ARTICLE

DEVELOPMENT AND VALIDATION OF THE TRAUMA-INFORMED PRACTICE SCALES

Lisa A. Goodman
Boston College

Cris M. Sullivan
Michigan State University

Josie Serrata
The National Latina Network for Healthy Families and Communities

Julia Perilla
Georgia State University & The National Latina Network for Healthy Families and Communities

Joshua M. Wilson, Jennifer E. Fauci, and Craig D. DiGiovanni
Boston College

Spurred by research demonstrating the pervasiveness and impact of trauma, domestic violence programs are increasingly adopting a trauma-informed approach. In the absence of measurement tools, however, they are unable to determine whether indeed clients experience their practices as trauma-informed. The aim of this study was to create and validate a set of scales that measure the degree to which DV programs are using trauma informed practices (TIP) from survivors' perspectives. We generated measure items based on a literature review, interviews with

This study received funding from Boston College and Michigan State University.
Points of view expressed in this manuscript do not necessarily represent the official position or policies of these entities.
We gratefully acknowledge the invaluable assistance of the expert interviewees, program participants, and program staff members who contributed to the development of this measure, as well as the undergraduate students at Boston College who helped with data management, including Theresa V. Grzeslo, Anna Marie Paolicelli, and Juliette E. Lloyd. We are especially indebted to Drs. Carole Warshaw and Eleanor Lyon for their support of this project and input on a prior version of this article.
Please address correspondence to: Lisa Goodman, Department of Counseling and Developmental Psychology; Lynch School of Education; 310 Campion Hall; Boston College; Chestnut Hill, MA 02461. E-mail: goodmalc@bc.edu

JOURNAL OF COMMUNITY PSYCHOLOGY, Vol. 44, No. 6, 747–764 (2016)
Published online in Wiley Online Library (wileyonlinelibrary.com/journal/jcop).
 DOI: 10.1002/jcop.21799

experts, and focus groups with survivors and advocates; administered the draft measure in the context of a survey of 370 survivors from 15 DV programs; and evaluated its factor structure, reliability, and validity. Results supported the validity and reliability of the TIP Scales, which assess six domains: Environment of Agency and Mutual Respect, Access to Information on Trauma, Opportunities for Connection, Emphasis on Strengths, Cultural Responsiveness/Inclusivity, and Support for Parenting. © *2016 Wiley Periodicals, Inc.*

Over the last several decades, an explosion of research on trauma has demonstrated its high prevalence (e.g., Elhai et al., 2012; Ford et al., 2012) and profound deleterious effects on emotional and physical well-being (Kubiak, 2005; Radford, Corral, Bradley, & Fisher, 2013). Typical psychological effects of trauma, including fear, distrust, and anger, can make it harder for trauma survivors to seek out and engage with the very services and supports designed to help them (Harris & Fallot, 2001). Even when survivors do access services, staff who fail to recognize or understand trauma can inadvertently increase clients' sense of vulnerability and disempowerment (Substance Abuse and Mental Health Services Administration [SAMHSA], 2014; Elliott, Bjelajac, Fallot, Markoff, & Reed, 2005; Warshaw, 2014).

As a result, many human service organizations have adopted a *trauma-informed* (TI) approach to working with clients. The two core ideas underlying a TI approach are that (a) any person seeking services or support might be a trauma survivor and (b) systems of care need to recognize, understand, and counter the sequelae of trauma to facilitate recovery (SAMHSA, 2014). A TI approach is not a narrowly defined treatment; rather, it is an overarching framework that guides the behavior of every actor in the system, from the receptionist to the direct service provider to the board member–a "'universal design' for serving trauma survivors, provided to all, by all" (DeCandia, Guarino, & Clervil, 2014, p. 8).

Since its introduction in 2001 in the context of mental health and substance abuse services (Harris & Fallot, 2001; Morrissey et al., 2005), TI principles have spread widely to systems ranging from child welfare agencies and schools (Ko et al., 2008) to homeless shelters (Hopper, Bassuk, & Olivet, 2010) and prisons (Harner & Burgess, 2011). Many characteristics of the TI approach, such as being respectful, holistic, and strengths-based, overlap with what have become general expectations for competent practice. However, TI care also includes numerous specific practices–such as universal screening for trauma history, training of all staff in the nature and effects of lifetime trauma, and educating clients about "trauma triggers" and potential coping responses–that are not standard practice in many human service programs (Harris & Fallot, 2001).

Programs serving victims of domestic violence[1] either want to incorporate a TI approach into their work or believe they are already doing so; however, there are no psychometrically sound measures to evaluate these efforts or to help generate research on their effects on survivor outcomes. The absence of such a measure hampers efforts to establish, refine, and promote TI approaches to working with survivors. We aimed to fill this gap by developing and validating a set of scales that can measure the degree

[1]We use the term "domestic violence" (DV) here to refer to violence committed by an intimate partner. Although the phrase "intimate partner violence" has become the preferred way to refer to this kind of abuse, we use the term DV to be consistent with the way that programs describe themselves (e.g.,"DV programs").

to which domestic violence (DV)[1] programs incorporate TI services into their work. The next section summarizes the main principles of a TI approach in the DV context.

TI PRINCIPLES IN THE DV CONTEXT

Those seeking help from DV programs have experienced some form of physical assault, stalking, sexual assault, or psychological abuse. Many have faced systematic intimidation and control and have been deprived of core liberties such as the freedom to move about, socialize, work, or parent (Stark, 2007). Not surprisingly, this kind of victimization results in a range of mental health difficulties, including depression, substance abuse, anxiety, and posttraumatic stress disorder (Bennice, Resick, Mechanic, & Astin, 2003; Dillon, Hussain, Loxton, & Rahman, 2013).

Further, DV survivors seeking services often struggle with additional challenges that precede or accompany the abuse itself, including poverty, homelessness, discrimination, immigration, or childhood, chronic, or historical trauma (Perilla, Serrata, Weinberg, & Lippy, 2012; Warshaw, Sullivan, & Rivera, 2013). Over the last 15 years, DV programs have reported that program participants are struggling with more severe, chronic, and varied traumatic experiences and mental health difficulties than ever before–challenges that many staff feel unprepared to address (Melbin, Smyth, & Marcus, 2014).

Propelled by a growing awareness of these challenges, many DV programs have made great efforts to incorporate TI principles into their work (Warshaw, Lyon, Phillips, & Hooper, 2014). Two recent projects aimed to capture and distill the various ways that DV-focused organizations have operationalized the principles underlying a TI approach: In 2013, the National Center on Domestic Violence, Trauma, and Mental Health conducted interviews with staff at 45 agencies serving DV survivors, 30 of which were DV-specific programs (Phillips, Lyon, Warshaw, & Fabri, 2013). Findings revealed a consensus view that a TI approach goes beyond basic "good advocacy" to incorporate a an understanding of the effect of trauma into every aspect of an agency's work, including its structure and policies; its training and supervision of staff; its approach to working with survivors; and its involvement with the larger community.

Focusing more specifically on TI *practices*, that is, direct interactions between staff and survivors, we (Wilson, Fauci, & Goodman, in press) conducted a qualitative content analysis of available publications describing TI approaches for DV programs to identify a set of common practices. Results of this analysis identified a set of practices that fall within six broad principles: (a) establishing emotional safety, (b) restoring choice and control, (c) facilitating connections, (d) supporting coping, (e) responding to identity and context, and (f) building strengths. These broad principles are consistent with general conceptualizations of TI principles (e.g., Harris & Fallot, 2001); however, the specific practices identified are often unique to the experiences and needs of DV survivors and the organizations that support them.

The rapid spread of TI principles and practice in DV programs across the country begs the question of whether a TI approach improves outcomes for survivors. In one of the few efforts to address this question, the Women, Co-Occurring Disorders, and Violence Study tested the effectiveness of integrating trauma, mental health, and substance abuse interventions for women with co-occurring disorders and a history of abuse. A meta-analysis of findings across nine settings indicated that women who had received TI treatment showed significant improvements in mental health and trauma symptoms relative to treatment as usual (Morrissey et al., 2005). Although this study lays important groundwork, it did not focus specifically on DV programs or survivors.

Given the strong and growing interest in bringing TI services to DV programs and the growing consensus on its core elements in a DV context, it is time to explore whether and to what degree programs are succeeding in their efforts to adopt a TI approach, and to what extent its adoption contributes to positive mental health and well-being for program participants. We therefore set out to develop and validate a set of scales in English and Spanish that would assess the degree to which survivors perceive DV programs to be TI, focusing specifically on staff practices that involve survivors directly rather than the infrastructure required to do this work.

METHODS

Initial Item Generation

Four data sources contributed to the development of items for the initial measure, each of which helped to ensure strong content validity. First, a qualitative analysis of relevant publications (see Wilson, Fauci, & Goodman, in press) produced a set of overarching principles and specific practices that became the basis for draft survey items. Second, interviews with 15 national experts on TI practices in the DV context helped to expand the item pool. This process started with a meeting organized by the National Center on Domestic Violence, Trauma & Mental Health and then expanded through snowball sampling.

Third, after generating initial items, we conducted separate focus groups with survivors and advocates.[2] The survivor focus group included 20 participants (eight Latina, eight African American/Black, and six White) from two separate DV support groups outside a major city in the Northeast. This process resulted in participants adding a number of additional items related to cultural sensitivity. Consistent with many of our interviewees, members of this focus group also suggested that we replace the word "trauma" with the words "abuse" or "abuse and other hardships," arguing that the word trauma would not resonate with many participants. The advocate focus group comprised 20 practitioners (11 Latina, eight Caucasian, and one Native staff member) from a national organization that provides technical assistance on working with Latino families experiencing DV. They tweaked language in several places to make items more culturally sensitive and accessible for Latina survivors, added items, and confirmed that the word "abuse" was preferable to "trauma."

Last, we tested the initial measure with the first 15 participants recruited into the validation study (described next). Based on an assessment of skipped items, we modified the wording of several items, clarified skip patterns, and added an "I don't know" response for items having to do with staff behavior that not all clients may have knowledge about (e.g., *staff understand the challenges faced by people who are immigrants*).

Given these changes, we divided the draft measure into three separate scales, each of which had a distinct set of response choices or was appropriate to only a subset of potential respondents. The main 4-point scale initially included 48 items and response choices ranging from 0 (*not at all true*) to 3 (*very true*). The nine "cultural responsiveness and inclusivity" items used the same response choices but included an "I don't know" response choice; and the five parenting items allowed anyone without children to skip those items. Once we had arrived at a final set of items, together called the TIP Scales, we examined its psychometric properties by administering it to 370 DV survivors.

[2]The word "advocate" is the preferred term for providers of services in most domestic violence programs

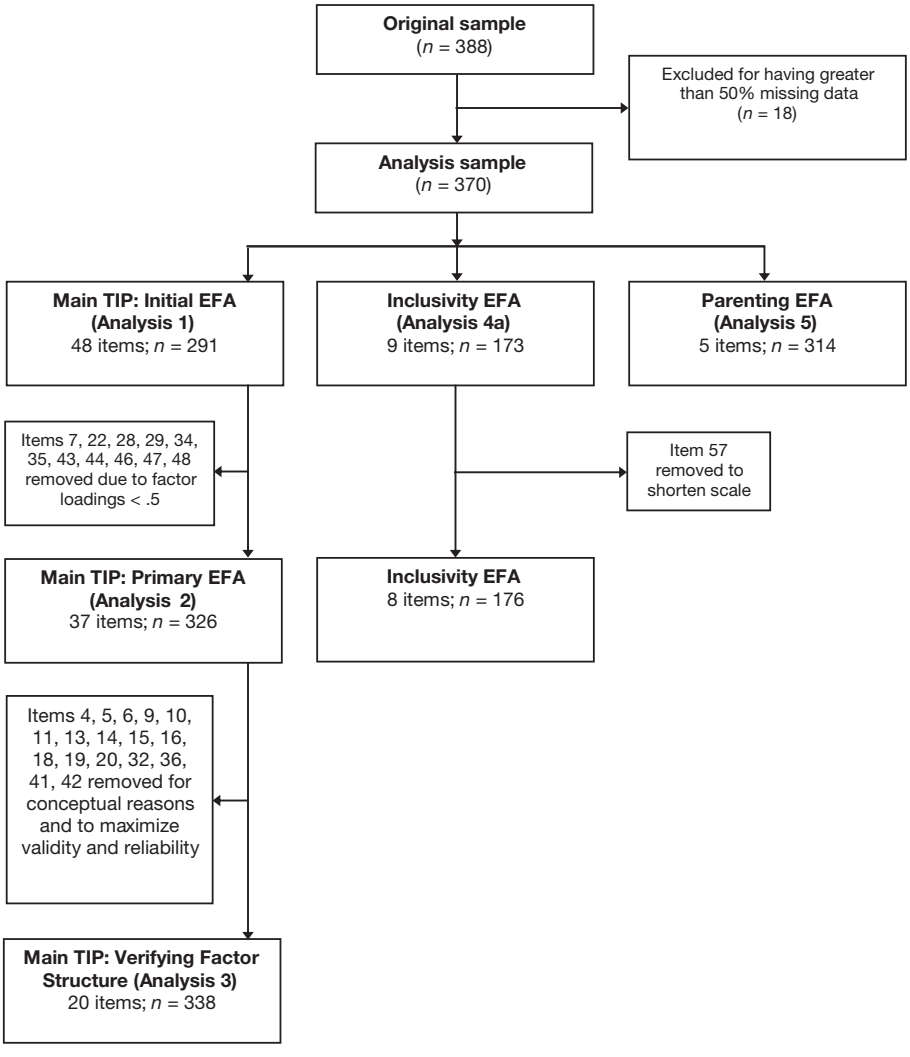

Figure 1. Progression of sampling and analyses.
Note. EFA = exploratory factor analysis; TIP = trauma-informed practice. Analyses 1, 4, and 5 were conducted separately using the analysis sample ($n = 370$).

Psychometric Evaluation

We collected data from a convenience sample of survivors seeking DV services at one of 15 urban and suburban organizations in five states across Midwestern and Northeastern United States. All programs provided safety planning, counseling, and information and referral to DV survivors. Most provided emergency shelter, with stay lengths of several weeks to 6 months. Several also offered transitional living programs, with stay lengths of up to 2 years.

Eligible participants were aged 18 years or older and English- or Spanish-speaking. We eliminated data from 18 participants who responded to less than 50% of the survey, leaving a sample of 370 (see Figure 1). The final sample had a mean age of 36.4 (standard deviation [*SD*] = 12.0), with a racial and ethnic composition of 39.1% White, 24.4%

African or Black/African American, 23.8% Hispanic/Latino, 7.8% Multiracial, and 5.0% Other. Most (74.9%) were born in the United States and most (86.2%) completed the survey in English.

Participants' socioeconomic status was mixed, with 38.8% reporting they attended at least some college or graduated college and 51.6% reporting that they were unemployed. Half (53.1%) reported that they either couldn't pay their bills or had trouble doing so. Regarding relationship and family characteristics, the majority (87.7%) identified as heterosexual and 16.5% reported being in a relationship. Most women (67.6%) had children; of those, 51.9% reported that their abuser was their children's father. The sample was evenly divided between survivors using community-based programs (50.8%) and those using residential programs (49.2%).

Procedures

Participants were recruited through staff at each program. We supplied all programs with flyers, descriptions of the study, consent forms, surveys, gift cards for participants, and stamped, self-addressed envelopes so that participants could return the surveys directly to the research team.

Program staff posted flyers about the project in Spanish and English and made announcements about it in client meetings, describing the study as an investigation of survivors' experiences seeking help from the program. Staff members stressed that participation was entirely voluntary and would not affect their program participation in any way. Participants could take the survey on their own or in groups in either English or Spanish. To prevent anyone with literacy issues from feeling uncomfortable, staff offered to read the survey to any participant who "did not feel like" doing the survey on their own (with the participant still responding to the questions privately). Although participants returned the consent forms to program staff, they sent their surveys directly to the first author using the self-addressed envelopes we provided. All participants received a $20 gift card to a local store. Survey administrations lasted approximately 30 minutes.

Measures

In addition to the initial TIP items, the survey included a range of demographic questions and established measures of satisfaction with services and alliance with staff, designed to test the convergent validity of the TIP Scales. All measures were translated and back-translated to ensure their appropriateness for use with monolingual Spanish-speaking participants.

Advocate–survivor alliance. As one test of convergent validity, we administered two subscales of the Short-Revised version of the Working Alliance Inventory (WAI-SR; Hatcher & Gillaspy, 2006). The WAI aims to assess the "collaborative and affective bond" between participants and staff (Martin, Gaske, & Davis, 2000, p. 438), which captures the relational dynamic that is at the heart of TI practice (Phillips et al., 2013; SAMHSA, 2014). The WAI is one of the most reliable and valid measures of the therapist–client relationship (Martin et al., 2000) and has been employed in community-based contexts (e.g., Florsheim, Shotorbani, Guest-Warnick, Barratt, & Hwang 2000).

The SR version is a 12-item scale that includes Goal, Task, and Bond subscales. We used the four-item Goal subscale (e.g., *My therapist and I collaborate on setting goals for my therapy*) and the four-item Bond subscale (e.g., *I believe my therapist likes me*) because these are most relevant to the advocate–survivor relationship. To adapt the measure to the

context of advocacy, we also changed the word "therapist" to the word "staff members" and removed the word "therapy" or changed it to the word "work" (e.g., *The staff and I collaborate on setting goals for our work*). Items were scored using a 5-point Likert scale ranging from 1 (*rarely/never*) to 5 (*always*). Cronbach's alpha was .95 for the full measure, .94 for the Goal subscale, and .92 for the Bond subscale.

Client satisfaction with services. As another test of convergent validity, we administered the Client Satisfaction Questionaire-8 (CSQ-8; Larsen, Attkisson, Hargreaves, & Nguyen, 1979), a measure of the participants' satisfaction with the program from which they were seeking help. The CSQ-8 was chosen because it reflects the survivor-centered focus of a TI approach. Participants responded to items (e.g., *How would you rate the quality of services you received?)* on a 4-point scale depending on the nature of the question (e.g., "poor" to "excellent" vs. "quite dissatisfied" to "quite satisfied"). Scores range from 8 to 32, with higher values indicating higher satisfaction. The CSQ-8 has been studied extensively and has been shown to have high internal consistency and predictive validity in a variety of samples, including inpatient and outpatient mental health settings (Attkisson & Greenfield, 2004). Cronbach's alpha for this scale for this sample was .96.

RESULTS

The development of the TIP Scales involved separate exploratory factor analyses (EFAs) of, respectively, (a) the main 48 items, (b) the nine items assessing cultural responsiveness and inclusivity, and (c) the five items assessing parenting support (see Figure 1). We chose this process for several reasons. First, we assessed the five Parenting Support items as a separate measure because not all participants in this sample had children. Second, we assessed the nine Cultural Responsiveness and Inclusivity items as a separate measure because, as noted earlier, they included an additional "I don't know" response category. Because this response was coded as missing, these nine items had a larger amount of missing data (9.5% to 41.4%, depending on the item). Including them in the main EFA would have reduced the sample size to 134, which was inadequate for a full EFA (Costello & Osborne, 2005).

Before beginning the factor analysis, we examined missing data on all 62 TIP items. The percentage of skipped responses ranged from .3% to 12.4%, with the majority of items having less than 3% missing data. Given the low percentage of missing data, we used a listwise deletion of cases for our EFAs (Hair, Black, Babin, & Anderson, 2010); participants that had missing data for any of the included items used in an analysis were removed from the analysis.

We used a principal axis-factoring extraction method throughout our process because of its robustness and the fact that responses on our items were moderately skewed ($M_{skew} = -1.79$, $SD = .39$; Costello & Osborne, 2005; Fabrigar, Wegener, MacCallum, & Strahan, 1999). We applied an oblique rotation because of the likelihood that the factors would be correlated.

Main TIP

Analysis 1: Initial EFA. We conducted a preliminary EFA on the main 48 items, which included 291 participants (those with no missing data on any items). The participant to item ratio was approximately 6:1, which is adequate for an EFA (Worthington & Whittaker, 2006). The Kaiser-Meyer-Olkin (KMO) value was .97, indicating excellent

sampling adequacy and Bartlett's test of sphericity was significant ($p < .001$), indicating the correlation matrix was appropriate for an EFA. We allowed the analysis to extract factors with an eigenvalue greater than one, which resulted in five factors. The scree plot was inconclusive, with no clear factor cutoff. Upon examining the factor loadings, the fifth factor contained only three items above .32 and all of them cross-loaded onto other factors. Extraction communalities ranged from .24 to .90.

To create a cleaner factor structure, we removed a set of items based on the following quantitative criteria: (a) if their factor loadings were less than .5 (Costello & Osborne, 2005) and (b) if their loadings were greater than .32 on two or more factors (i.e., they were cross-loaded; Tabachnick & Fidell, 2001). We removed 11 items that loaded below .5 (three of which had cross-loadings greater than .32), leaving 37 items (see Figure 1).

Analysis 2: Primary EFA. We ran another EFA on these 37 items ($N = 326$; Table 1), revealing a clearer four-factor structure with strong factor loadings (greater than .5) and no cross-loadings for any items. Extraction commonalities ranged from .50 to .87. Factor correlations ranged from .26 to .70, indicating that factors did not share more than 49% of the variance. Factor 1 had 21 items, factor 2 had eight items, factor 3 had three items, and factor 4 had five items.

Next, we sought to reduce the total number of items still further to keep the scale parsimonious, especially for factor 1, which had the highest number of items. To do this, we used a combined quantitative and qualitative approach. First, we examined the effect on factor reliability and validity of deleting each item, one by one (see e.g., Raubenheimer, 2004). Second, we then deleted an item if its presence caused factor correlations to go above .7 and/or if we agreed as a group that the item had confusing wording, was redundant in the presence of other items, or was a conceptual mismatch with other items. Ultimately, this process led to the deletion of 17 items, reducing factor 1 to nine items, factor 2 to five items, and factor 4 to three items; factor 3 remained the same (see Table 1 for specific items). We agreed that the four factors qualitatively retained their meaning after these reductions.

Analysis 3: Verifying factor structure. We conducted a final EFA on the remaining 20 items ($N = 338$) to ensure the item removal process did not affect the factor structure from analysis 2 (Worthington & Whittaker, 2006). An examination of the scree plot revealed a clear four-factor solution. In addition, although the fourth factor had an eigenvalue of .96, we considered this close enough to one when using a factor retention criterion of having eigenvalue greater than one. Extraction commonalities ranged from .50 to .89 and factor correlations ranged from .35 to .70. Factor 1 loadings ranged from .63 to .91, factor 2 from .66 to .99, factor 3 from .60 to .92, and factor 4 from .66 to .85. The final four-factor solution accounted for 79.3% of the shared variance. Initial eigenvalues for factors 1 to 4 were 11.63, 1.97, 1.31, and .96, respectively, and accounted for 58.13%, 9.84%, 6.54%, and 4.79% of the variance, respectively.

After examining the factors, the team agreed that factor 1 assessed an Environment of Agency and Mutual Respect (i.e., Agency), factor 2 assessed Access to Information on Trauma (i.e., Information), factor 3 assessed Opportunities for Connection (i.e., Connection), and factor 4 assessed an Emphasis on Strengths (i.e., Strengths). The final 338 participants differed from the 32 individuals not included in the analysis in that those included were more likely to be born in the United States, $\chi^2(1, 362) = 5.76$, $p < .05$, and to have taken an English version of the survey, $\chi^2\ (1, 370) = 6.06$, $p < .05$. There were no significant differences in employment status, race, or education.

Table 1. Factor Loadings and Extracted Communalities for Main TIP Items (Analysis 2)

	Factor analysis structure coefficients				
Item	*1*	*2*	*3*	*4*	h^2
20. I feel comfortable in my interactions with staff.	.93	−.06	.04	.03	.88
19. Staff explain clearly how things work in this program.	.90	.08	−.03	−.12	.73
26. Staff respect the choices that I make.	**.89**	.01	.12	−.10	.80
12. In this program, I can share things about my life on my own terms and at my own pace.	**.89**	.01	−.04	−.01	.76
13. Staff approach me with a caring attitude.	.89	.01	−.04	.07	.85
10. Staff respect me.	.87	−.02	−.07	.13	.84
9. I feel that staff are on my side.	.86	.02	−.08	.08	.80
24. Staff respect my privacy.	**.84**	.10	−.03	−.10	.69
11. Staff show me compassion even if I am having a very difficult time.	.84	.00	−.01	.10	.84
8. Staff treat me with dignity.	**.82**	−.06	−.05	.17	.78
17. I can trust staff.	**.79**	−.04	−.03	.18	.78
18. I can explore my feelings with staff.	.74	.04	.04	.17	.83
25. I decide what I want to work on in this program.	**.73**	.01	.11	−.14	.50
21. Staff can handle difficult situations.	**.71**	.13	.11	−.09	.65
27. Staff understand that I know what's best for me.	**.68**	.02	.22	.02	.69
16. I feel safe to bring up the challenges I face right now, whatever they are.	.65	.12	.01	.19	.77
23. Staff are supportive when I'm feeling stressed out or overwhelmed.	**.64**	.20	.10	.04	.75
32. Staff understand the relationship challenges I'm facing.	.63	.09	.24	.04	.74
14. Staff seem interested in hearing about my experiences.	.62	.08	.04	.24	.78
15. I feel safe to talk about what is really on my mind.	.58	.15	.04	.22	.76
6. I am able to communicate with staff in the language of my choice.	.56	−.06	−.08	.27	.50
38. I have the opportunity to learn how abuse and other difficulties affect responses in the body.	−.06	**.98**	−.01	−.05	.85
37. This program gives me opportunities to learn how abuse, and other difficulties affect peoples' ability to think clearly and remember things.	−.07	**.94**	−.04	.06	.82
39. I have the opportunity to learn how abuse and other difficulties affect peoples' mental health.	.06	**.91**	−.06	−.04	.82
40. This program creates opportunities for me to learn how abuse and other hardships affect peoples' relationships.	−.02	**.86**	−.02	.11	.78
36. This program gives me opportunities to learn how abuse and other difficulties, affect peoples' physical health.	.07	.84	.00	.01	.78

(*Continued*)

Table 1. Continued

Item	*Factor analysis structure coefficients*				
	1	*2*	*3*	*4*	h^2
41. I have the option to learn about the connections between abuse and substance abuse.	.05	.78	.03	−.04	.66
45. I am learning more about how to handle unexpected reminders of the abuse and difficulties I have endured.	.05	**.73**	.18	.00	.75
42. This program supports me to learn different ways of dealing with feeling overwhelmed.	.10	.65	.10	.14	.72
31. I have opportunities to help other survivors of abuse in this program.	−.08	.01	**.88**	.04	.73
33. In this program, I have the opportunity to connect with others.	.06	.05	**.68**	.03	.55
30. I have the option to get support from peers or others who have had experiences similar to my own.	.27	.03	**.56**	.12	.66
2. The strengths I bring to my relationships with my children, my family, or others are recognized in this program.	−.01	.12	.06	**.76**	.70
3. Staff respect the strengths I have gained through my life experiences.	.15	.05	.09	**.73**	.82
1. Staff respect the strengths I get from my culture or family ties.	.12	.07	.06	**.64**	.63
4. Staff are interested in the knowledge and skills I bring to coping with the life difficulties I've been through.	.20	.11	.13	.64	.84
5. Staff respect the specific ways I've found to cope with feeling upset.	.27	.10	.05	.56	.73

Note. $N = 326$. TIP = trauma-informed practice. Principal axis factoring with oblique rotation. Bolded items and coefficients were retained in the final scale. 1 = Environment of Agency and Mutual Respect, 2 = Access to Information on Trauma, 3 = Opportunities for Connection, 4 = Emphasis on Strengths. The following 11 items were removed in analysis 1: 7 "Staff share some aspects of my background (e.g., nationality, ethnicity, etc.)"; 22 "Staff check on how I'm feeling during our interactions"; 28 "I have the opportunity to provide feedback at this program"; 29 "Staff are flexible"; 34 "Staff support peoples' connections to family and community"; 35 "I am discovering how common my feelings are, given the abuse and other hardships I have gone through"; 43 "If I need more help than staff can provide, they will help me find it"; 44 "Staff help people find ways to manage stress"; 46 "Staff encourage people to take care of themselves spiritually"; 47 "Staff encourage people to take care of themselves emotionally"; and 48 "Staff encourage people to take care of themselves physically"

Supplementary TIP Scales

After finalizing the main TIP scale, we conducted two separate EFAs on the items intended to assess cultural responsiveness and inclusivity and those items intended to assess parenting support.

Analysis 4: Cultural responsiveness and inclusivity EFA. We conducted a separate EFA on the nine items related to cultural responsiveness and inclusivity ($N = 173$). These items were unique in that they included an "I don't know" response option, which we coded as missing (Furr, 2011). The KMO value was .92 and Bartlett's test of sphericity was significant ($p < .001$). A scree plot revealed the items loaded on one factor (eigenvalue = 7.64), which accounted for 84.86% of the variance. Extraction communalities ranged from .72 to .90

Table 2. Factor Loadings and Extracted Communalities for Inclusivity and Parenting Scales (Analyses 4 and 5)

Item	*Factor analysis structure coefficients*	h^2
Analysis 4		
51. Staff respect peoples' sexual orientations and gender.	**.95**	.90
50. People's religious or spiritual beliefs are respected in this program.	**.93**	.86
55. Staff recognize that some people or cultures have endured generations of violence, abuse, and other hardships.	**.93**	.86
49. Peoples' cultural backgrounds are respected in this program.	**.92**	.86
54. Staff understand how discrimination impacts peoples' everyday experience.	**.92**	.84
53. Staff understand the challenges faced by people who are immigrants.	**.92**	.84
56. This program treats people who face physical or mental health challenges with compassion.	**.89**	.79
52. Staff understand what it means to be in my financial situation.	**.88**	.78
57. Staff provide support for people who must interact with potentially difficult systems (e.g., courts, police, housing, child protective services, public assistance).	.85	.72
Analysis 5		
58. I am learning more about how children react emotionally when they have witnessed or experienced abuse, and other hardships.	**.84**	.71
59. Staff help me explore how children's relationships can be affected by witnessing or experiencing abuse and other life difficulties.	**.89**	.79
60. I am learning more about how my own experience of abuse can influence my relationships with my children.	**.88**	.78
61. The program provides opportunities for children to get help dealing with the abuse and other hardships they may have experienced or been affected by.	**.80**	.65
62. Staff support me to strengthen my relationships with my children.	**.72**	.52

Note. Table displays two separate factor analyses: analysis 4 (Cultural Responsiveness and Inclusivity; 9 items; $N = 173$) and analysis 5 (Support for Parenting; 5 items; $N = 314$). The one factor in analysis 4 accounted for 84.86% of the variance with an eigenvalue of 7.64; the one factor in analysis 5 accounted for 74.91% of the variance with an eigenvalue of 3.74. Both analyses were conducted using principal axis factoring with oblique rotation. Bolded items and coefficients were retained.

and unrotated factor loadings ranged from .85 to .95. Cronbach's alpha for the scale was .98. To reduce the scale length, we removed the only item that would have kept the alpha the same if deleted. This left us with eight items (see Table 2 for final factor loadings). As a team, we agreed these eight items measured Cultural Responsiveness and Inclusivity (i.e., Inclusivity).

Analysis 5: Parenting support EFA. We also conducted a separate EFA on the five parenting support items ($N = 314$). These items were also unique in that they excluded anyone who was not a parent. The KMO value was .85 and Bartlett's test of sphericity was significant

($p < .001$). A scree plot revealed the items loaded on one factor (eigenvalue = 3.74), which accounted for 74.91% of the variance. Extraction communalities ranged from .52 to .79 and unrotated factor loadings ranged from .72 to .89 (see Table 2 for factor loadings). Cronbach's alpha was .92. The team concurred that these five items measured Support for Parenting (i.e., Parenting) and retained all items.

Validity and Reliability

We assessed construct validity in three ways: First, we examined scale factor loadings; second, we compared subscale scores to other relevant measures; and third, we explored the extent to which the TIP Scales were able to capture program level differences above and beyond individual level differences. As noted above, we confirmed that factor loadings were generally high for the main TIP.

Regarding comparing the subscale score to other measures included in our survey, we expected that scores on the WAI-SR and the CSQ-8 would be correlated with each of the TIP Scales. Indeed, we found that the WAI-SR was strongly correlated with Agency ($r = .77$), Information ($r = .57$), Connection ($r = .50$), Strengths ($r = .65$), Inclusivity ($r = .73$), and Parenting ($r = .52$), with all correlations significant at $p < .001$. Although all the subscales were related to WAI-SR, we expected subscales more related to alliance (e.g., Agency and Inclusion) would be more highly correlated with WAI-SR compared to other subscales (e.g., Parenting and Information). Indeed, we found that Agency was more strongly correlated with WAI-SR than were Information ($z = 6.83$, $p < .001$) or Parenting ($z = 6.90$, $p < .001$).

In addition, Inclusion was more highly correlated with WAI-SR than were Information ($z = 4.93$, $p < .001$) or Parenting ($z = 5.73$, $p < .001$). Likewise, the CSQ-8 was strongly associated with Agency ($r = .76$), Information ($r = .62$), Connection ($r = .51$), Strengths ($r = .62$), Inclusivity ($r = .72$), and Parenting ($r = .54$), with all correlations significant at $p < .001$.

Finally, we assessed the extent to which the TIP subscales could indeed capture *program level* differences in practice. To do so, we conducted a series of random effects analyses of variance with agencies with at least 20 participants ($n = 9$) for each subscale. We first examined the F values to assess whether there were significant differences in subscale scores *across programs*. We then calculated the intraclass correlation coefficient (ICC) for each subscale to measure the proportion of variance between agencies relative to the variance among individuals (i.e., higher values indicate higher variability in the subscale score across agencies). Results indicated that there were significant differences in all subscale scores across agencies as follows: Agency, $F(8, 293) = 11.90$, $p < .001$, ICC = .24; Information, $F(8, 293) = 3.93$, $p < .001$, ICC = .08; Connection, $F(8, 292) = 5.22$, $p < .001$, ICC = .56; Strengths, $F(8, 288) = 8.83$, $p < .001$, ICC = .19; Inclusivity, $F(8, 286) = 12.59$, $p < .001$, ICC = .26; Parenting, $F(8, 272) = 2.28$, $p < .05$, ICC = .04.

Further, the fact that all subscales had an ICC above zero suggests that not all variability was due to individual differences; instead, components of the agency could explain some portion of the variability. This was most strongly the case for Connection, followed by Inclusivity, Agency, Strengths, Information, and Parenting. Suggestively, scores were significantly higher in smaller programs that accommodated longer stays than in larger programs whose residents were allowed to stay for shorter time periods. Means and standard deviations for each TIP scale can be found in Table 3.

Table 3. Correlations Among TIP Subscales

	M (SD)	*1*	*2*	*3*	*4*	*5*	*6*
1. Agency	2.56 (0.70)	—					
2. Information	2.45 (0.80)	.64***	—				
3. Connection	2.41 (0.81)	.57***	.55***	—			
4. Strengths	2.47 (0.83)	.75***	.51***	.44***	—		
5. Inclusivity	2.64 (0.67)	.82***	.59***	.51***	.66***	—	
6. Parenting	2.36 (0.82)	.44***	.65***	.42***	.39***	.49***	—

Note. TIP = trauma-induced practice; m = mean; SD = standard deviation. 1 = Environment of Agency and Mutual Respect; 2 = Access to Information on Trauma; 3 = Opportunities for Connection; 4 = Emphasis on Strengths; 5 = Cultural Responsiveness and Inclusivity; and 6 = Support for Parenting.
*** $p < .001$.

The factor analysis itself provided evidence of discriminant validity among main TIP subscales, with correlations ranging from .35 to .70, indicating that there was less than 50% of shared variance among the factors. Although our final correlation table (Table 3) revealed that the Agency and Strengths subscales were highly correlated ($r = .75$), we believe these factors are distinct in that when part of the factor analysis, their values were less than .7 in the factor correlation matrix and did not have any cross-loaded items. We also observed that the Agency and Inclusivity subscales were highly correlated ($r = .82$). This makes conceptual sense given that both measure aspects of the relationship between program staff and participants. Still, we think that they assess such different aspects of those relationships that we believe both subscales should be used as part of the overall package of scales comprising the TIP–DV Scales.

Reliabilities for the subscales ranged from .85 to .98, indicating generally high reliability across factors. Reliabilities for the 51 Spanish surveys were between .70 and .96, whereas the English surveys were between .86 and .98, indicating that the TIP Scales are generally consistent across the two languages.

DISCUSSION

Although *TI services* has become a buzz phrase in many human service and advocacy programs across the country, the extent to which programs have successfully translated TI principles into everyday practice remains unclear. Advocates, researchers, and policymakers are eager to learn whether programs have made this translation successfully, identify areas of relative strength and weakness, and investigate the relationship between specific TI practices and survivor well-being. Until now, those questions have remained unanswered given the absence of a way to measure TI practices. We set out to fill this gap through a university–community collaboration that included national experts, survivors, advocates, and programs.

A series of exploratory factor analyses based on 370 survivors seeking support from DV programs and an assessment of the resultant scales' relationships with other relevant variables supported their reliability and validity. The TIP Scales (available from first author upon request) are grounded in both theory and practice and assess six evidence-supported domains of trauma informed practice. The scales are meant to be used together because each one contributes uniquely to the comprehensive assessment of TI practices across diverse program contexts.

The structure of the main TIP and the two supplementary scales (together called the TIP Scales) that emerged from the EFAs correspond remarkably well with existing literature on TI practice in the DV context, capturing each of its main dimensions in a parsimonious way. This correspondence may be easiest to see in a comparison of the TIP Scales with the principles that emerged in a prior qualitative analysis of publications that delineate TI practices for use by DV programs (Wilson, Fauci, & Goodman, in press).

Specifically, each of the six TIP subscales corresponds to one of the six principles that emerged from that analysis, as follows: The factor Agency overlaps substantially with the qualitatively derived principle Restoring Choice and Control. Scores on this factor indicate the extent to which survivors feel that the program and its staff respect their agency and autonomy by offering opportunities for choice and control. The Information factor roughly corresponds to the Supporting Coping principle of the qualitative analysis. Scores reflect the extent to which survivors feel that their program offers information that increases their understanding of trauma and coping skills.

The Connection factor corresponds directly to the principle Facilitating Connection, with scores indicating the extent to which survivors perceive their program as one that creates opportunities for developing and strengthening mutually supportive relationships. The Strengths factor directly reflects the Building Strengths principle that emerged from the qualitative analysis. Scores on this subscale reflect the extent to which survivors perceive that their program and its staff explicitly recognize and value the unique strengths that they bring from their family, culture, relationships, and life experiences.

The Inclusivity Scale corresponds directly to the Responding to Identity and Context principle in the qualitative analysis. Scores on this factor reflect the extent to which survivors regard staff as understanding of and responsive to various aspects of their identity, including culture, religion, sexual orientation, socioeconomic status, and immigration status. Finally, the Parenting Scale maps onto a particular category within the principle Facilitating Connection, called staff support survivors' parenting relationships. Scores on this scale reflect the degree to which survivors feel the program helps them strengthen their relationships with their children through support and education. Overall, the correspondence between the TIP Scales and prior conceptualizations of TI practice in the DV context indicate that the scales have strong face validity.

The TIP also demonstrated satisfactory psychometric properties. Turning first to the main TIP, the communalities of the items were adequate (i.e., above .4; Costello & Osborne, 2005), and the factor structure was distinctive, as reflected by a lack of equally high loadings across factors. The construct validity of the main TIP was also excellent: As expected, each subscale was related to a widely used measure of client satisfaction with services as well as with a measure of the alliance between client and staff.

Moreover, those subscales that were more substantively related to the construct of alliance (e.g., Agency and Inclusion) were more highly correlated with the alliance measure than other subscales (e.g., Parenting and Information). Similar to the main TIP, our two supplementary scales of Inclusivity and Parenting also had adequate commonalities, reliability values, and high factor loadings. These two scales were also positively associated with participants' satisfaction with services and alliance with staff. Finally, for all six scales, participants within a program were significantly more likely to endorse the same items than were participants from different programs, suggesting that the TIP Scales assess a program-level phenomenon rather than simply a dyadic relationship between a particular survivor and a particular staff member.

Several methodological limitations should be considered when interpreting study findings. First, although the sample was somewhat ethnically diverse (primarily including

White, Black, and Latina participants) and was drawn from a wide variety of programs in the Northeast and Midwest, we were not able to recruit from other regions of the country with more and different diversity. Second, the small number of Spanish-speaking survivors in the sample made it impossible to conduct a separate EFA with this subsample.

In addition, we found that the 32 individuals omitted from the final analysis (because of having at least one missing value on a response) were more likely to have taken the Spanish version of the survey and to have been born outside of the United States. This study needs to be replicated with non-U.S.-born and non-English-speaking groups to better understand the reasons for skipped items and further build an inclusive measure of TI practice. It is important to note, however, that most participants who completed the Spanish version and were not born in the United States did complete the study.

Third, the sample was entirely female and almost entirely heterosexual. The extent to which the TIP Scales would be useful for men or gay, lesbian, bisexual, transgender and/or queer survivors is unclear, although when constructing them, we attempted to use language that would apply to all gender and sexual identities. Despite these limitations, it is worth noting that the demographics of this sample (including race/ethnicity, immigration status, and educational background) are quite similar to those of a national sample obtained in a study of community-based domestic violence programs (Lyon, Bradshaw, & Menard, 2011), suggesting that the scales may well be generalizable to survivors in residential and community-based DV programs.

Regarding the internal validity of the TIP Scales, although we emphasized to participating programs the need to recruit as many participants as possible without regard to length of time in or feelings about the program, it is possible that some program staff may have recruited participants with whom they felt most comfortable. Further, although we stressed to participants that no one at their agency would see their survey responses, some may nevertheless have felt pressure to describe the program positively. For these reasons, scores on the TIP may have been higher than they would be in other contexts, perhaps influencing the factor structure. Future studies using the TIP Scales should work to ensure the broadest possible representation of program participants. Finally, 370 participants was an insufficient amount for conducting a confirmatory factor analysis for cross-validation. For this reason, the findings presented here represent preliminary analyses only.

Implications for Program Evaluation and Practice

Many human service and advocacy-based organizations are committed to providing TI services to their participants, and DV programs are no exception. Some DV programs have a long history of using a TI lens in their work, while others are newer to the concept. Regardless of where on the continuum of TI practice such organizations lie, until now none has had a tool for examining, from the survivor's perspective, the extent to which they are truly engaging in the type of practice to which they aspire. The TIP Scales were intentionally created to be used easily by community programs so that they can (a) identify areas of relative strength and weakness, (b) improve their practices, (c) demonstrate to funders and other key stakeholders that they are incorporating TI principles, and (d) begin to understand the ways in which TI practice is related to survivor outcomes.

Implications for Research

The development and validation of the TIP Scales also enhances researchers' ability to explore a range of new questions and, most importantly, enable research on the long-term

effect of TI practices on survivor outcomes. Despite a growing emphasis on defining and delivering TI practice, almost no empirical evidence exists that documents how such an approach affects DV survivors' engagement with services or their long-term safety, mental health, and overall well-being. Longitudinal research using the TIP Scales could address this question. Moreover, larger scale validation studies of the TIP Scales with participants from a diverse range of social backgrounds would deepen our understanding of what such practices look like and mean to different groups of survivors. Finally, although this study focused on DV programs in particular, the TIP Scales do not specifically mention any particular form of trauma and could be applicable to a variety of human service and advocacy organizations.

Conclusion

This study was designed to fill a critical gap in our ability to assess the extent to which DV programs are engaging in TI practice. Preliminary evidence indicates that the TIP Scales are in fact theoretically grounded and ecologically valid tools that assess core aspects of survivors' experience in DV programs. We hope that DV programs will use the Scales to ensure that their very good intentions translate into practices that promote health and healing for the survivors who deserve our very best.

REFERENCES

Attkisson, C. C., & Greenfield, T. K. (2004). The UCSF Client Satisfaction Scales: I. The Client Satisfaction Questionnaire-8. In M. Maruish (Ed.), The use of psychological testing for treatment planning and outcome assessment (3rd ed., pp. 799–811). Mahwah, NJ: Erlbaum Associates.

Bennice, J. A., Resick, P. A., Mechanic, M., & Astin, M. (2003). The relative effects of intimate partner physical and sexual violence on post-traumatic stress disorder symptomatology. Violence and Victims, 18, 87–94.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical Assessment, Research & Evaluation, 10(7), 1–9. Available at: http://pareonline.net/getvn.asp?v=10&n=7

DeCandia, C. J., Guarino, K., & Clervil, R. (2014). Trauma-informed care and trauma specific services: A comprehensive approach to trauma intervention. Waltham, MA: The National Center on Family Homelessness.

Dillon, G., Hussain, R., Loxton, D., & Rahman, S. (2013). Mental and physical health and intimate partner violence against women: A review of the literature. International Journal of Family Medicine, 1–15.

Elhai, J. D., Miller, M. E., Ford, J. D., Biehn, T. L., Palmieri, P. A., & Frueh, C. (2012). Posttraumatic stress disorder in DSM-5: Estimates of prevalence and symptom structure in a nonclinical sample of college students. Journal of Anxiety Disorders, 26(1), 58–64.

Elliott, D. E., Bjelajac, P., Fallot, R. D., Markoff, L. S., & Reed, B. G. (2005). Trauma-informed or trauma-denied: Principles and implementation of trauma-informed services for women. Journal of Community Psychology, 33(4), 461–477.

Fabrigar, L. R., Wegener, D. T. MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. Psychological Methods, 4(3), 272–299.

Florsheim, P., Shotorbani, S., Guest-Warnick, G., Barratt, T., & Hwang, W. C. (2000). Role of the working alliance in the treatment of delinquent boys in community-based programs. Journal of Clinical Child Psychology, 29(1), 94–107.

Furr, M. (2011). Scale construction and psychometrics for social and personality psychology. Thousand Oaks, CA: SAGE Publications Ltd.

Hair, J., Black, W., Babin, B., & Anderson, R. (2010). Multivariate data analysis (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Harner, H., & Burgess, A. W. (2011). Using a trauma-informed framework to care for incarcerated women. Journal of Obstetric, Gynecologic, & Neonatal Nursing: Clinical Scholarship for the Care of Women, Childbearing Families, & Newborns, 40(4), 469–476.

Harris, M. E., & Fallot, R. D. (2001). Using trauma theory to design service systems. San Francisco: Jossey-Bass.

Hatcher, R. L., & Gillaspy, J. A. (2006). Development and validation of a revised short version of the Working Alliance Inventory. Psychotherapy Research, 16(1), 12–25.

Hopper, E. K., Bassuk, E. L., & Olivet, J. (2010). Shelter from the storm: Trauma-informed care in homelessness services settings. The Open Health Services and Policy Journal, 3(2), 80–100.

Kubiak, S. P. (2005). Trauma and cumulative adversity in women of disadvantaged social location. American Journal of Orthopsychiatry, 75, 451–465.

Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: Development of a general scale. Evaluation and Program Planning, 2, 197–207.

Lyon, E., Bradshaw, J., & Menard, A. (2011). Meeting survivors' needs through non-residential domestic violence services and supports: Results of a multi-site study. Washington, DC: Office of Justice Programs.

Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. Journal of Consulting and Clinical Psychology, 68(3), 438–450.

Melbin, A., Smyth, K. F., & Marcus, S. (2014). Domestic and Sexual Violence Cohort Demonstration Project: Taking action to move from services to social change. Greenfield, MA: The Full Frame Initiative.

Morrissey, J. P., Jackson, E. W., Ellis, A. R., Amaro, H., Brown, V. B., & Najavits, L. M. (2005). Twelve-month outcomes of trauma-informed interventions for women with co-occurring disorders. Psychiatric Services, 56(10), 1213–1222.

Perilla, J. L., Serrata, J. V, Weinberg, J., & Lippy, C. A. (2012). Integrating women's voices and theory: A comprehensive domestic violence intervention for Latinas. Women & Therapy, 35(1-2), 93–105.

Phillips, H., Lyon, E., Warshaw, C., & Fabri, M. (2013). Promising Practices and Model Programs: Trauma-informed approaches to working with survivors of domestic violence and other trauma. National Center on Domestic Violence, Trauma, and Mental Health.

Radford, L., Corral, S., Bradley, C., & Fisher, H. L. (2013). The prevalence and impact of child maltreatment and other types of victimization in the UK: Findings from a population survey of caregivers, children and young people and young adults. Child Abuse & Neglect, 37, 801–813.

Raubenheimer, J. (2004). An item selection procedure to maximize scale reliability and validity. South African Journal of Industrial Psychology, 30(4), 59–64.

Stark, E. (2007). Coercive control: How men entrap women in personal life. New York: Oxford University Press.

Substance Abuse and Mental Health Services Administration. (2014). SAMHSA's Concept of trauma and guidance for a trauma-informed approach. HHS Publication No. (SMA) Rockville, MD: Substance Abuse and Mental Health Services Administration.

Tabachnick, B. G., & Fidell, L. S. (2001). Using multivariate statistics. Boston: Allyn and Bacon.

Warshaw C., (2014). Thinking about trauma in the context of domestic violence: An integrated framework. Synergy, A Newsletter of the Resource Center on Domestic Violence Child

Protection and Custody, National Council of Juvenile and Family Court Judges. Available at: http://www.ncjfcj.org/sites/default/files/2410_NCJFCJ_Synergy_17_1_Final_Pages.pdf

Warshaw, C., Lyon, E., Phillips, H. P., & Hooper, M. (2014). Mental health and substance use coercion survey: Report on findings from the National Center on Domestic Violence, Trauma & Mental Health and the National Domestic Violence Hotline 2014. Available at: http://www.nationalcenterdvtraumamh.org/wp-content/uploads/2014/10/NCDVTMH_NDVH_MHSUCoercionSurveyReport_2014-2.pdf

Warshaw, C., Sullivan, C., & Rivera, E. (2013, February). A systematic review of trauma-focused interventions for domestic violence survivors. Retrieved from: http://www.nationalcenterdvtraumamh.org/wpcontent/uploads/2013/03/NCDVTMH_EBPLitReview 2013.pdf

Wilson, J. M., Fauci, J. E., & Goodman, L. A. (in press). Bringing trauma-informed practice to domestic violence programs: A Qualitative analysis of current approaches. American Journal of Orthopsychiatry.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research a content analysis and recommendations for best practices. The Counseling Psychologist, 34(6), 806–838.