

BUILDING EVIDENCE



DOMESTIC VIOLENCE

services & interventions

A FRAMING PAPER

January 2017

Submitted by

Center for Policy Research, in partnership with
National Resource Center on Domestic Violence



1. INTRODUCTION

To be most effective, policy, practice, and resource allocation should be informed by research and evaluation. **Yet, currently there is little rigorous evidence on the effectiveness of domestic violence (DV) program services.** As part of its effort to extend evidence-based practice, the Office of the Assistant Secretary for Planning and Evaluation (ASPE) contracted with Center for Policy Research (CPR), a Denver research and evaluation firm, and the National Resource Center on Domestic Violence (NRC DV) to develop this framing paper and organize the roundtable to which you have been invited.

The purpose of this framing paper is to set the stage for the roundtable and prepare participants to engage in conversations that will **generate innovative and concrete solutions** for the field. The roundtable is designed to facilitate the exchange of ideas and develop recommendations for conducting rigorous research to build the evidence for DV services.

To set the stage for a solution-driven roundtable, this framing paper:

1. Briefly **summarizes** and **critiques** the current state of **rigorous** evidence on core DV victim services
2. Identifies some **key challenges** that have limited the development of—and evidence base for—core services
3. Identifies a number of **strategies** that could assist in building the evidence base
4. Outlines some **areas of opportunity** for strengthening future research and evaluation

The FVPSA Program

The Family Violence Prevention and Services Program, located within the U.S. Department of Health and Human Services (HHS), administers the Family Violence Prevention and Services Act (FVPSA) funding nationally. FVPSA grants to states, territories, and tribes are awarded using a population-based formula. These formula funds are then provided as sub-awards to community-based programs.

Historically, core domestic violence services include emergency shelter, advocacy, counseling, safety planning, and support groups. However, since local programs are survivor- and community-driven and are not standardized, some programs use their FVPSA funding to offer a wider array of services including prevention and culturally specific programming that reflects the needs of survivors. The goal of programming is to promote both the immediate safety and the long-term social and emotional well-being of DV survivors and their children.

While FVPSA funding is important, it typically represents a relatively small portion of total program funding, and many DV programs rely on funding through multiple sources to provide services to victims and their families. Most local programs do not have the capacity to conduct program evaluations. While evaluation is not specifically prohibited, neither is it one of the stated statutory purposes of the FVPSA State and Tribal Formula Grant Program, which are dedicated to the provision of services. The National Domestic Violence Hotline and competitive national resource center grants also do not specify evaluation as an eligible activity. Although the FVPSA Program has

access to extremely limited discretionary funds, these have historically been directed to targeted competitive awards, technical assistance, and special projects designed to address pressing unmet services needs. Nevertheless, the FVPSA Program, like other HHS programs, is committed to supporting activities that are based on the best available evidence to maximize the likelihood that desired outcomes will be achieved through funded program services.

A Note about the Definition of “Core” Services

While the roundtable focuses primarily on the FVPSA-funded core services—**shelter, counseling, advocacy, support groups, and/or safety planning**—they are not the only services relevant for future evaluations. Several respondents interviewed for this framing paper (see section Appendix D for details about these interviews) said **it is important to consider additional services** in evaluations of program efficacy in the DV field. They note that some organizations, especially culturally specific programs, provide other services that do not fit these categories. These programs may respond to survivors' needs from a holistic perspective and engage in activities that aim to foster community engagement and/or build community capacity. Failure to include these efforts in evaluations provides an incomplete picture of program activities and precludes the ability to generate information on promising prevention or intervention approaches.

Importance of Building a Strong Evidence Base for Core DV Victim Services

Many activities in DV programs have been designed based on practitioner and survivor expertise and non-experimental studies. There are many different types of evaluations that can make up an evidence base. See the Center for Disease Control and Prevention's *Continuum of Evidence of Effectiveness* in Appendix A for more information about the different factors that contribute to evidence and some of the terms used throughout this framing paper.

While practice-based evidence and program evaluations provide useful information to the field, **methodologically rigorous research is another way of giving voice to the real experience of survivors and practitioners, and is needed to definitively indicate what works, for whom, and under what conditions.** For example, a systematic investigation of behaviorally informed messages might reduce the number of hang-ups by callers and improve the effectiveness of hotlines. Similarly, a rigorous empirical assessment of support group interventions could produce reliable guidance on the number of sessions and/or the curricula that are most effective at achieving desired outcomes.

Rigorous research has impacted policy and practice for DV programs and other areas. As one example, The Danger Assessment is an evidence-informed tool that is widely used by programs to measure risk in an abusive relationship and develop appropriate safety plans and other interventions (Campbell et al., 2008). The strong evidence that home visits by nurses positively impact a specific population of new mothers and their children (Olds et al., 2015; Olds et al., 2014) has resulted in widespread dissemination of home

visitation interventions and funding for Nurse-Family Partnership and other home visitation programs in the Affordable Care Act.

Local DV programs, state coalitions, funders, and policy makers are all committed to seeing that services are as beneficial as possible in helping DV survivors and their children improve their overall well-being. A strong body of evidence is critical to achieving this goal. Currently, however, the evidence base for DV services is limited.

Given the great diversity of individuals who receive services and the variety of contexts through which services can be provided, it is critical to support research that attempts to understand what works, and for whom, under what conditions. We are excited that the upcoming expert roundtable will result in concrete recommendations for strengthening the evidence base behind DV services, and in turn, strengthening the impact of such services for DV survivors and their children.

2. SYSTEMATIC REVIEW

To inform our roundtable discussions, we included a summary of the rigorous evidence that currently exists regarding DV services. The review located empirical articles examining the impact of advocacy, counseling, safety planning, shelter, and support group services (core services) on outcomes for adult survivors of DV. The inclusion criteria are below (see Appendix B for the methodology).

Inclusion criteria:

- Experimental or quasi-experimental design that included a comparison or control group to examine program impact
- The program or service specifically targeted adult survivors of intimate partner violence (IPV)
- The service was provided within or in collaboration with a DV program
- The study examined one or more health or psychosocial outcomes, including psychological, emotional, behavioral, and social factors

Limited Research Exists Linking DV Program Activities to Desired Outcomes

Fifteen studies of DV services met the criteria for inclusion: four for advocacy, seven for counseling, two for safety planning, two for support groups, and none for shelter (see Appendix C for the results of each study, organized by service type).

The terms domestic violence (DV) and intimate partner violence (IPV) are used interchangeably in this paper.

Shelter

No studies of shelter services met the criteria for inclusion because of the serious safety and ethical concerns of not providing shelter services in a control or a comparison group. Comparisons of survivors who use and do not use shelters are limited by the fact that these two groups of women differ on many variables other than shelter use (e.g., income level, education level, access to other options, severity of abuse).

Advocacy

The advocacy studies indicate that advocacy services that are offered for approximately 10 weeks in survivors' natural communities result in decreased risk of re-abuse, as well as increased access to community resources, higher social support, and improved mental health and well-being for survivors over time.

Counseling

The seven studies of counseling treatments designed for IPV survivors show that they hold promise for reducing depression and post-traumatic stress symptomatology. Evidence suggests that helpful components may include: (a) psycho-education about the causes and consequences of IPV; (b) attention to ongoing safety concerns; and (c) a focus on survivors' strengths.

Because the seven interventions differed from one another on whether they were offered in group settings or individually, the number of sessions offered, counselor education and training, and curriculum content, it is premature to determine if there are specific components that might be essential for all survivors, beneficial to some survivors, or even irrelevant to treatment success.

15 studies were included in the systematic review

Studies showed **positive outcomes**; however, there were **methodological limitations** across studies.

7 counseling

decreased depression symptoms
decreased PTSD symptoms

HOWEVER...

all different in terms of structure, content,
and counselor education & training

4 advocacy

less difficulty accessing resources
increased social support
less violence over time
higher quality of life

HOWEVER...

some studies had small sample sizes
or other methodological limitations
that limit generalizability

2 support groups

reduced psychological distress
increase in feeling supported
increase in healthcare utilization

HOWEVER...

there are serious methodological
weaknesses of these studies

2 safety planning

increased safety planning

HOWEVER...

they did not assess actual
safety over time

MAJOR METHODOLOGICAL LIMITATIONS ACROSS STUDIES

Small **sample sizes**

Lack of adequate representation of
people from different cultural
backgrounds, various geographic
areas, and people of color

Brief **follow-up** time frames

Designs that failed to account
for **confounding variables**

Measurement concerns

Safety Planning

Two studies examined the effectiveness of promoting safety strategies with IPV survivors. One used repeated telephone calls, while the other used a mobile app, but both methods resulted in increased safety planning. Neither study, however, linked this new behavior to whether survivors were, in fact, safer over time as a result.

Support Groups

One of the two studies of support groups for survivors of IPV reported significant changes for support group members with regard to psychological distress, feeling supported, and healthcare utilization. Methodological weaknesses of the study, however, limit confidence in the findings.

Limitations of the Included Studies

Although the 15 studies included in this review used experimental or quasi-experimental designs, **most had one or more serious limitations**. There is a paucity of evidence on the effectiveness of program services for Latinas and Native American women. Some studies lack any follow-up assessment or are limited to a few weeks. Other studies are unable to differentiate between the effects of several simultaneous interventions. Across studies, methodological limitations include: (a) small sample sizes; (b) lack of adequate representation of people from different cultural backgrounds, various geographic areas, and people of color; (c) brief follow-up time frames; (d) designs that failed to account for confounding variables; and (e) measurement concerns. As a result, even the most rigorous evaluations provide less than optimal guidance for programs.

The Body of Evidence

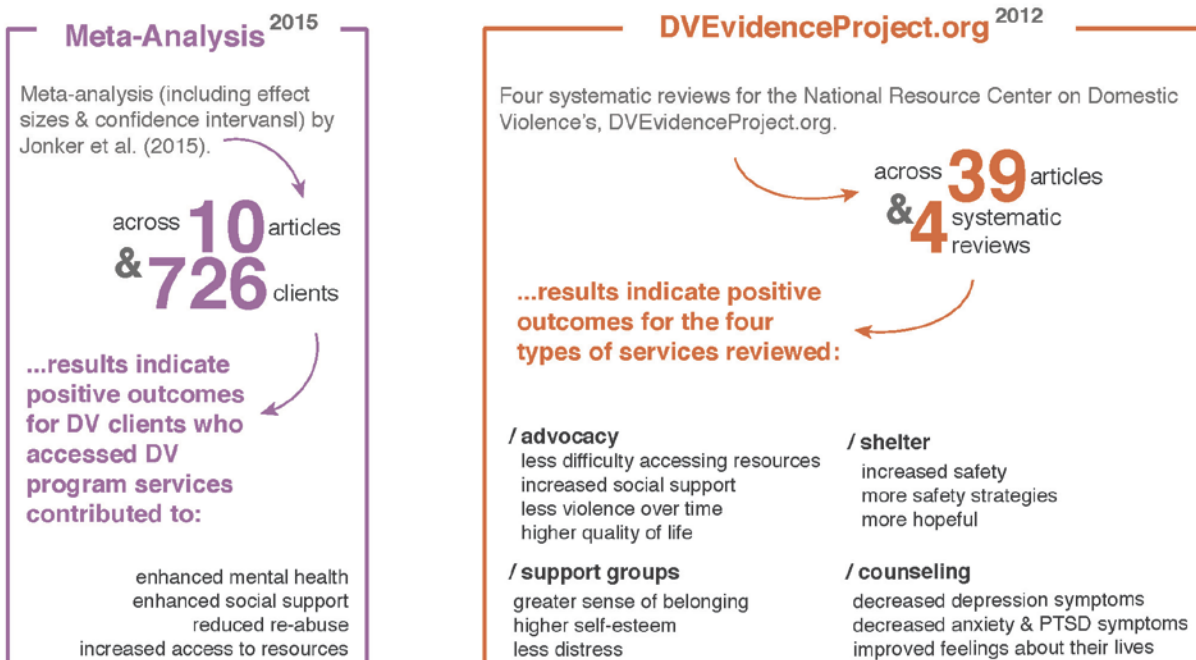
This is not to say that the studies fail to provide us with useful information. **In fact, these studies share a common theme: they each found that DV services and interventions contribute to enhanced survivor well-being, which is ultimately the goal.** Taken together, these studies establish the value of DV programs and services. And even if many of the causal connections are diffuse, intertwined, and otherwise difficult to establish, we must not give up on assessing effectiveness.

Indeed, other evidence reviews that were more inclusive in their methodological criteria than this review found similar patterns about the positive impact that services have (see DVEvidenceProject.org for these reviews). See below for a brief summary of these reviews, as well as a recent meta-analysis of DV services.

Taken together, this systematic review, combined with other evidence reviews, indicate that DV programs are heading in the right direction and achieving promising results. What the field needs now is more rigorous evidence to understand why (e.g., the specific components), for whom, and under what conditions these services and interventions work.

Taken Together, Multiple Evidence Reviews Indicate DV Services & Interventions Enhance Survivors' Social & Emotional Well-being

In addition to the systematic review conducted for this framing paper, 5 other systematic reviews also demonstrate the positive impact of DV services & interventions.



For the review articles, please see Sullivan (2012a-c); Sullivan, Warshaw, & Rivera (2013); and Jonker et al. (2015); full citations provided in Appendix E.

3. CHALLENGES TO BUILDING THE EVIDENCE BASE

There are many challenges to conducting rigorous and relevant evaluations of DV services. Some are unique to the DV field, while others are widespread across the study of human services provision with people who are low-income, transient, and/or have experienced trauma. Some of the most salient challenges are provided next.

Serious Safety Concerns

The dynamics of domestic violence contribute to **unique safety concerns** for participants, their children, project staff, and researchers. As part of the dynamics of abuse, one partner actively seeks to control their partner or ex-partner, often using a combination of physically and sexually violent, psychological, economic, social, and legal tactics. Clients' and their children's lives can literally be in danger from an individual who has intimate knowledge of and easy access to them.

Relatedly, it may be inappropriate or unsafe to recruit clients into an evaluation and/or collect data when survivors first access services and are in crisis. Many DV clients access services (e.g., shelter) at the time they and their children are also most likely to be murdered by their abusive partner—when trying to leave.

For more information about the challenges involved with evaluating DV services, please review Sullivan, 2011; Sullivan & Cain, 2004; and the World Health Organization, 2001 (See Appendix E for citations).

Thus, DV research entails some issues that go **beyond the protections normally addressed in IRB-approved studies**. In addition to aggregating and de-identifying data, separating and guarding identifying information, and destroying audiotapes, DV researchers may, for example, need to consider and discuss with research participants whether survivors' responses are protected from subpoena by a perpetrator's attorney. As in all research settings, DV researchers also need to reveal that disclosures of child abuse or neglect and/or threats to harm oneself or another person must be reported to the authorities.

Lack of Practitioner-Researcher Trust and Relationships

A lack of trust between practitioners and researchers is particularly salient in the DV field.

Researchers strive to conduct the most rigorous studies that are feasible to generate statistically valid, unambiguous results. Practitioners are understandably concerned that researchers may design studies that reflect an inadequate understanding of how services are structured or what meaningful service outcomes look like. They are also concerned that the process of data collection might be re-traumatizing or coercive for survivors. When researchers do not share the same cultural understanding as program staff or survivors—especially if they do not recognize this as a concern—developing relationships and trust is difficult. A relationship built on trust, shared power, and transparency needs to be developed over time and among the researcher and the agency's director,

frontline staff, and survivors. This level of engagement and collaboration requires an extensive time commitment and dedication to shared learning from both researchers and community partners.

Statutory Safeguards on Data Protection and Safety

To protect the safety and confidentiality of survivors, **DV programs are prohibited from disclosing identifiable information about their clients** by the Violence Against Women Act and the Family Violence Prevention and Services Act. While important to victim protection, this policy means that large-scale databases from which researchers might gather secondary data for analysis neither exist nor can be generated at the state or national level. Nor can researchers use program data to link with other agency data to assess long-term outcomes in health, criminal or civil justice, or employment earnings.

Limited Funding Structures Dedicated to Services Research

The FVPSA Program is able to use 2.5% of its annual \$150 million budget for administration, evaluation, and monitoring, which includes overseeing the 1,600 local programs it funds to deliver core program services. Local programs receive an average award of \$50,000 per year limiting resources to support program evaluation.

There are no federal funding streams dedicated to examining the broad impact of DV services on survivor

safety and well-being. **DV research and evaluation funding is spread across multiple agencies and programs that each have specific outcomes of interest.** For example, the National Institute of Justice (NIJ) funds studies with a criminal and legal focus, and specific Institutes within NIH, such as the National Institute of Drug Abuse (NIDA) fund studies focused on the intersections of substance use and violence against women.

The designation of separate agencies for research and practice reinforces “silos” that often preclude the generation of more holistic, practice-oriented research and collaborations between researchers and practitioners. The division of research funding across multiple agencies, institutes, and centers reinforces the mission of the various funding entities without

The designation of separate agencies
for **research & practice**
reinforces
“silos”

necessarily promoting the examination of DV services’ impact on survivor safety and well-being.

Extensive Training Requirements for Research Staff

Evaluating DV program services requires staff training and support over and above what is typically provided when conducting research.

This includes training on the **dynamics** of DV and identifying and rectifying staff **biases** and **stereotypes**. Research staff need to understand **safety planning** strategies both to discuss with project participants, as necessary, as well as for themselves. The **emotional toll** and potentially

triggering nature of this type of research also need to be addressed with research staff, and with proper support provided.

Other Methodological Challenges

There are many methodological challenges to conducting rigorous research and evaluation unique to DV. There are also challenges for economically vulnerable, transient, and traumatized populations more generally (Schorr & Farrow, 2011). Building the rigorous evidence requires us to account for these challenges as well as the challenges most salient to DV settings, as described above.

- **Individualized services** based on survivors' needs and life experiences make it difficult to compare services across clients and programs. Because programs often tailor interventions to clients' individual circumstances, treatments are unique and dosages unstandardized
- Program **attrition** is high among low-income, traumatized, and marginalized populations, who are often geographically mobile and/or sporadically homelessness. Some also distrust research in general or have safety concerns that make them hesitant to engage in research. These issues make it difficult to engage and retain such groups in interventions and evaluation, and makes retention efforts expensive and time consuming
- **Longitudinal research** is expensive, time consuming, and requires specialized training so survivors are neither endangered nor re-traumatized by their participation
- **Randomizing** clients into a no-treatment control group can be **ethically and practically unfeasible** with populations that require immediate services, safety, or housing,
- **Small program size** makes it difficult to detect significant differences, especially for traditionally marginalized groups such as persons of color, adolescents, elders, lesbians and gay men, trans survivors, immigrants and refugees, and those with disabilities or multiple needs
- **Unclear service end dates** make it difficult to time the accurate collection of post-program data
- **Treatment heterogeneity** due to differences in curriculum, dosage, facilitator, and other features makes it difficult to compare similarly named interventions

4. ADDRESSING THE CHALLENGES TO BUILDING THE EVIDENCE BASE

Addressing the challenges described in the prior section will require innovative strategies, and the roundtable has been designed to foster more critical thinking in this area. This section includes initial ideas for addressing these issues, generated from our own experiences, the expertise shared through seven key informant interviews (see Appendix D for list of key informants), and prior published work.

Beginning with an Evidence-Supported Theory of Change

Before researching whether a DV intervention has resulted in positive change for a survivor, the first step is to explicate what the intervention was intended and can reasonably be expected to accomplish. This requires articulating the **theory of change** that guides services. A theory of change is an empirically justified explanation of how and why one expects a desired change to occur. It involves identifying the desired long-term objectives (e.g., what are we hoping to accomplish?), and then working backwards to identify how to achieve measurable outcomes tied to the goals (e.g., how do we get there?).

A theory of change for DV services was developed in 2012 by numerous experts, including practitioners, advocates, survivors, funders, researchers, and policy makers. It first notes that the long-term objective of domestic violence programs is to **enhance survivors' and their children's well-being**. There is ample empirical evidence demonstrating that social and emotional well-being is impacted by: (a) self-efficacy; (b) hopefulness; (c) social connectedness; (d) safety;

(e) having adequate social and economic opportunities; (f) economic stability; (g) enhanced justice; and (h) good physical, emotional, and spiritual health. DV programs are invested in impacting these eight factors for survivors and their children through efforts targeting multiple levels of change. Their work is designed to create **intrapersonal** change in survivors' thinking, belief systems, and emotions, and to create the **interpersonal** and social changes necessary for the outcomes to be achieved. This theory of change is presented on the next page.

Addressing the Methodological Challenges

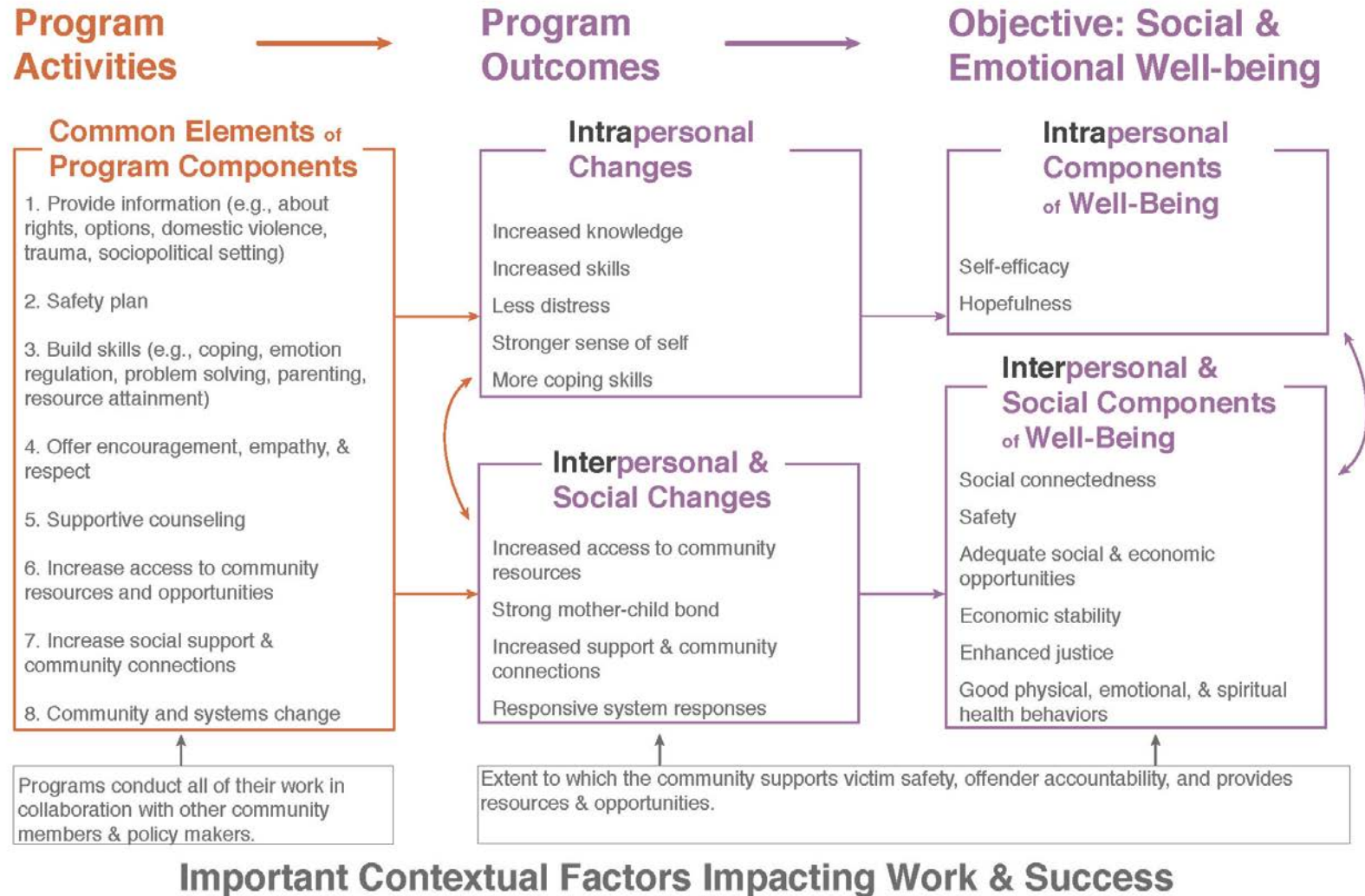
The next generation of efforts to reliably measure outcomes of DV program services should utilize approaches that address methodological challenges for the field. Some key challenges include:

- Lack of diversity in DV survivors included in research and evaluation studies
- Inability to randomize certain services and identify appropriate comparison groups
- Limited feasibility of implementing longitudinal designs
- Variation in the amounts and types of services that clients receive*

Some strategies that address these key challenges are presented next. However, **our goal is to have a creative, solution-focused discussion about addressing these challenges at the roundtable.**

* For example, *shelter only vs. shelter and counseling*

Theory of Change Underlying How Domestic Violence Program Activities Impact Adult & Child Survivors' Well-Being



Sullivan, C.M. (2016, January). Examining the Work of Domestic Violence Programs Within a "Social and Emotional Well-Being Promotion" Conceptual Framework. Harrisburg, PA: National Resource Center on Domestic Violence at <http://www.dvevidenceproject.org>

Challenge

Possible Method to Address the Challenge

Lack of Diversity in Study Samples

- Use electronic devices (e.g., tablets) that are faster, easier, and of higher quality than paper surveys; allow for easier language translation; and can include an audio component (e.g., audio computer-assisted self-interviews).
- Prioritize evaluations of culturally specific services, particularly by those who are from the same communities being served.
- Consider mixed method designs that use indigenous research methodology (e.g., sharing circles), which prioritizes generating and disseminating knowledge that will be culturally accurate and meaningful to Indigenous communities.
- Implement community-based research collaborations that engage survivors and practitioners in identification of critical practice-generated questions, development of relevant outcome measures, and the design of safe research and data collection approaches

Inability to Randomize a Service

- Consider adaptive randomized designs that include providing research participants with choices throughout the research process
- Implement sequential multiple assignment randomized trials (SMART), which rigorously assess the sequencing and delivery of intervention components and dosage variations but do not deny treatment
- Employ rigorous quasi-experimental designs that control for important, potentially confounding factors
- Consider a factorial design that investigates intervention components, such as varying number of sessions or curriculum modules

When Longitudinal Designs Are Not Feasible

- Use Life History Calendar, a data collection tool designed to provide memory cues that promote more accurate retrieval of prior life events. It can be used retrospectively to identify temporal and causal relationships

Different Amounts and Types of Services

- Record interactions between staff and clients and code them for quality and quantity to understand variability in what clients receive
- Incorporate fidelity measurement in the design and analyses
- Conduct rigorous, in-depth case studies of promising practices
- Use naturalistic, cohort designs to examine the match among need, services received, and impact of the services
- Conduct cross-site evaluations that could identify key program characteristics
- Evaluate the components within interventions that hold the most promise for generating positive change

5. AREAS OF OPPORTUNITY FOR BUILDING THE EVIDENCE

HHS is interested in generating stronger evidence in the DV field about effective practice and promoting the use of evidence-based services among programs. **We discussed possible areas of opportunity in interviews with seven key informants** (See Appendix D for information about these interviews). We also reviewed materials prepared by the Coalition for Evidence-Based Policy, which tracks the characteristics of evidence-based initiatives in six specific areas of social intervention, including home visiting and teen pregnancy prevention. This section includes a summary of the information gathered from these sources about areas of opportunity to build stronger evidence. **The ideas presented here are a starting point to spark discussion among roundtable participants.**

As noted earlier, community-based domestic violence programs receive critical, but limited, funding under FVPSA and therefore have limited capacity and resources to conduct program evaluations. At the federal level, the FVPSA Program is authorized under that statute to coordinate domestic violence research initiatives within HHS and with other federal agencies and researchers, and has provided important leadership in efforts to build the evidence base for domestic violence services. However, given the current FVPSA funding framework, many of the opportunities and innovative approaches highlighted in this section will not be possible to implement without expanded capacity to provide grants, additional funding, and new federal partnerships.

Cultivate & Train Researchers to Engage in DV Research and Be Responsive to the Logistical, Ethical, & Cultural Issues of DV

With its emphasis on publishing in top-tier journals and obtaining extramural research funds, the academic reward system is not particularly compatible with doing evaluation research, especially in the DV field. Moving beyond the difficulty of attracting researchers to the field, there is the challenge of training them well. Several respondents spoke about the **importance of researchers understanding and being responsive to the many logistical, ethical, cultural, and linguistic issues that come into play when engaging survivors in research dealing with DV services.** Indeed, one respondent suggested creating an “Angie’s List” of vetted researchers. More typically, stakeholders suggested that initiatives be developed to engage, train, mentor, and support researchers in the DV field.

Build Collaborations Between Researchers & Programs

Getting researchers and programs to know and trust one another is critical to effective evaluations and building rigorous evidence. Overworked and underpaid program staff sometimes resent the resources and time that research requires, have had negative

experiences with “drive-by data collectors,” fear that survivors will be triggered by data collection activities, and worry about the impact of inconclusive or negative findings on program funding. External requirements to evaluate often overshadow consideration of the benefits of evaluation.

Meaningful and respectful collaboration with programs throughout the evaluation process is a critical component of an effective evidence-building strategy.

Practitioners **must** be involved in identifying research questions & priorities, creating the outcome measures, and selecting the research design & data collection strategies

Simply put, practitioners must be involved in identifying the research questions and priorities that the field needs to address, creating the outcome measures that reflect client and program goals, and selecting the research design and data collection strategies. Researchers can earn respect by working with practitioners to identify ways to conduct research while protecting survivors. Adequate time should be allocated for such project planning and relationship-building activities in future funding announcements.

Spotlight: DVPERC

One example of this approach is the Domestic Violence Program Evaluation and Research Collaborative (DVPERC), which consists of representatives from approximately 25 DV programs and researchers in the

New England area who have met regularly for six years to identify research questions of mutual interest, share the results of relevant research, and conduct needed research projects in the field. To date, the group has developed and published measures of survivor-defined advocacy, safety-related empowerment, and trauma-informed practices that are gaining traction nationally (See [DVEvidenceProject.org](https://dvevidenceproject.org) for examples).

Spotlight: Federal Initiatives

Other examples of community-based participatory approaches that are supported by the federal government are exemplified in the many research centers ACF has created to work with marginalized populations, including Native and Tribal and Latino/a communities. One example of this is the Tribal Early Childhood Research Center, which collaborated with tribal programs to foster a community of learning and support, conduct research and measurement development, and promote evaluation and research-to-practice activities. Other ACF center grants to build rigorous evidence in other fields include the National Center on Research on Hispanic Children and Families and the Center for Early Care and Education: Dual Language Learners (CECER-DLL).

Generate a Reliable and Adequate Support for Evaluation Research on DV Services

There is **no statutory set-aside for evaluation research on domestic violence**. Research on domestic violence is scattered across multiple agencies, with allocation levels and priorities that vary

from year to year, and no funding dedicated to evaluation.

Spotlight: Agency Coordination

One approach to supporting DV services research is to promote collaboration and coordination across multiple, relevant federal funding agencies. For example, in 2012, NIJ and NSF signed a Memorandum of Understanding to reduce crime and promote justice by giving both agencies flexibility to sponsor research and evaluations in similar subject areas.

In late 2015, three federal agencies (the Department of Justice, the Department of Housing and Urban Development, and the Department of Health and Human Services) agreed to pool resources and expertise to launch a \$2.3 million federal Domestic Violence and Housing Technical Assistance Consortium to provide national training, technical assistance, and resource development on domestic violence and housing. As another example, FVPSA and other offices within ACF have collaborated to study DV hotlines and to assess the evidence on DV screening in programs offering marriage and relationship education.

Sponsor an Array of Research

ACF's evaluation policy (2012) stresses use of the best scientific methods that are appropriate and feasible in all evaluation activities. This policy also acknowledges the value of multiple types of evidence. Thus, in addition to impact evaluations that use random assignment techniques, ACF recognizes the benefits of multiple types of evidence including high-quality descriptive studies, performance measures, qualitative studies, financial and cost data, survey statistics, program administrative data, and performance management data. Future funding announcements in

the DV area might mirror grant-making initiatives in other human services areas that reflect a commitment both to **rigor and methodological heterogeneity**. An array of research and evaluation efforts that have been used in other arenas are described next.

Cross-Site, Random Assignment

Cross-site randomization studies can be implemented when study participants are offered one of two or more acceptable interventions. The HUD-sponsored Family Options Study, for example, used a random assignment strategy that compared three enhanced treatment strategies (rapid rehousing, housing subsidy, or transitional housing) with services as usual to assess the relative benefits of various interventions to homeless families across 12 sites.

To address the migration of participants from one treatment group to another, several National Institute of Health studies have promoted **adaptive intervention designs** such as Sequential Multiple Assignment Randomized Trials (SMART; see footnote).

Common Program Performance and Outcome Measures

When full-control, multi-site projects are not feasible, some aggregation across multiple studies conducted by local evaluators can be achieved by having programs use a limited number of core outcome measures. For example, part of the ACF-funded third-party evaluation portfolio of Health Profession Opportunity Grants (HPOG) includes uniform data measures across all 32 grantees on participant characteristics, outputs and outcomes. Data are collected at the grantee, program, and individual participant levels and input into a management information system (MIS). The MIS also automatically generates the performance progress reports grantees must submit to the Office of Family Assistance and has robust reporting functionality to allow grantees to understand and use their own data.

The 2015 Healthy Marriage and Responsible Fatherhood Grants (HMRG) awarded by ACF's Office of Family Assistance permit tracking of results across different programmatic approaches, settings, and populations.

As part of the Fatherhood and Marriage Local Evaluation and Cross-Site Project (FaMLE), Mathematica Policy Research developed a common set of instruments and a management information system for all 90 grantees to use. In addition to evaluation activities by local evaluators, Mathematica will ultimately aggregate the grantee-specific data on program operations, performance, and outcomes to yield both program-specific and big-picture results.

In a SMART design, there is a separate stage for each of the critical decisions involved in the adaptive intervention. At each stage, all participants are randomly assigned to a treatment option. By randomizing participants multiple times, one can test the effectiveness of each stage and treatment option, including dose and timing of treatments (Lei et al., 2012).

Project Spotlight: FRPN

The Fatherhood Research and Practice Network (FRPN) seeks to build rigorous evidence in the fatherhood field by building the evaluation capacity of programs and researchers and sponsoring evaluations conducted by researchers who collaborate with local programs. FRPN is funded through a cooperative agreement between OPRE and Temple University/Center for Policy Research. FRPN components include:

- Solicit proposals from researcher-practitioner teams to conduct evaluations
- Fund rigorous evaluations and systematic descriptive studies that fill research gaps
- Mentor selected researchers and practitioners through monthly check-in calls
- Conduct quarterly webinars on evaluation issues and ways to improve participant recruitment, retention, and data collection
- Conduct day-long trainings on research at national and regional practitioner conferences leading to a certificate from Temple University and continuing education credits in social work
- Develop and validate new measures in the fatherhood field and encourage their use by FRPN grantees and other fatherhood evaluators
- Create and update frpn.org to connect interested practitioners and evaluators with one another, and provide access to fatherhood evaluation resources and new literature
- Provide technical assistance to other fatherhood researchers, including local evaluations funded through the 2015 Health Marriage and Responsible Fatherhood Grants

Local Evaluations

Local evaluations conducted by trusted research partners afford the opportunity to investigate unique, **program-generated questions and culturally specific populations and interventions**. Numerous funding initiatives support these types of investigations. For example, nearly half of the 90 HMRF grants are also implementing evaluations (some grantee-led, some federally led), with designs ranging from descriptive studies to randomized control trials. In both initiatives, grant funds could be used for evaluation.

Evaluation Technical Assistance

Investing in Innovation (i3) is another example of how evaluation can be conducted at the local level and supported through technical assistance, if funding is provided to support local evaluators and a cross-site technical assistance provider. The i3 grantees are required to conduct an independent evaluation and must agree, along with their independent evaluator, to cooperate with the evaluation technical assistance provided by the Department of Education and its evaluator, Abt Associates, Inc. The purpose of the technical assistance is to **maximize the strength of the impact studies and the quality of their data and performance feedback**. Depending on the level of prior evidence supporting the intervention, the design of the i3 projects and their evaluations vary and range from innovative studies, validation studies, or scale-up efforts.

Small, Time-Limited Evaluations

While there is a strong need for large-scale, longitudinal studies that examine what works, for whom, and under what conditions in the DV area, there is also growing recognition of the **value of smaller, time-limited evaluations** that are methodologically rigorous (e.g., Rapid Cycle Evaluations; see note).

Rapid cycle evaluations (also known as usability testing) are rigorous, short-term, and low-cost assessments of critical components of an intervention to find out what works. The short-term framework allows for tests of different iterations of the intervention and different methods of implementation to generate causally valid impact estimates within a year (Cody et al., 2015).

Solicitation Structures

Funding initiatives that will build the evidence base about DV program services **include thoughtfully crafted announcements regarding funding opportunities**. One example is a recent NIJ solicitation to fund measurement research related to teen dating violence. Another area of opportunity is the development of solicitations that invite:

- The development of evaluation toolkits for culturally specific populations and programs
- The creation of community-based and survivor-informed research advisory boards to provide input and technical assistance to researchers and funders
- The provision of evaluation technical assistance to programs
- The creation of networks to facilitate relationship building between researchers and practitioners
- Initiatives to develop and validate new measurement tools that can be used to evaluate DV services

6. SUMMARY & NEXT STEPS

This purpose of this framing paper is to set the stage for upcoming conversations and idea generation at the roundtable in April 2016. It is expected that the roundtable will result in concrete and innovative recommendations for building the evidence base for DV services. To that end, this paper described the current state of the evidence for DV services and interventions, presented the most salient obstacles to building the evidence base for services designed for DV populations, and provided some preliminary strategies for addressing those obstacles. It concluded by presenting a number of areas of opportunity that either exist or could be created to build a strong evidence base for DV services.

By convening intervention researchers—in and out of the DV field—DV practitioners, policy makers and funders, we expect to generate concrete suggestions for the necessary next steps. Key feedback we hope to obtain includes: (a) major evaluation gaps; (b) questions the field should address to move the field forward; (c) the most rigorous research strategies that are feasible and appropriate to answer targeted evaluation questions; (d) how to most effectively sequence future research; and (e) the actions that ASPE and other HHS entities can take to further the goals of generating reliable evidence that is used to inform practice. The results of the roundtable will be distilled and incorporated in a report to ASPE on areas of opportunity that will help to inform the agency in policy decisions and resource allocation.

Key feedback we hope to obtain during the roundtable includes:

- (a) major evaluation gaps
- (b) questions the field should address to move the field forward
- (c) the most rigorous research strategies that are feasible and appropriate to answer targeted evaluation questions
- (d) how to most effectively sequence future research
- (e) the actions that ASPE and other HHS entities can take to further the goals of generating reliable evidence that is used to inform practice

APPENDICES

Table of Contents

Appendix A: A Continuum of Evidence

Appendix B: Systematic Literature Review Method

Appendix C: Empirical Evidence Tables: Results from the Systematic Review

Appendix D: Key Informant Interviews Method

Appendix E: References

Appendix A: A Continuum of Evidence

We encourage readers to review the entire Centers for Disease Control and Prevention (CDC) Guide to the Continuum of Evidence of Effectiveness. A graphic of the Continuum of Evidence from this report is on the **next page** for quick reference. There are many different types of evidence that can make up an evidence base. Each type of evidence has a unique contribution to the overall body of evidence, as well as strengths and limitations (Puddy et al., 2011; Schorr & Farrow, 2011). For example, practitioner experience and high quality qualitative studies provide useful information on the experiences of individuals and programs. While sometimes they are representative of a larger group and can be generalized to broader populations, they often have small sample sizes. Evaluation studies that incorporate quantitative data may have larger sample sizes, however other limitations (e.g., the lack of a comparison group) may also limit the ability to generalize to other populations. In all, each research design and data type have strengths and limitations as well as different quality standards. For the purposes of this paper, we focused on studies that met some of these quality standards (i.e., randomized control trials or quasi-experiments that included a comparison group). In the paper, when we refer to building the rigorous evidence base, however, we are referring to high quality research and evaluation (i.e., the “well-supported and supported” category on the CDC’s continuum of evidence).

Continuum of Evidence of Effectiveness

	Well Supported		Promising Direction / Emerging / Undetermined More Research Needed			Unsupported	Harmful
Effect	Found to be effective		Some evidence of effectiveness	Expected preventive effect	Effect is undetermined	Ineffective	Practice constitutes risk of harm
Internal validity	True experimental design	Quasi experimental design	Non-experimental design	Sound theory only	No research No sound theory	True or quasi experimental design	Any design with results indicating negative effect
Type of evidence/ research design	Randomized control trials and meta-analysis / systematic review	Quasi experimental design	Single group design	Exploratory study	Anecdotal / Needs assessment	Randomized control trials or quasi experimental design	Any design with results indicating negative effect
Independent replication	Program replication with evaluation replication		Program replication without evaluation replication	Partial program replication without evaluation replication		Program replication with evaluation replication	Possible program replication with/without evaluation replication
Implement- ation guidance	Comprehensive		Partial	None		Comprehensive	Comprehensive/ partial
External and ecological validity	Applied studies—different settings (2+)	Applied studies—similar settings (2+)	Real-world informed	Somewhat real-world informed	Not real-world informed	Applied studies—same/different settings	Possible applied studies—similar/different settings

Appendix B: Systematic Review Method

Process of Conducting a Systematic Review of the Evidence behind Core Services

A systematic review of the scientific literature was undertaken to locate all empirical articles examining the impact of advocacy, counseling, safety planning, shelter, and support group services on the lives of adult survivors of domestic violence. Articles were located through computerized journal databases (PsycINFO, PubMed, JSTOR, and Web of Science) using multiple combinations of keywords, such as: domestic violence, intimate partner violence, gender-based violence, effectiveness, evaluation, longitudinal, intervention, randomized, rigorous, trial, and impact. For each review, these keywords were crossed with the appropriate service: either advocacy, counseling, therapy, safety planning, shelter, or support group. Additional articles were then located using a backward search through relevant articles' reference lists.

Results were limited to peer-reviewed, empirical articles published after 1994 and written in English. To be included in this review, each study had to meet the following criteria:

- The program or service specifically targeted adult survivors of IPV;
- The service was provided within or in collaboration with a domestic violence program;
- It empirically examined one or more health or psychosocial outcomes, including psychological, emotional, behavioral, and social; and
- It included a comparison or control group to examine program impact.

Initial searches yielded 382 results for advocacy interventions; 1,243 results for counseling; 725 results for shelter; 97 for safety planning; and 628 for support groups. However, the vast majority of these articles did not meet the inclusion criteria for this review. After removing studies that failed to meet the inclusion criteria, this review was based on 15 articles: four for advocacy, seven for counseling, two for safety planning, none for shelter, and two for support groups.

Appendix C: Empirical Evidence Tables: Results from the Systematic Review

Advocacy (n=4)

Citation	Sample	Study Eligibility	Study Design	Intervention Components	Race / Ethnicity	Retention Rate	Results	Limitations / Caveats
Bell, M.E., & Goodman, L.A. (2001). Supporting battered women involved with the court system: An evaluation of a law school-based advocacy intervention. <i>Violence Against Women</i>, 7, 1377-1404.	81 abused women seeking protection orders at DC Court Domestic Violence Intake Center	Women seeking protection order for a heterosexual relationship, who met poverty guidelines, and were not receiving help from an attorney or other community agency, who did not admit to initiating the violence, who did not appear 'out of it' (talking to themselves or intoxicated), and access to a phone	Quasi-experimental, researcher invited 59 women to participate in surveys, while law school professor invited 22 women to work with 2 law students as advocates. Surveys at study entry and 6 weeks later.	Two law students worked with participants until they had a protection order. Students provided legal advice and advocacy	93% African American	70% after 6 weeks (95% intervention group, 61% comparison group)	After 6 weeks, women in the advocacy condition reported significantly lower levels of psychological and physical re-abuse and marginally significant increases in levels of emotional support in relation to women in the comparison condition. No group differences in tangible social support or depression across time.	Small sample, quasi-experimental, and retention was not high for comparison group. Short follow-up time (6 weeks).
Bybee, D.I. & Sullivan, C.M. (2002). The process through which an advocacy intervention resulted in positive change for battered women over time. <i>American Journal of Community Psychology</i>, 30, 103-132.	Follow-up analyses to Sullivan & Bybee, 1999	Women exiting a Midwest domestic violence shelter program, who planned to stay in the area for at least three months	Longitudinal Randomized Controlled Trial; latent structural equation modeling to examine process of change	Community Advocacy Program: advocates work with survivors in their communities, 6-8 hrs/wk over 10 weeks. Survivor-driven, strengths-based, focus on obtaining community resources and mobilizing social support for survivors	45% African American 42% European American 7% Latina 2% Asian American; 4% other	94% at 6 and 12-months, 95% at 18 and 24-months	The positive long-term effects were mediated by the planned short-term intervention effects. Initial access to resources and social support led to increased quality of life and safety	Reabuse measure focused only on physical violence; reoccurrence of other types of abuse was not assessed. All participants had been shelter residents, most had low incomes and all were urban dwellers – limiting the generalizability of the findings.

Citation	Sample	Study Eligibility	Study Design	Intervention Components	Race / Ethnicity	Retention Rate	Results	Limitations / Caveats
Sullivan, C.M. & Bybee, D.I. (1999). Reducing violence using community-based advocacy for women with abusive partners. Journal of Consulting and Clinical Psychology, 67, 43-53.	278 women exiting a Midwest domestic violence shelter program	Experienced abuse in prior three months; planned to stay in the area for at least three months	Longitudinal Randomized Controlled Trial; 143 women in E, 135 in C; participants interviewed six times over two years (pre, 10 wks, 6 mo, 12 mo, 18 mo, 24 mo)	Community Advocacy Program: advocates work with survivors in their communities, 6-8 hrs/wk over 10 weeks. Survivor-driven, strengths-based, focus on obtaining community resources and mobilizing social support for survivors	45% African American 42% European American, 7% Latina, 2% Asian American 4% other	94% at 6 and 12-months 95% at 18 and 24-months	Women in the experimental group experienced less abuse over time, higher quality of life and social support, and had less difficulty obtaining community resources. More than twice as many women receiving advocacy services experienced no violence across 2 years compared with control group. Of those wanting to end the relationship, those with advocates were more successful.	All participants had been shelter residents, most had low incomes and all were urban dwellers – limiting the generalizability of the findings. Some measures were created specifically for this study.
Sullivan, C.M., Bybee, D.I., & Allen, N.E. (2002). Findings from a community-based program for battered women and their children. Journal of Interpersonal Violence, 17, 915-936.	80 mothers and their 80 children, recruited from 3 community sources: (1) after exiting a DV shelter, (2) community-based family services program or (3) state Social Services.	At least one child living with the mother had to be between the ages of 7-11 and interested in participating in a 12-week children's program. Mother had to have experienced DV within previous four months and family had to plan on staying in the area for the next eight months	Longitudinal Randomized Controlled Trial; 45 families randomized into E; 35 into control.	Advocates provided women and children with 16-weeks of intensive community-based advocacy. Children also attended a 12-week support and education group	49% non-Hispanic white 39% African American 5% Hispanic/Latina 5% Multiracial 1% Asian 1% Native American	95% over 8 months	Mothers in the experimental group experienced reduced depression and improved self-esteem. Children in the experimental group showed increased self-competence and self-worth, while children in the control group showed little change.	Sample size small and follow-up time period (8 months) relatively short.

Counseling (n=7)

Citation	Sample	Study Eligibility	Study Design	Intervention Components	Race/ Ethnicity	Retention Rate	Results	Limitations/ Caveats
Chronister, K.M., & McWhirter, E.H. (2006). An experimental examination of two career interventions for battered women. Journal of Counseling Psychology, 53, 151-164.	72 women recruited from a small Northwest urban community using flyers	At least 18 years-old female, able to speak and read English, IPV victimization in past five years, interested in participating in a career group	Randomized Controlled Trial; participants randomized into one of three groups: standard career intervention, career intervention plus consciousness raising; or wait-list control	Five week programs; both included standard career intervention components (Standard). One group also included exercises designed to enhance critical consciousness (Standard Plus).	75% European American 7% Latina 5% Native American 5% Biracial 4% Multiracial 1% Pacific Islander, 1% other	Not reported, but authors noted that retention was a major issue. Some women were prevented from attending sessions by their abusive partners, others too busy to attend all sessions	Standard participants had higher career-search self-efficacy and Standard-Plus had higher critical consciousness post intervention. At followup, Standard Plus had higher critical consciousness and had made more progress toward goal achievement than Standard participants.	Small sample size; measures created specifically for the study; brief follow-up time frame
Crespo, M., & Arinero, M. (2010). Assessment of the efficacy of a psychological treatment for women victims of violence by their intimate male partner. The Spanish Journal of Psychology, 13(2), 849-863.	53 IPV survivors, recruited through DV agencies around Madrid, Spain	Women 18 years of age or older; having suffered IPV by a male partner; presenting posttraumatic symptoms without meeting all the diagnostic criteria for PTSD; and receiving no other current treatment.	Randomized Controlled Trial Post-traumatic symptoms, anxiety, depression, self-esteem and anger expression were analyzed pre, post, and at 1-, 3-, 6- and 12-mo follow-up	One group involved exposure technique (n = 28) while the other included communication skills training (n = 25). The treatment was carried out in 8 weekly sessions	Not reported	68%	Across both groups, severity and intensity of their main mental health symptoms were reduced, and most changes persisted over time. Exposure group more likely to reduce Avoidance subscale of PTSD compared to Communication group.	Small sample size; differential attrition for women experiencing physical abuse and involved in legal actions may indicate that this program is not applicable for all survivors of IPV; study did not have an untreated control group so differences from untreated populations unknown.

Citation	Sample	Study Eligibility	Study Design	Intervention Components	Race/ Ethnicity	Retention Rate	Results	Limitations/ Caveats
Hernández-Ruiz, E. (2005). Effect of music therapy on the anxiety levels and sleep patterns of abused women in shelters. Journal of Music Therapy, 42, 140-158.	28 adult female shelter residents	In shelter at least 2 days but no more than 1 week.	Randomized Controlled Trial with pre and post measures. Participants matched on sleep quality at pre.	Across 5 consecutive days, 20-minute recordings of participant-selected music and a Progressive Muscle Relaxation script. Control group lay quietly for 20 minutes.	Not reported.	Not reported.	Intervention reduced anxiety and increased sleep quality.	Small sample size with no followup period. Limited detail about participants and retention. Can not separate out effect of music from relaxation script.
Johnson, D. M., Zlotnick, C., & Perez, S. (2011). Cognitive behavioral treatment of PTSD in residents of battered women's shelters: Results of a randomized clinical trial. Journal of Consulting and Clinical Psychology, 79(4), 542-551.	70 adult female residents of one of two inner-city battered women's shelters in a mid-sized Midwestern city.	Had to experience incident of IPV 1 month prior to shelter admission, meet diagnostic criteria for IPV-related PTSD or sub-threshold. Exclusion criteria: symptoms of psychosis, lifetime bipolar diagnosis, suicidal ideation, on psychotropic medication with a change in medication type or dose in last month or were in concurrent individual therapy.	Randomized Controlled Trial comparing Helping to Overcome PTSD through Empowerment (HOPE; n =35)with standard shelter services (Control; n=35). Measurement at 1-week, 3- and 6-months post-shelter	HOPE included 12 sessions, approximately twice weekly, while in shelter. The first five sessions focus on psycho-education regarding inter-personal violence, PTSD, safety planning, and empowerment. Later sessions incorporate established CBT skills to manage PTSD and its associated features.	50% African American 42.9% Caucasian 7.1% Other 4.3% Hispanic	97% at 1-week post shelter, 94% at 3-months, and 95% at 6-months. However, most participants left shelter prior to completing HOPE (62.9%), with many participants leaving before receiving a minimal dose (33.3%).	Women in control group were 12 times more likely to be reabused relative to participants who received a minimal dose of HOPE. HOPE did not reduce overall PTSD severity, but findings suggest that HOPE significantly reduced effortful avoidance, emotional numbing, and arousal symptom severity. HOPE participants reported fewer depression symptoms over follow-up relative to control group.	Small sample size. Interviewers not masked to participant's treatment condition. Relatively short follow-up period – no information on durability of treatment effects. No ability to generalize to survivors who do not reside in shelters.

Citation	Sample	Study Eligibility	Study Design	Intervention Components	Race/ Ethnicity	Retention Rate	Results	Limitations/ Caveats
Kim, S., & Kim, J. (2001). The effects of group intervention for battered women in Korea. Archives of Psychiatric Nursing, 15(6), 257-264. doi: 10.1053/apnu.2001.28682	Participants selected from two shelters in Seoul, South Korea	Not reported.	Quasi experimental, nonequivalent design. 30 women from one shelter received intervention; 30 from second shelter were comparison group. Pre-test, post-test eight weeks later	Eight 90-minute group sessions attending to: trauma, feelings, understanding self, identifying batterer's characteristics, stress management, action plans, and empowerment.	Not reported.	Treatment group lost 14 participants (47%) and the comparison group 13 (43%). 16 women in treatment group (53%) completed all 8 sessions.	No significant group differences detected	Small sample; non-random assignment; low retention rates. Some measures not validated with Korean populations.
McWhirter, P.T. (2011). Differential therapeutic outcomes of community-based group interventions for women and children exposed to intimate partner violence. Journal of Interpersonal Violence, 26(12), 2457 – 2482.	Women (n=46) and their children (n=48) residing in temporary family homeless shelters in a large Southwestern metropolitan area	History of exposure to violence by an intimate partner within one year rated with a score of 15 or higher on the HITS tool of IPV and reported a child aged between 6 and 12 years to be present during at least one incident of IPV within the past year	Multiple focused experimental construct validity design. Participants randomly assigned to either emotion-focused or goal-oriented group; no control group	Groups met weekly for 5 weeks – women-only for 60 minutes, and children for 45 minutes and then a facilitated 60 minute joint session. Goal-oriented groups focused on increasing internally guided goal-oriented change. Emotion focused groups include an initial cognitive behavioral psycho-educational segment that was then processed via a gestalt approach.	47% White 20% Latino, 16% African American 11% Native American 2% Asian American	Not reported.	Participants in both groups demonstrated decreased family conflict and improved quality of social support; however, significantly greater reduction in family conflict was reported among women who participated in the goal-oriented intervention, and significantly greater increases in social support were reported among women who participated in the emotion-focused intervention. Children in both groups reported decreases in family and peer conflict and increases in state of emotional well-being and self-esteem. Women in both groups reported decreases in depression and increases in family bonding and self-efficacy.	No control group; small sample size. Limited generalizability to other populations of survivors and their children.

Citation	Sample	Study Eligibility	Study Design	Intervention Components	Race / Ethnicity	Retention Rate	Results	Limitations / Caveats
Miller, L.E., Howell, K.H., & Graham-Bermann, S.A. (2014). The effect of an evidence-based intervention on women's exposure to intimate partner violence. American Journal of Orthopsychiatry , 84, 321-328.	120 women recruited from DV shelters and the general community in the Midwest US and Southern Ontario, Canada	Experience of IPV within previous two years and have a child between the ages of four and six years-old.	Raandomized Wait-List Control Trial; 58 women randomized into treatment and 62 into wait list control. Participants interviewed at baseline, five weeks later and at 6-8 month follow-up	10, 1 hour sessions delivered twice weekly for 5 weeks, with 5-7 participants per group. Main foci of the intervention include empowerment, safety, social support, and safety as well as topics related to the inter-generational transmission of violence, the woman's family of origin and history of violence, and the effects of violence exposure on young children. Skill development re: conflict resolution, assertive communication, stress management, and emotion regulation.	48% European American 37% African American 8% Biracial 6% Hispanic American 1% Asian American	59% at the 6-8 month follow-up	Both groups reported decreased IPV over time, with those in treatment group reporting even lower rates	Small sample size; no data collected on what other services participants may have used outside the intervention. Study attrition very high.

Safety Planning (n=2)

Citation	Sample	Study Eligibility	Study Design	Intervention Components	Race / Ethnicity	Retention Rate	Results	Limitations / Caveats
Eden, K.B., Perrin, N.A., Hanson, G.C., Messing, J.T., Bloom, T.L., Campbell, J.C., Gielen, A.C., Clough, A.S., Barnes-Hoyt, J.S., & Glass, N.E. (2015). Use of online safety decision aid by abused women: Effect on decisional conflict in a randomized controlled trial. American Journal of Preventive Medicine, 48, 372-383.	826 women from Arizona, Maryland, Missouri & Oregon, recruited from the community through flyers and social media	Women 18 or older who could speak English, who had experienced IPV with the past six months, and who had safe access to the internet	Randomized Controlled Trial to test the effectiveness of a safety decision aid compared with usual safety planning (control), both delivered through a secure website. After randomization, survey questions within website occurred at baseline and after completion of web intervention.	Online Safety Decision Assessment (SDA) has three main components: (1) safety priority-setting activity with feedback; (2) Danger Assessment with feedback; and (3) safety planning strategies and resource information tailored to a woman's responses to the first two components, considering her previous use of protective strategies.	63% white 25% African American 5% Multiracial 10% Hispanic/Latina	After randomization into intervention, 87.3% completed baseline and 86.5% completed website. Of those randomized into control condition, 85% completed baseline and website	Women in intervention group, who received more individualized safety planning than is typical, reported less decisional conflict about their safety planning.	Not longitudinal – unable to capture change over time. Unable to tie reduced decisional conflict to actual increase in safety. Safe access to the internet excludes some survivor populations.
McFarlane, J., Malecha, A., Gist, J., Watson, K., Batten, E., Hall, I., & Smith, S. (2004). Increasing the safety-promoting behaviors of abused women. The American Journal of Nursing, 104, 40-51.	150 abused women who applied and qualified for a protection order in an urban district attorney's office	English or Spanish speaking women applying and qualifying for a protection order	Two-group repeated measures Randomized Controlled Trial to test a phone intervention (6 calls over 8 weeks) intended to increase "safety-promoting behavior" of abused women. 75 women in each group; Phone interviews at pre, 3mo, 6mo, 12mo and 18mo followup	Participants received the usual services from the District Attorney's office, as well as six intervention calls. The first intervention call occurred between 48 and 72 hours of intake; the remaining were made at one, two, three, five, and eight weeks afterward. Each call began with the safety-promoting behavior checklist, noting behaviors adopted since the previous call. During the intervention calls, the investigator discussed specific safety-promoting behaviors.	33% African American 26% White, 41% Latina	99%	Adoption of safety-promoting behaviors significantly increased over time among women in the intervention group, and they continued to practice these behaviors for 18 months.	Sample limited to women seeking and qualifying for a protection order; unknown if the intervention increases actual safety

Support Groups (n=2)

Citation	Sample	Study Eligibility	Study Design	Intervention Components	Race / Ethnicity	Retention Rate	Results	Limitations/Caveats
Brownell, P., & Heiser, D. (2006). Psycho-educational support groups for older women victims of family mistreatment: A pilot study. Journal of Gerontological Social Work, 46(3-4), 145-160.	16 women aged 69 to 83 identified by community partners	Experience of physical, psychological, and/or financial abuse; no significant cognitive impairment; able to communicate in English; able and willing to attend a weekly support group meeting, 2 hours in length, for 8 consecutive weeks	Randomized Controlled Trial; 9 randomized into intervention, 7 into control In-person interviews at pre, post, and 2mo follow-up	Psycho-educational support group; content based on a curriculum used by NOVA House, an elder abuse shelter program in Manitoba, Canada. The 8-week, 2-hour per week support group included topics such as: domestic violence against older women; dealing with depression, anxiety and stress; coping skills; dealing with substance abuse; and identifying community resources.	50% White 44% Black 6% Asian Pacific Islander or Hispanic	Not reported	While eight of the nine support group members said that the group was helpful in increasing their self-esteem and well-being, the authors noted no significant group differences two months post-intervention between those who did and did not receive the intervention.	Study did not report their analyses nor discuss attrition, and this sample was extremely small, making it difficult to draw any conclusions about the efficacy of this group.
Constantino, R., Kim, Y. & Crane, P.A. (2005). Effects of a social support intervention on health outcomes in residents of a domestic violence shelter: A pilot study. Issues in Mental Health Nursing, 26, 575-590.	24 women recruited from domestic violence shelter	First-time shelter users	Randomized Controlled Trial; Experimental design; Users were randomized into either the support group (n=13) or a no-treatment group that involved free-flowing weekly discussion in a group setting (n=11). Measures at pre and post	Groups met twice weekly for four weeks to coincide with the shelter's 30-day time limit. The support group was led by a trained nurse and focused on helping women increase their social support networks, access to community resources, and self-esteem.	71% White 29% Black	Not reported	Post intervention, those in the experimental group showed greater improvement in psychological distress symptoms and reported higher feelings of social support. They also showed less health care utilization than did the women who did not receive the intervention.	Small sample with no follow-up. The authors did not report attrition rates nor the number of sessions attended by women in either group, although the Discussion section implied that attrition may have been a problem.

Appendix D: Key Informant Interviews Method

In December 2015, we interviewed seven key informants to provide additional context and ideas for the framing paper. To the extent possible (given their respective expertise), key informants were asked about:

- Challenges to evaluating DV services
- Research methods (e.g., designs, data collection techniques, analytic strategies) that would be especially useful in evaluating DV services
- Example projects to build the evidence base in the DV and other fields, including those that target capacity building, networks or partnerships, and cross-site evaluations
- Example initiatives or ideas for strategic grant and funding opportunities

Key Informant	Affiliation
Deborah Bybee, PhD	Professor, Psychology Department Michigan State University
Lisa Goodman, PhD	Professor, Counseling, Developmental and Educational Psychology Department Lynch School of Education, Boston College
Marylouise Kelley, PhD	Director Family Violence Prevention and Services Program Family and Youth Services Bureau Administration for Children and Families U.S. Department of Health and Human Services
Robin Miller, PhD	Professor, Psychology Department Michigan State University
Laura Peck, PhD	Principal Scientist, Social and Economic Policy Abt Associates, Inc.
Josie Serrata, PhD	Assistant Director of Research for the National Latin@ Research Center on Family and Social Change, National Latin@ Network for Healthy Families and Communities/Casa de Esperanza
Lauren Supplee, PhD	Division Director, Administration for Children and Families, Office of Planning, Research and Evaluation, Division of Family Strengthening

Appendix E: References

- Bell, M. E., & Goodman, L. A. (2001). Supporting battered women involved with the court system: An evaluation of a law school-based advocacy intervention. *Violence Against Women*, 7, 1377-1404.
- Brownell, P., & Heiser, D. (2006). Psycho-educational support groups for older women victims of family mistreatment: A pilot study. *Journal of Gerontological Social Work*, 46(3-4), 145- 160.
- Bybee, D.I. & Sullivan, C.M. (2002). The process through which an advocacy intervention resulted in positive change for battered women over time. *American Journal of Community Psychology*, 30, 103-132.
- Campbell, J.C., Webster, D., Koziol-McLain, J., Block, C.R., Campbell, D., Curry., M.A., Gary, F., McFarlane, J., Sachs, C., Sharps, P., Ulrich, Y., & Wilt, S.A. (2003). Assessing Risk Factors for Intimate Partner Homicide. *National Institute of Justice Journal*, 250, 14-19.
- Cody, S., Perez-Johnson, I., & Joyce, K. (2015). Administrative experiments: Unlocking what works better and what works for whom. *American Journal of Evaluation*, 36(4), 547-557.
- Chronister, K.M., & McWhirter, E.H. (2006). An experimental examination of two career interventions for battered women. *Journal of Counseling Psychology*, 53, 151-164.
- Constantino, R., Kim, Y. & Crane, P.A. (2005). Effects of a social support intervention on health outcomes in residents of a domestic violence shelter: A pilot study. *Issues in Mental Health Nursing*, 26, 575-590.
- Crespo, M., & Arinero, M. (2010). Assessment of the efficacy of a psychological treatment for women victims of violence by their intimate male partner. *The Spanish Journal of Psychology*, 13(2), 849-863.
- Eden, K.B., Perrin, N.A., Hanson, G.C., Messing, J.T., Bloom, T.L., Campbell, J.C., Gielen, A.C., Clough, A.S., Barnes-Hoyt, J.S., & Glass, N.E. (2015). Use of online safety decision aid by abused women: Effect on decisional conflict in a randomized controlled trial. *American Journal of Preventive Medicine*, 48, 372-383.
- Frangakis, C.E., & Rubin, D.B. (2002). Principal stratification in causal inference. *Biometrics*, 58, 21–29.
- Hernández-Ruiz, E. (2005). Effect of music therapy on the anxiety levels and sleep patterns of abused women in shelters. *Journal of Music Therapy*, 42, 140-158.
- Johnson, D. M., Zlotnick, C., & Perez, S. (2011). Cognitive behavioral treatment of PTSD in residents of battered women's shelters: Results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology*, 79(4), 542-551.
- Jonker, I.E., Sijbrandij, M., van Luijckelaar, M. J. A., Cuijpers, P., & Wolf, J. R. L. M. (2015). The effectiveness of interventions during and after residence in women's shelters: A meta-analysis. *The European Journal of Public Health*, <http://dx.doi.org/10.1093/eurpub/cku092>.
- Kim, S., & Kim, J. (2001). The effects of group intervention for battered women in Korea. *Archives of Psychiatric Nursing*, 15(6), 257-264. doi: 10.1053/apnu.2001.28682
- Lei, H. Nahum-Shani, I., Lynch, K., Oslin, D., & Murphy, S.A. (2012). A 'SMART' design for building individualized treatment sequences. *Annual Review of Clinical Psychology*, 8, 21-48. doi: 10.1146/annurev-clinpsy-032511-143152.
- McFarlane, J., Malecha, A., Gist, J., Watson, K., Batten, E., Hall, I., & Smith. S. (2004). Increasing the safety-promoting behaviors of abused women. *The American Journal of Nursing*, 104, 40-51.

- McWhirter, P.T. (2011). Differential therapeutic outcomes of community-based group interventions for women and children exposed to intimate partner violence. *Journal of Interpersonal Violence*, 26(12), 2457 – 2482.
- Miller, L.E., Howell, K.H., & Graham-Bermann, S.A. (2014). The effect of an evidence-based intervention on women's exposure to intimate partner violence. *American Journal of Orthopsychiatry*, 84, 321-328.
- Olds, D. L., Baca, P., McClatchey, M., Ingoldsby, E. M., Luckey, D. W., Knudtson, M. D., Loch, J.M., & Ramsey, M. (2015). Cluster randomized controlled trial of intervention to increase participant retention and completed home visits in the Nurse-Family Partnership. *Prevention Science*, 16(6), 778-88.
- Olds, D. L., Kitzman, H., Knudtson, M. D., Anson, E., Smith, J. A., & Cole, R. (2014). Effect of home visiting by nurses on maternal and child mortality: Results of a 2-decade follow-up of a randomized clinical trial. *JAMA pediatrics*, 168(9), 800-806.
- Peck, Laura R. (2013). On analysis of symmetrically-predicted endogenous subgroups: Part one of a method note in three parts. *American Journal of Evaluation*, 34, 225-236.
- Puddy, R. W., & Wilkins, N. (2011). *Understanding evidence part 1: Best available research evidence: A guide to the continuum of evidence of effectiveness*. Retrieved from the Centers for Disease Control and Prevention website: http://www.cdc.gov/violenceprevention/pdf/understanding_evidence-a.pdf
- Reardon S.F., Unlu, F., Zhu, P., & Bloom, H. (2014). Bias and bias correction in multi-site instrumental variables analysis of heterogeneous mediator effects. *Journal of Educational and Behavioral Statistics*, 39, 53-86.
- Schorr, L. & Farrow, F. (2011). *Expanding the evidence universe: Doing better by knowing more*. Paper prepared for the 2011 Harold Richman Public Policy Symposium. Washington, DC: Center for the Study of Social Policy. Retrieved from <http://lisbethschorr.org/doc/ExpandingtheEvidenceUniverseRichmanSymposiumPaper.pdf>
- Sullivan, C.M., Warshaw, C., & Rivera, E.A. (2013). Counseling services for domestic violence survivors: Review of the empirical evidence. <http://www.dvevidenceproject.org/wp-content/uploads/CounselingResearchSummary.pdf>
- Sullivan, C.M. (2012). Examining the work of domestic violence programs within a “social and emotional well-being promotion” conceptual framework. <http://www.dvevidenceproject.org/wp-content/uploads/ConceptualFramework.pdf>
- Sullivan, C.M. (2012). Advocacy services for women with abusive partners: A review of the empirical evidence. <http://www.dvevidenceproject.org/wp-content/uploads/AdvocacyResearchSummary.pdf>
- Sullivan, C.M. (2012). Domestic violence shelter programs: A review of the empirical evidence. <http://www.dvevidenceproject.org/wp-content/uploads/ShelterResearchSummary.pdf>
- Sullivan, C.M. (2012). Support groups for women with abusive partners: A review of the empirical evidence. <http://www.dvevidenceproject.org/wp-content/uploads/SupportGroupResearchSummary.pdf>
- Sullivan, C.M. & Bybee, D.I. (1999). Reducing violence using community-based advocacy for women with abusive partners. *Journal of Consulting and Clinical Psychology*, 67, 43-53.
- Sullivan, C.M., Bybee, D.I., & Allen, N.E. (2002). Findings from a community-based program for battered women and their children. *Journal of Interpersonal Violence*, 17, 915-936.
- Sullivan, C.M., & Cain, D. (2004). Ethical and safety considerations when obtaining information from or about battered women for research purposes. *Journal of Interpersonal Violence*, 19(5), 603-618.

World Health Organization (2001). Putting women first: Ethical and safety recommendations for research on domestic violence against women. Geneva, Switzerland: World Health Organization.
<http://www.who.int/gender/violence/womenfirstseng.pdf>