s and 1980s, discussions about their efficacy have proliferated within research, professional, and policy circles.¹⁻³ In attempting to answer the general question, “Do these programs work?” a number of sub-surface debates have emerged, highlighting points of contention about the nature of intimate partner violence (IPV), about the multiple levels and types of influences that may contribute to abusive behavior, about what “success” means in terms of batterer treatment, and about the “right” approach to both achieving and evaluating this success.^{1, 4} These differences in perspective lead to conflicting conclusions about the effectiveness of these programs, making the job of navigating the literature that surrounds batterer intervention or treatment programs challenging.

Yet, as part of the official response to domestic violence across North America, sentencing guidelines or other codified judicial requirements frequently require individuals who are convicted of crimes against intimate partners to attend treatment or educational programs (the content and format of which can vary widely) as a condition to receiving a deferred sentence, probation, or parole.⁵⁻⁷ This, among other factors, has spurred a proliferation of program evaluations and systematic reviews, most sharing a defined goal of determining whether or not the programs that currently exist can be proven to directly reduce subsequent violence and criminal behaviors.^{3, 8} While there has been significant debate over what theoretical approach(s) should be used to guide these programs (e.g. feminist theory, family systems theory, cognitive behavioral theory), the nature of these discussions tends to be as political (e.g., pro- or anti-feminist in rhetoric) as it is scientific⁹⁻¹². Few systematic reviews have attempted to examine the underlying

programmatically and understand how, why, and in what contexts these programs work – or do not work.

Focusing on Theory Rather Than on Programs

Drawing upon the work of Sayer¹³ and other realist philosophers, Pawson¹⁴ describes interventions as “complex process that are inserted into complex structures” (pg. 79). Budgets get cut, referrals increase or decrease, participants resent attending mandated programs staff feel overworked and underappreciated, and the list goes on. Even under the best of conditions, behavior change is difficult, and programs often use multiple theoretical and strategies to help clients move along the path to improvement.

Intervention programs for batterers are particularly tricky. While multiple models for batterer intervention exist, most function within the framework of a larger community and criminal-justice oriented response to domestic violence.¹ Participants are primarily – although not exclusively – required to attend BIPs as part of probationary or deferred sentencing agreements,⁶ and while the programmatic details vary across jurisdictions, most BIPs are designed as a series of educational and skills building group sessions that run from 12 – 52 weeks.⁷ Regardless of the specific therapeutic, philosophical, or political framework used, these programs are impacted by a variety of internal and external factors, including the characteristics and experiences of participants and staff, the mission of the lead organization, the levels of communication between the programs, the local courts or probationary departments, and victim-centered domestic violence services, and the social and political climate of the larger community. In turn, these factors influence how

1
2
3 programs run, how closely they adhere to the program design, how the strategies are
4
5 received by participants, and more.
6
7

8 Yet, a great deal of the research describing the “effectiveness” of batterer
9
10 intervention programs has been designed to minimize the influence of these real-world
11
12 contextual factors, generally by controlling for many of the very forces that could explain
13
14 program success or failure (e.g., cultural backgrounds, income, substance use). Often,
15
16 experimental and quasi-experimental evaluations are held up as the gold standard; these
17
18 models endeavor to link the intervention – and ONLY the intervention – to a narrowly
19
20 defined outcome, usually recidivism or re-offense^{3, 8, 15, 16}. A result has been the
21
22 proliferation of a vast body of literature that shows mixed evidence of program success
23
24 with little explanation of why. It was our frustration with this lack of explanation (and the
25
26 frustration that program administrators and domestic violence advocates expressed to us
27
28 about not knowing what to do to improve treatment) that led us to conduct this realist
29
30 synthesis of the literature.
31
32
33
34
35

36
37 What follows is a description of a protocol we have developed for undertaking this
38
39 review. As with other realist reviews, the purpose of this synthesis is explanatory: to
40
41 articulate underlying program theories and use evidence to determine their usefulness and
42
43 relevance for batterer treatment.¹⁴ In writing and sharing this protocol, we set forth three
44
45 goals. First, based on our initial understanding of BIPs, to propose and explain a
46
47 “preliminary rough theory” that captures the framework used by the majority of BIPs with
48
49 the aim of identifying both strengths and gaps. Second, we wish to illustrate how a realist
50
51 perspective can add value to our understanding and interpretation of BIP evaluation by
52
53 shifting focus from whether programs work to how and why they [should] work. Finally,
54
55
56
57
58
59
60

we desire to provide transparency for the forthcoming review, enabling readers to know the specific steps that will be followed throughout the review process, and to highlight some of the specific challenges we face as we move forward.

Part of what makes realist synthesis unique is the emphasis it places on proposing, testing, and ultimately refining theory. Initially, reviewers conceptualize a *rough, preliminary theory* that explains “what is supposed to happen?” and “why is that supposed to work”¹⁷ for the program in question (a program theory). In the language of realist synthesis, the key elements that need to be identified and understood in relation to one another include the program *strategies* (the activities that comprise the program), *contextual factors* that influence how and for whom the programs operate (such the individual-level characteristics mentioned above as well as the structural context in which programs operate, such as funding and statutory requirements), *participant outcomes* (e.g., participants reduce their use of violence or recidivism), and the hidden *mechanisms* or “generative process” that often take place within participants’ minds and trigger the outcome (or, as Wong and colleagues (2013, pg. 6) put it, “what it is about a program that generates change”).¹⁷⁻¹⁹ For each intervention strategy that a program includes, there may be multiple pathways that lead to multiple outcomes; each pathway has its own set of contextual factors and mechanisms at play, and the outcomes can be visible or hidden, intermediate or final, and intended or unintended.²⁰ At the end of a realist synthesis, the primary goal is the generation of a refined theory that takes the shape of a set of Context-Mechanism-Outcome configurations (CMOs), a heuristic used to illustrate these relationships and pathways.²⁰ It is at this point that our theoretical lens becomes less

1
2
3 focused on the programmatic elements of BIPs and more focused on the mechanisms of
4
5
6 change.
7

8 This protocol paper is organized into several sections, following the steps of realist
9
10 synthesis laid out by Pawson and others ¹⁹ (Figure 1). The core of any realist synthesis or
11
12 evaluation is the description of the programmatic theory that underpins the activity being
13
14 evaluated (Step 1). We began this process by formulating our preliminary (rough) theory
15
16 describing the strategies most BIPs employ, the circumstances in which they operate, what
17
18 they aim to accomplish and how it appears to us that these strategies will lead to those
19
20 outcomes; essentially, this is a program theory that describes, in general terms, what is
21
22 supposed to change and why as a result of the program. As we develop this protocol, we
23
24 are in the initial stages of the review process, and what is reflected here reflects our initial
25
26 thinking about BIPs. Furthermore, as Jagosh and colleagues remind us, realist reviews are
27
28 abductive in nature, meaning that we infer “to the best explanation,” iteratively “examining
29
30 evidence and developing hunches or ideas about the causal factors linked to that evidence)
31
32 (pg. 135)²⁰ After laying out our preliminary program theory, we go on to describe the
33
34 remaining steps (2 – 5) that will be taken to conduct this review. Finally, we conclude with
35
36 a discussion of why we believe this approach will contribute to our understanding of
37
38 batterer treatment programs and how it can inform not only program implementation but
39
40 also larger policy.
41
42
43
44
45
46
47
48
49
50

51 *** FIGURE 1 ABOUT HERE ***
52
53
54
55
56
57
58
59
60

METHODS/DESIGN The impetus for this review was a series of conversations with program facilitators and judicial personnel who had grown frustrated with the lack of conclusive information about what their systems could do to reduce the perpetration of partner violence. As they saw it, the evidence was insufficient to declare batterer interventions useless, yet these programs clearly did not work in all circumstances. Using this to inform our approach, we framed our research question as *for whom, and under what conditions, will batterer intervention programs help men who have been identified as perpetrators of partner violence reduce their violent behaviors and why?*

Step 1a. Establishing the Scope of our Work

Before refining our scope and articulating the processes we believe to be at work, we needed to understand what the bigger picture of batterer response looked like in North America. After selecting a handful of evaluations, systematic reviews, and reports from both the scientific and grey literature,^{2, 3, 5, 8, 15, 21-27} we learned that the majority of individuals who attend batterer intervention in the U.S. and Canada undergo a two-part process: first, they enter the criminal justice system and are adjudicated for an offense against a partner, and second, they are mandated to attend an educational/therapeutic “treatment” program – what we refer to as a BIP – as part of their sentence or agreement with the court. Because the vast majority of participants in BIPs first have contact with the criminal justice system (and because our interest is in explaining how and why these programs work), we consider this to be part of the larger environment in which these programs exist, rather than as one of the strategies or interventions that make up BIP programs. We recognize that this decision may become a limitation for this study, and that to gain a true understanding of the mechanisms that underlie BIPs, we may need to look

more critically at the systems-level issues at work (for example, whether or not the level of communication between BIP program staff and court officials has an impact on BIP outcomes); if this appears to be the case as we proceed, we will revisit this decision. However, the ongoing debate over the efficacy of the programs themselves (and the lack of information about community-level systems contained within most program evaluations) leads us to begin this review by focusing on BIPs.

Another issue we encountered while considering the scope of the review was the enormity of what we are calling “individual perpetrator characteristics” and how they may interact with program strategies to impact whether participants respond to the program (positively or negatively). A substantial amount of literature supports the contention that men who engage in violence against intimate partners and family members are not all alike.²⁸ Over the past 30 years, researchers have attempted to categorize batterers according to psychological, behavioral, attitudinal, and/or motivational characteristics using descriptors such as *family-only* or *typical* batterers, *dysphoric/borderline* or *passive-aggressive dependent* batterers, and *sociopathic* or *generally violent/antisocial* batterers.²⁹⁻

³² Likewise, while Johnson does not provide a batterer typology, per se, he differentiates between types of IPV (*coercive-controlling violence*, *situational partner violence*, and *violent resistance*), which suggests that the motivation behind these categories of violence would be different for different instigators.³³ Finally, substantial evidence points to an overlap between substance abuse, a history of trauma, neglect, and/or victimization, and other psychological conditions, some of which are captured in the batterer typologies described above but which may also emerge as independent issues that programs may or may not be prepared to address.^{34, 35}

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

In no way are we claiming that individual-level characteristics *cause* someone to engage in abusive behaviors; rather, those factors may interfere with the effectiveness of program strategies.^{36, 37} Unfortunately, individual programs are often unable to assess for and address the myriad of issues that participants may bring to the intervention, and even in jurisdictions where completing an intake assessment for substance abuse and mental health problems is mandatory, the availability of and coordination between various treatment modalities can vary widely. Understandably, much of this detail is left out of evaluation write-ups and formal reports, yet these are influential factors that need to be acknowledged and addressed. For the purposes of this review, we have decided that we will consider them to be among the contextual factors that can influence program effectiveness, and where they are mentioned, we will note them accordingly. However, because of the wide breadth of possible influences and the relative dearth of information collected about them in most evaluations, we have chosen not to limit our review to only those evaluations that address these issues.

As we began developing our rough program theory for BIPs, we first drafted a flow chart illustrating the “big picture” processes at work in most perpetrator interventions (Figure 2). After reviewing various program descriptions and evaluations, we differentiated two sets of outcomes that were generally discussed: “proximal” outcomes, or changes that happen within the participants as a result of participation in the program activities (such as changes in attitudes, skills, and intentions); and “final” outcomes, which would include recidivism and re-assault and which are generally measured as longer term consequences of program participation.

The majority of batterer intervention evaluations are concerned primarily or exclusively with recidivism and/or re-assault, which we found problematic for several reasons. First, we contend that for BIPs to impact recidivism or re-assault, they first must achieve these more immediate changes in attitudes, motivations, and skills. The final outcomes are at least one step removed from the BIP itself in that any changes in recidivism or violent behavior that are not preceded by changes in attitudes, motivations, and skills may not be the result of the BIP program. Furthermore, even after the proximal outcomes are met (if and when they are met), numerous factors unrelated to the BIP program itself may influence whether participants reoffend. For example, if a perpetrator with co-occurring substance abuse or clinical depression receives little or no additional treatment for those problems (whether because he refuses treatment or because treatment is not available, affordable, or accessible), the progress he makes towards realizing a non-abusive relationship (that may be achieved through BIP participation) may be offset by the lack of assistance for these other problems.

Another problem with using recidivism or re-offense measurements as the sole indicators of success or failure of BIPs is that definitions of these outcomes often differ across studies and may not be limited to assaults against a partner but can include a conviction for any violent crime, an arrest for a violent crime (regardless of conviction) or even any subsequent arrest.¹ Finally, because many – if not most – acts of abuse do not result in law enforcement intervention, using this as an outcome likely underestimates the recurrence of these behaviors.^{1, 3, 38} Unfortunately, other measures of re-assault are likely to be equally unreliable, in that unless evaluators contact the perpetrators' current partners, these data rely primarily on participants' self-reported descriptions of their

behaviors. Even when victim reports are included, the numbers are often small and subject to self-report bias. For these reasons, we chose to limit our review to the relationship between BIPs and these proximal outcomes, which are addressed in more detail below.

Step 1b. Forming an Initial Theory of Change: Looking for the Mechanisms

Because of the lack of agreement in the scientific community about whether partner violence is primarily a psychological, cognitive, developmental, or social problem, the literature is replete with a variety of intervention approaches. One of the challenges we faced was in identifying a universal program model that reflects all or most BIPs; not only do different state and provincial jurisdictions outline different standards for program length and content,^{6 7} but even at a local level, programs can vary tremendously in how they are implemented and who they serve. In spite of this heterogeneity, the vast majority of programs across the U.S. and Canada employ what could be called a **feminist-informed cognitive-behavioral therapy (CBT)* model.^{10, 28, 39} While the specific content, philosophical emphases, and details differ across jurisdictions and programs, our initial review of program descriptions indicate that most BIPs include a common set of elements, including:

a) the use of educational strategies that challenge beliefs about gender equity, relationships, and the impact of abuse; b) skills-building activities intended to provide alternatives to abuse and violence; and c) a facilitated group process that offers participants both support and accountability.⁴⁰

* Although this approach is often referred to as “the Duluth Model” because of the influential community-based program that emerged from that city in the early 1990s, we are referring to it as a *feminist-informed CBT model* because it is likely that very few current programs are true replications of the Pence and Paymar’s Duluth program.¹⁶

Based on this understanding of “generic” BIPs, the review team constructed our initial (rough) programmatic theory. After identifying what we felt were the most essential program-level (proximal) outcomes – those that would be necessary to lead to further change in longer-term (final) outcomes like recidivism – we worked backwards, asking ourselves what strategies were likely to be linked to each outcome and how that strategy triggers the outcome, and what contexts are the most relevant allow that to happen? We recognized that each strategy may be linked multiple outcomes, that multiple mechanisms are often at work within each strategy, and that sometimes change has to happen in a particular order. For example, one “education strategy” that appears common to most BIPs is to discuss the negative impact that violence and abuse has on one’s partner and children. For participants who are capable of feeling remorse and empathy (i.e., who do not have sociopathic tendencies or an antisocial personality disorder) (*context*), learning about these damaging impacts triggers both shame and guilt for past behaviors (*mechanism*) and a desire to stop hurting people who he loves (*mechanism*). Both of these mechanisms can lead to the perpetrator feeling motivated to stop using abusive behaviors (*outcome*) and to desire alternatives to violence (*outcome*). Another strategy commonly employed is to teach perpetrators skills they can use to avoid becoming violence (such as recognizing emotional triggers and calmly walking away). Learning and then practicing these behaviors triggers a level of self-confidence in participants (*mechanism*) that leads to the eventual adoption of these skills (*outcome*). However, before a participant is likely to truly benefit from these skills-building sessions, he most likely needs to already feel motivated to stop using abusive behaviors and to desire alternatives to violence. Figure 3 illustrates several pathways in which program strategies may link to our proximal outcomes; this is a partial

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

model intended to exemplify our process and is by no means a complete outline of our preliminary, rough theory.

Figure 3 about here

At this stage, we wish to acknowledge that we fully anticipate revising and refining this as we proceed with the review. In their reflection on realist review, Jagosh and colleagues describe their struggle with identifying a singular theoretical construct that guided the subject of their review, and observed the ways in which context, mechanisms, and outcomes often overlap, with an outcome in one chain of evidence serving as a context in a subsequent one.²⁰ We also recognize that using proximal outcomes may pose a challenge, as the nature of these outcomes – especially motivations and attitudes – can be subjective in nature and difficult to capture, and some may not be captured at all. As we move forward with our review, we will assess how well these constructs are assessed, and where we believe critical gaps may exist.

As we constructed our protocol, we identified several existing theories of behavior change (e.g., social cognitive theory, stages-of-change, etc.) that appear useful for understanding how BIPs (are supposed to) work,^{25, 41} yet our attachment to these is preliminary. Through ongoing discussions within the team and continuously asking what causes a particular strategy to lead to a particular outcome during the synthesis phase of the review, we will continue to refine our models and flesh out the mechanisms at work.

Step 2 – Search for Evidence

Search Strategy. In partnership with a medical reference librarian (who is conducting the searches but is not part of the review team), we selected search terms that have been identified in previous literature searches (examples include variations on words such as “batterer,” “perpetrator,” “intervention” “evaluation”, etc.). The disciplinary and interdisciplinary databases in our scope include (but are not limited to): Medline, EBM Reviews (including Cochrane Database of Systematic Reviews), Embase, PsycINFO, CINAHL, Criminal Justice Abstracts, Social Sciences Abstracts, International Bibliography of the Social Sciences (IBSS), Applied Social Sciences Index and Abstracts (ASSIA), ProQuest Criminal Justice, ProQuest Dissertations & Theses Full Text, ProQuest Social Services Abstracts, and Sociological Abstracts. Additionally, databases and electronic resources such as the Minnesota Center Against Violence and Abuse (MINCAVA) Electronic Clearinghouse and Google will be used to search the “grey literature” for unpublished and informal evaluations. Based on prior experience with realist reviews as well as the literature describing them, we expect that as we proceed with the review, we will return to this step several times as we expand and/or refine our scope as necessary. All searches will be limited to English language articles published from 1995 to present.

Inclusion/Exclusion Criteria. Unlike more traditional syntheses, we are less concerned with whether or not an evaluation meets certain methodological standards (e.g., is a randomized trial or includes a control/treatment group design) and more with the type of information it can provide about how, why, and for whom BIPs work; this is reflected in our initial inclusion and exclusion criteria. Quantitative, qualitative, and mixed methods evaluations, regardless of study design, are to be included if the programs they assess:

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- a) *Are offered in community-based rather than institutionalized settings, such as the military or prison.* Even if the program activities and format resemble those found in community-based settings, BIPs conducted in these institutions likely involve a unique set of contextual factors and mechanisms that impact how and for whom the programs work (e.g., a soldier's commanding officer often ensures compliance with program requirements, unlike in civilian settings);³
 - b) *Include some form of facilitated group treatment/education component,* the most common format required across North America. Most importantly, for purposes of this review, we are intentionally excluding research involving couples counseling or individual psychotherapy, as these are sufficiently different from BIP programs, are considered controversial by many practitioners, and are often specifically prohibited by judicial statute;⁷
 - c) *Run at least 8 weeks or 16 hours in duration.* Most judicial statutes require at least 12 hours,^{7, 24} and we do not believe that programs that are shorter could be comparable in scope or content;
 - d) *Involve primarily male, court-mandated partner violence perpetrators.* Both research and observation suggest that the use of violence may be different for women than for men, necessitating different approaches.⁴² Likewise, men who voluntarily choose to participate in perpetrator treatment likely have different characteristics, and programs designed to cater primarily to this population may be addressing different causes of violence; and
 - e) *Measure at least one proximal outcome,* such as skills, attitudes, intentions. This criterion emerged as we developed our preliminary program theory (described

above) and reviewed articles identified during an early search. At the start of the project, we planned to include evaluations that measured at least one of the final outcomes, but the iterative process of theory identification led us to shift to a more limited scope.

Articles that are program descriptions or evaluative reviews will be set aside for use as background, but will not be included in the formal synthesis process. Articles that focus solely on reducing program attrition, increasing completion rates, or are limited to identifying the impact that participant's stage of change has on program success will also be excluded.

Article Screening. Using the search strategies outlined above, our librarian will generate a list of articles and (when available) abstracts, which we will divide among members of the review team, who will review the titles and abstracts to determine if the paper is (a) focused on domestic violence perpetrator programs at all, and if it appears to (b) fit within the aforementioned inclusion/exclusion criteria. Screeners will be asked to categorize each article as "include," "exclude," and "maybe." When an abstract is not available, titles will be used to determine if the article is appropriate for the review (e.g., does the title mention BIPs?); if a title is insufficient to make this determination, the article will remain in the list of potential evaluations to be included until the complete text can be reviewed.

To ensure inner rater reliability, we will randomly select a handful of titles and abstracts that all screeners will review; as a group, we will discuss each screener's categorization and, as necessary, come to consensus about articles in which screeners

disagree. Once we are satisfied that all screeners share an understanding of the criteria and screening objectives, each team member will complete her/his assignments.

After this initial screening phase, all articles labeled as “include” and “maybe” will be re-distributed among the review team members, who will complete a second screen of the remaining the titles and abstracts. Once all members completed this task, the review team will again discuss this process. Articles that the first screener labels as “include” but the second screener decides to “exclude” will discussed, and consensus reached.

Finally, the complete article or paper will be obtained for all remaining titles. Once again, inner rater reliability will be assessed by having all reviewers read the same set of five articles, individually make recommendations on exclusion or inclusion, and then meet as a group to discuss the process. Each remaining article will be distributed among the review team members and skimmed in order to make a final determination of whether to include or exclude it based on our screening criteria. If screeners are uncertain about whether or not to include a particular article at this stage, the article will be shared among other team members, and consensus will be reached.

Step 3 – Study Appraisal and Data Extraction

For the appraisal process, each article will be read carefully by two reviewers, each assessing the *relevance* of the document to our inquiry (i.e., how much information can it contribute to our development of program theory?) and the *rigor* (i.e., whether that information was generated using credible and trustworthy methods).⁴³ Reviewers will use a tool designed to identify and record the following information: a) the program strategies that are described, b) what proximal outcomes are measured, c) how the proximal

outcome(s) is/are measured, d) the contextual factors that are mentioned in the article (e.g., if participants with addictions need to be in treatment or recovery prior to joining the program), e) whether the authors describe possible mechanisms that could lead to the outcome(s), and if so, what those mechanism are; and f) the study design/fit for purpose making a clear note if it seems to bias the results (e.g. if only successful program participants were included in the evaluation). Based on these findings, the reviews will give an overall impression of the richness of the data available from this article and how much it can contribute to our understanding of program theory.

We expect additional articles will be excluded after this in-depth review process if it is decided that they cannot contribute to our understanding of BIPs. If the two individuals who review a single article come to different conclusions, the larger team will discuss the issues and, if necessary, others will be asked to review the article(s) as well. Finally, we will comb through the citations of our articles as well as through our initial search results for additional articles or reports describing the same program and will review these sets or “families” of articles as a single unit.

As reviewers read and re-read these papers, particularly pertinent passages will be directly extracted and included in the spreadsheet, and other data summarized and annotated as necessary. The reviewers will meet on a regular basis to discuss their findings.

Step 4: Analysis and Synthesis

Having started with a hypothesized theoretical model for BIPs, this stage will involve examining the evidence gathered and determining whether it supports or

contradicts our proposed programmatic theory. Using our proximal outcomes as our organizing framework, we will look carefully at each evaluation that pertains to a particular outcome (or, if multiple publications describe the same program, we will look at these as a “family” of articles) and will assess how the data that were extracted from the studies inform our understanding of how batterer intervention works. Specifically, we will use the data to construct CMOs for each program. Because of the emphasis that many evaluators have placed on looking at final, rather than proximal, outcomes and at whether BIPs lead to reductions in recidivism/re-offense (instead of how they lead to them), we anticipate that the data describing mechanisms that underlay BIPs may be thin; thus, we will apply abductive reasoning as necessary to formulate our series of CMOs.²⁰ We will be particularly cognizant of the ways in which different contextual factors – when addressed – appear to influence the mechanisms that lead to these proximal outcomes.

Relying on the interdisciplinary perspectives and expertise of our research team, we will look at the information that arises from the construction of each CMO, as well as across programs to identify similarities and differences. Ultimately, we anticipate using the synthesis process to refine our original theoretical model in light of our review findings.

Step 5: Presentation and dissemination

The findings from this process will be presented in at least two formats: through at least one peer-reviewed article that conforms to the RAMESES publication standards put forth by Wong et al.⁴³ that is intended to inform implementation scientists and others in academic settings; and through targeted outreach and conversations intended to reach decision-makers and practitioners: program directors, policy makers, and community

1
2
3 coalitions charged with overseeing coordinated responses to partner violence. This will
4
5 include presentations at domestic violence and batterer treatment coalitions and
6
7 conferences and the preparation of plain-language reports and briefs for dissemination
8
9 through national and international networks.
10
11
12
13
14

15 DISCUSSION

16
17
18 This review was conceived after informal conversations with batterer intervention
19
20 treatment providers, domestic violence advocates, and judicial system personnel revealed
21
22 frustration over the current understanding of batterer intervention programs. Results
23
24 from evaluations and systematic reviews have been found to be conflicting, inconclusive,
25
26 and without substantial guidance about how approaches to batterer treatment could be
27
28 improved.^{2 44} In response, our team decided to complete this realist synthesis of
29
30 perpetrator treatment program evaluations with the aim of clarifying how and why BIPs
31
32 work for some men and the role that certain contextual factors may play in that success (or
33
34 lack of success).
35
36
37
38

39
40 This is the first realist review of the perpetrator treatment evaluation literature that
41
42 we know of, and we believe it will offer key insights into to the debate over how
43
44 communities can respond to partner violence. By focusing on the mechanisms that lead
45
46 participants to change (rather than only looking at whether or not change was achieved),
47
48 we believe we will provide much-needed insight into promising (and not-so-promising)
49
50 theoretically informed strategies that can lead to the types of changes that will allow men
51
52 to reduce their use of violent and/or controlling behaviors.
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

One of the key contributions of this review, in relation to the majority of BIP program evaluations and systematic reviews that have been done, is that we will focus on the impact that programs have on *proximal* outcomes, rather than on re-offending or recidivism. We believe this is important for several reasons. To truly understand *why* programs are or are not successful (and for whom and under what conditions), we need to gain a clear picture of the processes that lead to these outcomes. In the case of batterer treatment, it is unlikely that participants are magically transformed into nonviolent partners simply because they attended a program; rather, the program promotes certain outcomes within participants that then lead to these more distal behavioral changes. Identifying what these proximal outcomes are and how programs can achieve them is a key part of understanding this process.

This is not to say that we think the elimination of violent and controlling behaviors on the part of perpetrators should not be the ultimate goal of community responses; it most definitely should. We anticipate that one of the conclusions that may be drawn by this review is that significantly more work needs to be done to show how the achievement of these proximal outcomes ultimately can lead to the cessation of violence. However, as we look more closely at the literature surrounding BIPs, it has become apparent that factors unrelated to the programs themselves also contribute to the likeliness that men will cease to engage in violence against partners and family members. Influences at both the individual and interpersonal level as well as the community and social level are at play, and these need to be identified and accounted for within the larger coordinated response.

Authors' Contributions: AV conceptualized the study, reviewed the overview literature on batterer intervention and completed the initial program process and theoretical models. AV and PO led the design and drafting of the review protocol, assisted by RC, DF, and WD. AV and RC wrote the first draft of this manuscript, which was critically reviewed by PO, WD, and DF. All authors have approved submission of the manuscript to BMJ Open.

Acknowledgements: We extend our thanks to Pamela Dittman, staff to the Washington State Supreme Court Gender and Justice Commission, whose initial questions to the lead author sparked this work and who continues to provide ongoing review and input to ensure relevance to judicial systems across the U.S. Our appreciation also goes to Dr. Maritt Kirst who provided methodological advice during the early stages of protocol development, and to St. Michael's Hospital Information Specialist Carolyn Ziegler, who designed and conducted the search strategy.

Competing Interests: The authors declare no competing interests.

Funding Statement: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Gondolf EW. *Batterer Intervention Systems: Issues, Outcomes, and Recommendations*. Thousand Oaks, CA: Sage, 2002.
2. Carter LS. *Batterer Intervention: Doing the Work and Measuring the Progress*. A Report on the December 2009 Experts Roundtable. San Francisco: Family Violence Prevention Fund & National Institute of Justice, 2010.
3. Babcock JC, Green CE, Robie C. Does batterers' treatment work? A meta-analytic review of domestic violence treatment. *Clinical Psychology Review* 2004;**23**(8):1023-53.
4. Bowen E, Gilchrist E. Comprehensive Evaluation: A Holistic Approach to Evaluating Domestic Violence Offender Programmes. *International Journal of Offender Therapy and Comparative Criminology* 2004;**48**(2):215-34.
5. MacLeod G, Pi R, Smith D, et al. *Batterer intervention systems in California*. California, Judicial Council of California/Administrative Office of the Courts 2009.
6. Pepin D, Hoss L. *Menu of State Batterer Intervention Program Laws*. Atlanta, GA: Center for Disease Control and Prevention Office for State, Tribal, Local and Territorial Support, Public Health Law Program, 2015.
7. Maiuro RD, Eberle JA. State standards for domestic violence perpetrator treatment: current status, trends, and recommendations. *Violence Vict* 2008;**23**(2):133-55.
8. Miller M, Drake E, Nafziger M. *What works to reduce recidivism by domestic violence offenders*. Olympia, WA: Washington State Institute for Public Policy, 2013.
9. Dutton DG, Corvo K. Transforming a flawed policy: A call to revive psychology and science in domestic violence research and practice. *Aggression and Violent Behavior* 2006;**11**:457 – 83.

10. Gondolf EW. Theoretical and research support for the Duluth Model: A reply to Dutton and Corvo. *Aggression and Violent Behavior* 2007;**12**:664-57.
11. Dixon L, Archer J, Graham-Kevan N. Perpetrator programmes for partner violence: Are they based on ideology or evidence? *Legal and Criminological Psychology* 2012;**17**:196-215.
12. Gelles RJ. The politics of research: The use, abuse, and misuse of social science data—The cases of intimate partner violence. *Family Court Review* 2007;**45**(1):42-51.
13. Sayer A. Critical realism and the limits to critical social science. *Journal for the Theory of Social Behaviour* 1997;**27**(4):473-88.
14. Pawson R. *Evidence-Based Policy*. London: SAGE Publications Ltd, 2006.
15. Jackson S, Feder L, Forde DR, et al. Batterer Intervention Programs: Where do we go from here. Washington, D.C.: U.S. Department of Justice, Office of Justice Programs, 2003.
16. Stover CS, Meadows AL, Kaufman J. Interventions for Intimate Partner Violence: Review and Implications for Evidence-Based Practice. *Prof Psychol-Res Pr* 2009;**40**(3):223-33.
17. Wong G, Westhorp G, Pawson R, et al. Realist synthesis: RAMESES training materials: RAMESES Project, 2013.
18. Connelly JB. Evaluating complex public health interventions: theory, methods and scope of realist enquiry. *J Eval Clin Pract* 2007;**13**(6):935-41.
19. Pawson R, Greenhalgh T, Harvey G, et al. Realist Review - A New Method of Systematic Review Designed for Complex Policy Interventions. *Journal of Health Services Research & Policy* 2005;**10**(Suppl 1):21-34.

20. Jagosh J, Pluye P, Wong G, et al. Critical reflections on realist review: insights from customizing the methodology to the needs of participatory research assessment. *Research Synthesis Methods* 2014;**5**(2):131-41.

21. Sartin R, Hansen D, Huss M. Domestic violence treatment response and recidivism: A review and implications for the study of family violence. *Aggression and Violent Behavior* 2006;**11**(5):425-40.

22. Saunders D. Group Interventions for Men Who Batter: A Summary of Program Descriptions and Research. *Violence and Victims* 2008;**23**(2):156-72.

23. Paymar M, Garnes G. Countering Confusion about the Duluth Model. Duluth, MN: Domestic Abuse Intervention Programs, 2007.

24. Price BJ, Rosenbaum A. Batterer Intervention Programs: A Report From the Field. *Violence and Victims* 2009;**24**(6):757-70.

25. Mankowski ES, Haaken J, Silvergleid CS. Collateral Damage: An Analysis of the Achievements and Unintended Consequences of Batterer Intervention Programs and Discourse. *Journal of Family Violence* 2002;**17**(2):167-84.

26. Gondolf EW. How batterer program participants avoid reassault. *Violence Against Women* [HWWilson - SSA] 2000;**6**(11):1204.

27. Rosenbaum A, Leisring PA. Group intervention programs for batterers. *Journal of Aggression, Maltreatment and Trauma* 2001;**5**(2):57-71.

28. Gilchrist E. The Cognition of Domestic Abusers: Explanations, Evidence and Treatment. In: Gannon T, Ward T, Beech A, et al., eds. *Aggressive Offenders' Cognition: Theory, Research, and Practice*: John Wiley & Sons, 2007:247-66.

29. Holtzworth-Munroe A, Stuart GL. Typologies of male batterers: three subtypes and the differences among them. *Psychological bulletin* 1994;**116**(3):476.
30. Carlson RG, Jones KD. Continuum of conflict and control: A conceptualization of intimate partner violence typologies. *The Family Journal* 2010.
31. Gondolf EW. Who are those guys? Toward a behavioral typology of batterers. *Violence and victims* 1988;**3**(3):187-203.
32. Hamberger LK, Hastings JE. Personality correlates of men who abuse their partners: A cross-validation study. *Journal of family violence* 1986;**1**(4):323-41.
33. Johnson MP. *A Typology of Domestic Violence: Intimate Terrorism, Violent Resistance, and Situational Couple Violence*. Boston, MA: Northeastern University Press, 2008.
34. Buzawa ES, Buzawa CG, Stark E. *Responding to Domestic Violence: The Integration of Criminal Justice and Human Services*. Thousand Oaks, CA: Sage, 2017.
35. Stuart GL, Temple JR, Moore TM. Improving batterer intervention programs through theory-based research. *JAMA* 2007;**298**(5):560-62.
36. Day A, Chung D, O'Leary P, et al. Programs for men who perpetrate domestic violence: An examination of the issues underlying the effectiveness of intervention programs. *Journal of Family Violence* 2009;**24**(3):203-12.
37. Saunders DG. Feminist-cognitive-behavioral and process-psychodynamic treatments for men who batter: interaction of abuser traits and treatment models. *Violence Victims* 1996;**11**(4):393-414.
38. Bennett L, Violence PCAD. *Controversies and recent studies of batterer intervention program effectiveness*: National Resource Center on Domestic Violence VAWnet, 2001.

39. Diefenbeck C. Group therapy for male batterers: Comparison of cognitive-behavioral & object relations approaches. *Journal of Psychological Nursing & Mental Health* 2003;**41**(10):18-25.

40. MacLeod D, Pi R, Smith D, et al. Batterer Intervention Systems in California: An Evaluation. San Francisco, CA: Judicial Council of California, 2009.

41. Scott KL, Wolfe DA. Readiness to change as a predictor of outcome in batterer treatment. *J Consult Clin Psychol* 2003;**71**(5):879-89.

42. Melton H, Belknap J. He Hits, She Hits: Assessing Gender Differences and Similarities in Officially Reported Intimate Partner Violence. *Criminal Justice and Behavior* 2003;**30**(3):328-48.

43. Wong G, Greenhalgh T, Westhorp G, et al. RAMESES publication standards: realist syntheses. *BMC medicine* 2013;**11**(1):21.

44. Gondolf EW. Evaluating batterer counseling programs: A difficult task showing some effects and implications. *Aggression and Violent Behavior* 2004;**9**(6):605-31.

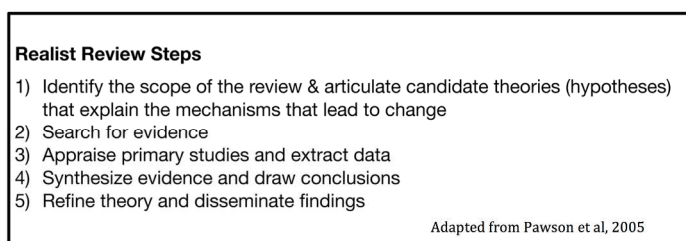


Figure 1: Key Steps in Realist Review

Figure 1: Key Steps in Realist Review
215x279mm (300 x 300 DPI)

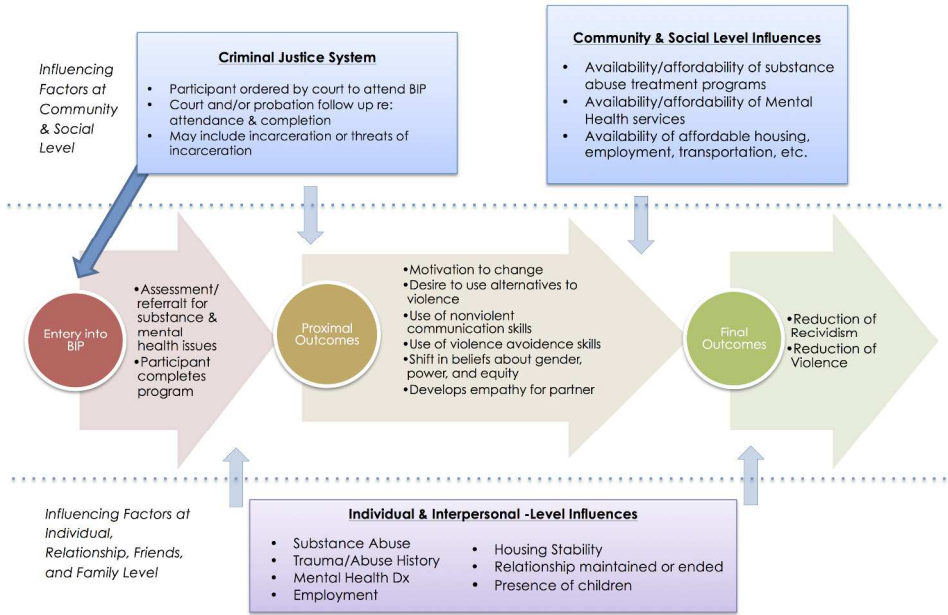


Figure 2: Batterer Intervention Process, Outcomes, and Influencing Factors

Figure 2: Batterer Intervention Process, Outcomes, and Influencing Factors
279x215mm (300 x 300 DPI)

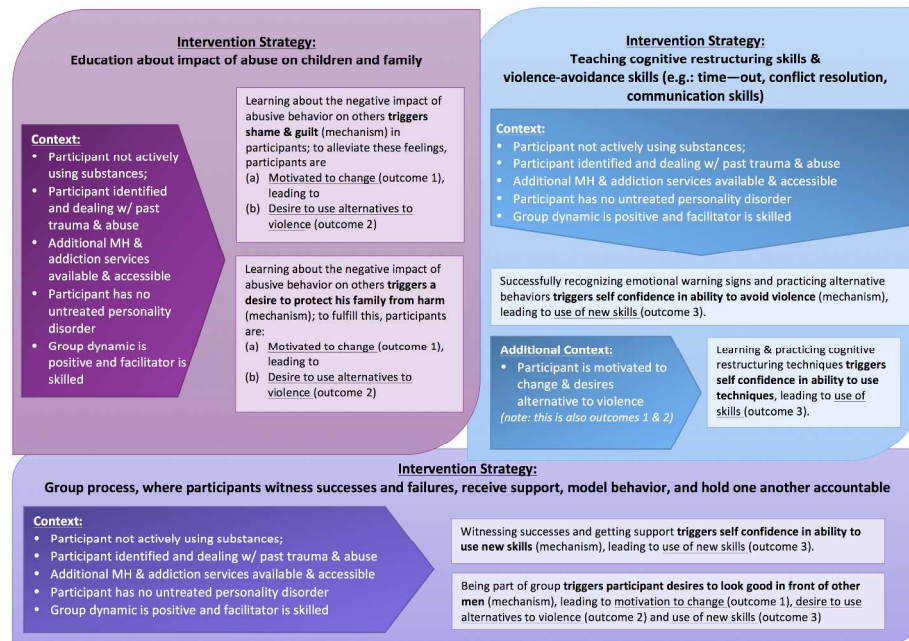


Figure 3: Sample CMO Configurations from the Preliminary Rough Theory of Batterer Intervention Treatment Programs

Figure 3: Sample CMO Configurations from the Preliminary Rough Theory of Batterer Intervention Treatment Programs

279x215mm (300 x 300 DPI)