The American Journal on Addictions, 27: 210–216, 2018

ISSN: 1055-0496 print / 1521-0391 online
DOI: 10.1111/ajad.12709

A Randomized Controlled Trial of a Gender-Focused Addiction Model Versus 12-Step Facilitation for Women Veterans

Lisa M. Najavits, PhD[1,2] **Justin Enggasser, PhD,**[1,2] **Deborah Brief, PhD,**[1,2] **Edward Federman, PhD**[3]

[1]Veterans Affairs Boston Healthcare System, Boston, Massachusetts
[2]Boston University School of Medicine, Boston, Massachusetts
[3]Edith Nourse Rogers Memorial Veterans Hospital, Boston, Massachusetts

Background and Objectives: Substance use disorder (SUD) has increased among women, including military veterans, yet SUD treatment was historically designed for males. This randomized controlled trial compared 12 individual sessions of a gender-focused SUD recovery model, *A Woman's Path to Recovery* (WPR) to an evidence-based, non-gender-focused SUD model, *12-Step Facilitation* (TSF) for 66 women veterans with current severe SUD.
Methods: The primary outcome was substance use; secondary outcomes were associated problems (e.g., psychological); coping skills, and 12-step attendance, with assessment at baseline, end-of-treatment, and 3-month followup.
Results: Substance use decreased over time, with no difference between conditions. Decreases occurred from baseline to end-of-treatment and baseline to followup and, for drug severity, also from end-of-treatment to followup. Effect sizes were large for alcohol and medium otherwise. Secondary outcomes were largely consistent with this pattern of improvement. Urinalysis/breathalyzer supported self-report. Treatment attendance was 62% for WPR and 57% for TSF (not significantly different). Twelve-step group attendance, surprisingly, did not increase in either condition.
Discussion and Conclusions: WPR provides a useful addition to women's SUD treatment options, with outcomes no different than an established evidence-based model, TSF. Both showed positive impact on substance use and related areas. Our lack of differences based on gender-focus may reflect women veterans being acculturated to a male military environment. Limitations include lack of an untreated control, a sample limited to veterans, and use of a large effect size for power assumptions.
Scientific Significance: This is the first RCT of a gender-focused approach for women veterans with SUD. (Am J Addict 2018;27:210–216)

Received July 18, 2017; revised February 7, 2018; accepted March 3, 2018.

Address correspondence to Lisa M. Najavits, Veterans Affairs Boston Healthcare System, 150 South Huntington Ave Boston, MA 02130. E-mail: najavits@bu.edu

INTRODUCTION

Women's rate of substance use disorder (SUD) has increased over the past several decades[1,2] and women military veterans, especially younger ones from the Iraq-Afghanistan cohort, have notable rates of hazardous drinking, binge drinking, and SUD.[3,4]

Moreover, compared to men, women suffer greater impact from SUD[5,6]: they are more likely to die from alcohol use disorder, for example, attempt suicide related to it, and have physical health problems from it. There is also some research indicating that they may have a *telescoped course*, a more rapid progression from use to disorder than men.[1]

Yet SUD treatments have historically been designed for males, who have higher rates of SUD.[1] Lack of childcare, transportation problems, and less encouragement to enter treatment are also barriers for women.[6] There have long been calls for more gender-focused SUD treatment but limited research on it.[6–12] Two randomized controlled trials (RCTs) using treatment manuals both found no difference between a gender-focused versus non-gender focused comparison.[13–15,a] We could not find any clinical trial of gender-focused treatment in a sample of women veterans.

A Woman's Path to Recovery (WPR) arose as a self-help workbook for women with SUD[16] and then was piloted as a clinician led group model, with positive outcomes.[17] The current RCT was designed to compare it to an evidence-based but not gender-focused model, *12-Step Facilitation* (TSF).[18] TSF performed just as well among women as it did among men in a major multisite trial, Project MATCH.[9] We thus viewed it as a legitimate comparison for our study.

[a]Messina et al. had an error in covarying out baseline characteristics that did not differ between groups, and without which there were no difference in outcomes between the treatment conditions.[15] We are thus interpreting the results based on the non-covaried outcomes.

We hypothesized that: (1) *WPR would outperform TSF on SUD* based on the long history of theoretical formulations on the need for gender-based SUD treatment[7,8]; (2) *TSF would increase AA attendance more than WPR* as TSF is specifically designed to increase AA attendance whereas WPR is not; and (3) *WPR would increase coping skills more than TSF* as WPR provides explicit education in coping skills whereas TSF does not.

METHOD

Participants

Patients were 66 female military veterans aged 18–65 who met DSM-IV criteria for current SUD and used a substance within 90 days of baseline. The study was begun while DSM-IV was still current and thus we did not change to DSM-5 when the latter was available in 2013 as the SUD criteria changed substantially between these two DSM editions and we did not want to create a potential difference in inclusionary criteria. Exclusion criteria were: (a) pregnant or planning pregnancy; (b) acute medical condition that could affect attendance or psychological functioning; (c) currently symptomatic bipolar I or psychotic disorder (as these would require referral for medication or other immediate intervention for safety reasons), or mental retardation, or organic mental disorder (as these could impact ability to understand the treatment materials in this study); (d) a primary treater's advice that participation was contraindicated; (e) dangerousness (eg, recent serious assault); (f) mandated to treatment; (g) upcoming major psychopharmacology change. Each of these exclusionary criteria were necessary because otherwise the data would be overly biased in a positive direction (criterion F); not able to capture patients' typical psychiatric presentation (criteria A, B, G); was unfeasible (criterion D); or had serious psychiatric issues in need of stabilization and/or our treatments were not appropriate (criteria B, E). We did not exclude for suicidality, self-harm, personality disorder, drug use disorder, homelessness, or recent substance use as excluding for these real-world vulnerabilities would have biased our sample toward more positive outcomes.

See Figure 1 for participant flow and assessment completion rates.

Procedures

Participants were recruited via study staff and flyers from anywhere in the greater urban area. They were not recruited from a particular specialty clinic or referral service of any kind. They did not have to be VA patients as our project was solely a research study and not embedded in any type of standard VA care. Block randomization was generated by the study statistician to achieve equal numbers of participants per condition by the end of the study and was implemented by non-blind staff. Study treatments were manualized and delivered individually in 12 weekly 1-hour sessions by three female study clinicians (two psychologists, one social worker). The clinicians were solely study clinicians and did not have regular clinical roles at the VA. To avoid a confound of therapist effects by treatment type, all therapists learned and administered both study conditions. We verified as part of each assessment that there was no contamination, that is, participants in each condition did not receive the other treatment during the study.

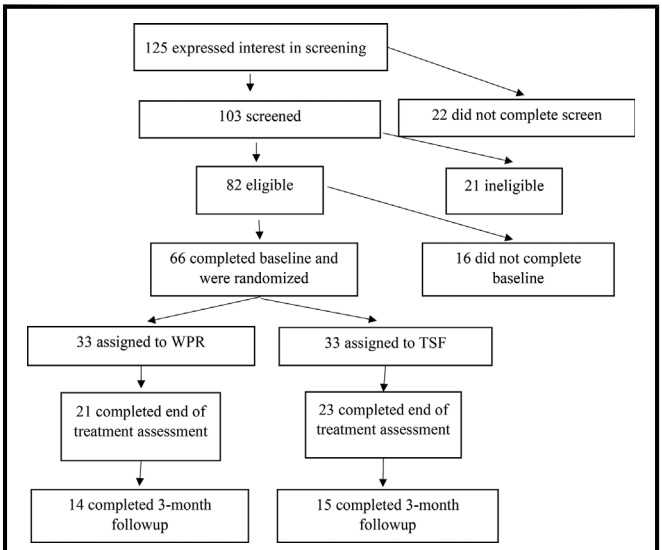

FIGURE 1. Flow of participants.

WPR is a clinician-led program for women with SUD using *A Woman's Addiction Workbook.*[16] The workbook provides psychoeducation on why a gender-based approach may be helpful, and how to identify addiction problems and co-occurring mental health disorders. The book then offers two major sections. *Exploration* identifies gender differences in addiction and recovery, barriers to treatment faced by women, and subgroups of women at risk for addiction. It highlights five key themes: body and sexuality, stress, relationships, trauma, and violence, and thrill-seeking. The *healing* section offers recovery methods in four domains–relationships, beliefs, actions, and feelings. WPR is based on cognitive-behavioral therapy, interpersonal, and emotive methods, and gender-focused theory of women's SUD (how SUD development and recovery differs for males and females).

The TSF condition used the manual for that treatment from Project MATCH.[18] It is "grounded in the concept of alcoholism as a spiritual and medical disease. The content of this intervention is consistent with the 12 Steps of Alcoholics Anonymous (AA) with primary emphasis given to Steps 1 through 5. In addition to abstinence from alcohol, a major goal of the treatment is to foster the patient's commitment to participation in AA" (pg. x). TSF is structured and enhanced with readings from the AA literature.

In both treatment conditions homework was suggested but not required. For example, in WPR homework might include reading an additional chapter from the WPR book. In TSF homework might include journal writing or reading from the AA literature.

Training in each study treatment occurred via a combination of reading, group discussion and roleplays, listening to

samples of each other's audiotaped sessions, and feedback on fidelity by the study leaders. All of these efforts occurred throughout the trial.

Measures

Assessments occurred at baseline, end-of-treatment, and 3-month followup. Our primary outcomes were the *Addiction Severity Index-Lite (ASI)*[19] alcohol and drug composites, which measure problem severity for alcohol/drugs, and the *Brief Addiction Monitor* use subscale,[20] which measures past 30 days level of substance use. Urinalysis/breathalyzer testing occurred weekly and baseline SUD diagnoses were obtained on the *Mini Neuropsychiatric Interview,*[21] a structured clinical interview for DSM-IV that was also used for exclusionary diagnoses.

Additional measures were the *Global Severity Index* from the *Brief Symptom Inventory*,[22] which measures intensity of mental health problems, scaled 0-4; the *AA Intention Measure,*[23] which measures attitudes and intentions toward attending 12-step groups, scaled 1–7, the *BASIS-24* total score, which measures psychiatric functioning and symptoms, and the substance abuse subscale), scaled 1-5; the *Brief Addiction Monitor* 12-step attendance item, which measures number of days attendance at 12-step meetings, rated 0–30 days[20]; the *Coping Scale*, which measures a variety of positive coping strategies, rated 1-5; and the *Addiction Severity Index-Lite (ASI)*[19] legal, employment, family/social, psychological, and medical composites, which measure addiction-related problems in those domains.

Fidelity was assessed for randomly selected audiotapes of full sessions, using the Adherence-Competence Scale for WPR[24] and the Yale Adherence and Competence Scale-II for TSF (YACS-II).[25] The former was adapted from the Seeking Safety Adherence Scale[26] with content changed to reflect the WPR model. It has 15 items and is scaled 0–3 with separate scores for adherence (0 = not done to 3 = done thoroughly) and competence (0 = harmful to 3 = extremely helpful); 2 or higher is considered adequate. The YACS-II measure is publicly available and was scaled 1 (not at all) to 7 (extensively); 4 or higher is considered adequate.[25] From the YACS-II we used the nine TSF-specific items; five assessment items; and four general support items. Each therapist was reviewed until she was found adequate on the respective scale, and could not treat study patients until then. After that, to avoid fidelity drift, at least one session of each therapist was reviewed monthly, rotating among patients within caseloads in group supervision throughout the trial. No therapist was "redlined,"[27] that is, fell below the adequate level on the relevant scale during the trial.

Data were collected between 2012 and 2015. Assessments were collected by research assistants hired for this study except for the *Mini Neuropsychiatric Interview*, which was conducted by a team member with a formal advanced degree in mental health. Patients received $40 for each assessment plus $3 for weekly urinalysis/breathalyzer. The primary outcomes were obtained by blinded raters. On all measures, higher scores reflect worse pathology unless noted otherwise. Participants were assessed at all timepoints regardless of treatment attendance, that is, we used an intent-to-treat design.

Analyses

Power calculation to determine sample size was based on the WPR pilot study[17] and a comparable sample of women from another treatment trial.[28] We used the ASI drug composite score as the basis for the power analysis. For 95% significance and 80% power, assuming different levels of intracluster correlation coefficient, the sample size for each arm to detect an effect size of .8 (end of treatment vs. baseline) are: 31 for ICC of .05, 32 for ICC of .10, and 33 for ICC of .15. We selected the highest of these (33 per group). The .8 level is for a large effect and often used for clinical trials.[29,30]

See Table 1 for baseline sample descriptive statistics and two-tailed independent samples *t*-tests or chi square tests to compare by study condition. We tested whether all outcome measures were missing completely at random (MCAR) using Little's test, which was non-significant for all outcome variables. We used multiple imputation procedures in SPSS version 24 to address missing data, using the pooled result of 100 imputations. In a clinical trial with a difficult-to-follow population such as ours, multiple imputation is widely recommended.[31] Imputed data were analyzed with two-way mixed analysis of variance (ANOVA) on the intent-to-treat

TABLE 1. Participant characteristics at baseline

	Women's path to recovery mean (SD) or percent	12-step facilitation mean (SD) or percent	Test statistic
Age	46.97 (9.92)	48.45 (8.13)	$t = .64$, df = 59, $p = .52$
Gender—female	100%	100%	not applicable
Education: 4-year college or higher	29%	26.7%	$\chi2 = .04$, df = 1, $p = .84$
Race/ethnicity: any minority	25.8%	26.7%	$\chi2 = .01$, df = 1, $p = .94$
Married	6.5%	13.3%	$\chi2 = .81$, df = 1, $p = .37$
Employed (full or part-time)	30%	36.7%	$\chi2 = .30$, df = 1, $p = .58$
Current substance dependence			
Alcohol	76.9%	74.2%	$\chi2 = .06$, df = 1, $p = .81$
Any drug	45.5%	56%	$\chi2 = .52$, df = 1, $p = .47$

sample; see Table 2. All continuous variables passed Levine's test of homogeneity of variance. Mauchley's test of sphericity was conducted and if significant, Greenhouse-Geisser estimates were used. Effect sizes are reported as partial eta-squared.

RESULTS

Participant Characteristics

There were no differences between conditions on baseline characteristics. See Table 1. Breakdown by ethnicity was Hispanic 3%, Caucasian non-Hispanic 89.4%, and 7.6% missing. Race was Caucasian 68.2%, African-American 18.2%, American-Indian or Alaskan Native 4.5%, multiracial 1.5%, missing 7.6%; we reported by the full sample as there was no difference between treatment conditions. For SUD, most (94.7%) met criteria for a current substance dependence (alcohol or drug); the few who did not did meet criteria for substance abuse (alcohol and/or drug). Of the 51.1% of the full sample who met for any drug dependency, the breakdown was 54.2% cocaine, 12.5% each cannabis or amphetamines, 8.3% opioids, 4.2% each heroin, methadone (illegally taken), or "other"; we reported by the full sample as there was no difference between treatment conditions.

Attendance

The average number of sessions, from a maximum of 12, was 7.40 for WPR (SD = 3.87) and 6.87 for TSF (SD = 3.75), indicating that for both conditions, participants attended a majority of sessions at 61.7% of available sessions for WPR and 57.2% for TSF, with no difference between them ($t = -.54$, df = 59, $p = .59$).

Outcomes

See Table 2.

Primary Outcomes

All three primary outcomes evidenced improvement over time, with no difference between conditions. To understand the main effect for time we analyzed pairwise differences. All three measures showed improvement from baseline to end-of-treatment and baseline to 3-month followup. On the ASI drug composite there was also improvement from end-of-treatment to 3-month followup. Effect sizes were large for the ASI alcohol composite and medium for the other two measures.

Our ASI results are supported by the weekly urinalysis/breathalyzer testing. The majority showed concordance (biological testing consistent with ASI): 13% were positive on both and 49% negative on both. When discordance occurred, moreover, it was non-problematic, that is, honest self-reporting: negative urinalysis/positive self-report. We presume these discordances occurred because breathalyzers and urinalysis, depending on the drug, have a limited timeframe. Also there were no negative consequences for patients' reporting substance use.

Secondary Outcomes

Overall the secondary outcomes show impact on a broad array of domains—*coping, employment, psychological problems/global severity, medical problems,* and *psychiatric functioning*. The pattern of results for the first three was consistent with the primary outcomes: improvement over time (baseline to end of treatment, and no deterioration at 3-month followup), medium effects sizes, and no difference between conditions. The *medical* composite had a similar pattern except for showing impact primarily at 3-month followup; and the measure of *psychiatric functioning* (BASIS-24) showed positive impact at all timepoints, including improvement from end-of-treatment to 3-month followup.

One secondary outcome had a somewhat aberrant pattern. The ASI family/social composite showed the same improvement from baseline to end-of-treatment as the other measures, but worsening from end-of-treatment to followup. One measure, the ASI legal composite, could not be computed due to lack of variance (low legal involvement in the sample).

Finally, and surprisingly, the two 12-step measures (attendance and intent to attend) showed no improvement over time and no difference between conditions, indicating that whatever in the treatments had positive impact, it was not by increased interest in or attendance at 12-step groups.

DISCUSSION

This RCT evaluated two SUD models in a sample of women military veterans: *A Woman's Path to Recovery* (WPR) and *12-Step Facilitation* (TSF), an evidence-based model that encourages 12-step group involvement. We hypothesized that the gender-focused approach, WPR, would outperform the non-gender-focused TSF. We also expected that TSF would lead to greater 12-step attendance and that WPR would increase coping skills more than TSF. None of these predictions was borne out.

Both treatments showed consistent positive impact on multiple domains, including our primary outcome, substance use, as well as secondary measures of medical, employment, and psychological problems; coping; and psychiatric functioning. Effect sizes were medium to large. There were no differences between conditions on any measure. This finding of no difference among active, manualized treatments is a pattern highly typical of behavioral therapy RCTs for alcohol use disorder[32] as well as a robust, consistent finding dating back many decades in the general psychotherapy field.[33]

It was heartening to find that WPR did as well as TSF (or in statistical terms, "no worse than" TSF). TSF has a strong evidence base, whereas WPR had previously been studied only in one small pilot.[17] WPR may be an especially important option for women who have tried 12-step and decided against it. Some of the women in our study said, "I'm sick of people telling me to go to AA" or "I went and it didn't work for me." Yet other women reported great benefit from 12-step groups. We can note, however, that in our sample, women in both

TABLE 2. Intent-to-treat outcomes for a woman's path to recovery (WPR) versus 12-step facilitation (TSF)

Measure	Baseline *M (SD)* [or % if noted]		End of treatment *M* (SD) [or % if noted]		3-month follow-up *M* (SD) [or % if noted]		*Treatment* x Time *F* (*df*) *p* η_p^2 effect size	*Treatment F* (*df*) *p* η_p^2 effect size	Time *F* (*df*) *p* η_p^2 effect size	Comparisons for significant effects
Primary outcomes	WPR	TSF	WPR	TSF	WPR	TSF				
Addiction severity index:	.36	.46	.18	.26	.23	.21	1.95 (1.69,87.88)	1.20 (1,52) $p = .28$	**25.29 (1.69,87.88) $p = .0005$**	**Time a < .0005 b < .0005**
alcohol composite	(.24)	(.35)	(.21)	(.24)	(.19)	(.11)	$p = .15$ $\eta p^2 = .036$	$\eta p^2 = .023$	**$\eta p^2 = .327$ (large)**	c = .87
Addiction severity index:	.10	.10	.04	.06	.04	.04	.17 (1.26,65.34)	.42 (1,52) $p = .52$	**12.24 (1.26,65.34) $p < .0005$**	**Time a=.004 b < .0005**
drug composite	(.11)	(.12)	(.05)	(.06)	(.05)	(.05)	$p = .74$ $\eta_p^2 = .003$	$\eta_p^2 = .008$	**$\eta_p^2 = .191$ (medium)**	**c = .049**
Brief addiction monitor:	4.61	4.48	1.96	2.33	2.33	1.92	.37 (1.31,78.69)	.01 (1,60) $p = .91$	**18.73 (1.31,78.69) $p < .0005$**	**Time a <.0005 b <.0005**
substance use subscale	(3.50)	(3.85)	(2.08)	(3.17)	(1.92)	(2.36)	$p = .60$ $\eta_p^2 = .006$	$\eta_p^2 = .000$	**$\eta p^2 = .238$ (medium)**	c = .94
Secondary outcomes										
Brief addiction monitor:	1.25	1.28	1.26	1.31	1.57	1.60	.002 (1.38,82.66)	.01 (1,60) $p = .90$	2.35 (1.38,82.66) = .12	Not applicable
12-step attendance item	(1.24)	(1.65)	(1.20)	(1.27)	(1.13)	(1.35)	$p = .99$ $\eta_p^2 = .000$	$\eta_p^2 = .000$	$\eta p^2 = .038$	
AA intention measure	5.21	5.26	4.90	4.97	4.99	4.97	.05 (2,110) $p = .95$	.02 (1,55) $p = .89$	2.24 (2,110) $p = .11$ $\eta p^2 = .04$	Not applicable
	(1.27)	(1.10)	(1.42)	(1.34)	(.98)	(1.07)	$\eta_p^2 = .001$	$\eta_p^2 = .000$		
BASIS-24: total score	36.22	38.05	33.40	34.12	29.80	29.57	.26 (1.43,85.94)	.22 (1,60) $p = .64$	**13.85 (1.43,85.94) $p < .0005$**	**Time a = .036 b < .0005**
	(12.32)	(10.63)	(9.15)	(6.90)	(7.68)	(6.31)	$p = .69$ $\eta_p^2 = .004$	$\eta_p^2 = .004$	**$\eta p^2 = .190$ (medium)**	**c < .0005**
BASIS-24: substance abuse	5.92	6.46	4.84	4.54	4.38	4.22	.40 (1.54,92.19)	.002 (1,60) $p = .97$	**7.64 (1.54,92.19) $p = .002$**	**Time a =.019 b < .0005**
subscale	(3.83)	(4.33)	(3.03)	(2.97)	(3.49)	(2.55)	$p = .62$ $\eta_p^2 = .007$	$\eta_p^2 = .000$	**$\eta p^2 = .113$ (medium)**	c = .289
Coping scale	2.25	2.92	2.88	2.95	2.85	2.92	.07 (1.75,104.96)	.06 (1,60) $p = .80$	**11.72 (1.75,104.96) $p < .0005$**	**Time a <.0005 b <.0005**
	(1.09)	(.40)	(.99)	(1.01)	(.69)	(.40)	$p = .91$ $\eta_p^2 = .001$	$\eta_p^2 = .001$	**$\eta p^2 = .163$ (medium)**	c = .81
Addiction severity index:	.64	.59	.58	.54	.58	.56	.41 (2,106) $p = .66$	.21 (1,53) $p = .64$	**7.86 (2,106) $p = .001$ $\eta_p^2 = .129$**	**Time a < .0005 b=.004**
employment composite	(.27)	(.34)	(.29)	(.33)	(.28)	(.32)	$\eta_p^2 = .008$	$\eta_p^2 = .004$	**(medium)**	c=.461
Addiction severity index:	.38	.43	.38	.33	.21	.20	1.55 (1.53,80.91)	.001 (1,53) $p = .98$	**23.83 (1.53,80.91) $p < .0005$**	**Time** a=.156 **b < .0005 c**
medical composite	(.31)	(.38)	(.29)	(.30)	(.28)	(.28)	$p = .22$ $\eta_p^2 = .028$	$\eta_p^2 < .0005$	**$\eta_p^2 = .31$ (large)**	**< .0005**
Addiction severity index:	.43	.48	.37	.41	.31	.39	.50 (1.71,90.90)	1.39 (1,53) $p = .24$	**12.34 (1.71,90.90) $p < .0005$**	**Time a=.002 b < .0005**
psychological composite	(.21)	(.18)	(.21)	(.18)	(.21)	(.19)	$p = .58$ $\eta_p^2 = .009$	$\eta_p^2 = .026$	**$\eta_p^2 = .189$ (medium)**	c=.110
Addiction severity index:	.20	.29	.14	.22	.23	.30	.28 (1.64,85.22)	2.50 (1,52) $p = .12$	**6.41 (1.64,85.22) $p = .005$**	**Time a = .017** b = .459
social/family composite	(.21)	(.22)	(.15)	(.19)	(.23)	(.24)	$p = .71$ $\eta_p^2 = .005$	$\eta_p^2 = .046$	**$\eta_p^2 = .110$ (medium)**	**c < .0005**
Brief symptom inventory:	1.55	1.85	1.37	1.34	1.62	1.26	**5.95 (1.65,99.27)**	.88 (1,60) $p = .88$	**6.98 (1.65,99.27) $p = .003$**	**Time a = .002 b = .018**
global severity index	(1.21)	(1.12)	(1.14)	(.92)	(1.01)	(.88)	**$p = .006$ $\eta_p^2 = .09$**	$\eta_p^2 = .000$	**$\eta_p^2 = .104$ (medium)**	c = .241

Higher scores indicate worse pathology on all measures except the Brief Addiction Monitor 12-step attendance item, AA Intent Scale, the Coping Scale, on which higher scores are better. Effect size (partial eta squared, η_p^2) are interpreted as .01 (small), .09 (medium), and .25 (large) per Watson40; they are reported to three decimal places to avoid misinterpretation due to rounding. "a" refers to baseline compared to end of treatment. "b" refers to baseline to 3-month follow-up. "c" refers to end of treatment to 3-month followup. Significant results at $p < .05$ are bolded. SPSS p values of ".000" are listed as <.0005.

conditions attended 12-step groups a mean of less than 2 days per month, from before the trial through 3-month followup, with no increase over time. Future research would benefit from further understanding of how prior exposure to 12-step and/or prior therapy may have had an impact on their response to our treatment conditions. It would also be interesting to explore the degree to which elements of 12-step groups that are gender-based (women-only groups and same-sex sponsorship) may impact recovery.

There may be various reasons why we did not find a result based on the gender-focused content of the WPR treatment versus the non-gender focused content of TSF. Both of our treatments were delivered by women therapists, which may have diminished the gender impact of WPR. Also structural barriers to treatment, which are known to accrue for women, such as childcare and transportation, were addressed equivalently in that all participants received the same payment for completing assessments, which may have helped them overcome these barriers to a degree. In general, it is widely recommended that helping women address such instrumental issues may be an important facet of gender-focused care.

We can further note that women military veterans are different from non-veterans in being generally less traditionally gender-focused by virtue of having volunteered for the extremely male-dominated culture that the military represents.[34] Military service is physically demanding, emotionally depriving, and sometimes dangerous. Some women specifically enter it out of a desire to escape traditional gender norms, although for others the reasons may include economics, lack of opportunities, or growing up in a military family. Whatever the motivations for any particular woman, however, once she enters the service her experience differs dramatically from her female counterparts in civilian life, generally in a direction that diminishes a traditional female-focus.[34]

The WPR model was not specifically adapted for women veterans and that is something to potentially consider in future trials. Such adaptation might include a focus on how military culture responds to feminine gender norms and how women change as a result of being in the military. It could also include the strong culture of alcohol use in the military, which has been documented by the Institute of Medicine.[4] Put simply by Katrina, an Army veteran:

"Drinking was all but encouraged. We had beer machines in the barracks. In the movies, you could drink. Even though there were rations for some stuff, you could get as much alcohol as you wanted· · ·And it wasn't seen as a problem. . .. It was only a problem if you had clear and obvious consequences, like not showing up for duty or getting into a big fight or a DUI" (pgs. 132–133).[35]

Despite not being specific to women veterans, the women in WPR stated that its content felt relevant to them, especially if they had experienced gender discrimination or military sexual trauma in the military.

This is the first RCT of a gender-focused approach in women veterans. Strengths of our design include the minimal exclusionary criteria designed to obtain a real-world sample; power analysis; intent-to-treat design; multiple imputation to address missing data; blind raters; fidelity monitoring; and validated measures. Limitations include all of those that accrue to an RCT, including highly trained clinicians with ongoing monitoring, patients who were willing to be randomized and assessed, payment for assessments, and limits on generalizability due to exclusions such as being mandated to treatment or having current bipolar 1 disorder or a psychotic disorder. We also had only one followup point after the end of treatment. There were also missing assessments at end of treatment (33%) and 3-month followup (56%), although we mitigated this by the use of multiple imputation. Our power analysis assumed a large effect size; future studies may benefit from a more conservative assumption. Also our fidelity ratings were limited to ongoing random monitoring rather than a quantitatively reported result. Interpretation of study results would have benefitted from a third study condition such as an untreated sample or TAU-alone to evaluate naturalistic outcomes without any study-specific intervention being offered. However in VA an untreated sample is not possible as women with SUD are routinely offered treatment as there is a strong focus within VA to outreach to veterans and strive to provide as much care as possible; also TAU in the current era often includes various evidence-based approaches and thus has limitations as a control. Our study results should be interpreted as reflecting how the two models fared in the context of VA. Also replication could be useful given recent literature on the need to replicate social science studies.[36]

WPR has strong promise given this first RCT on it. Our results also further validate the TSF model. Yet questions remain: Which clients are most likely to benefit from each? What characteristics of therapists and settings matter for successful outcomes? What training is necessary? What amount of change per measure would be clinically significant? How would the models compare outside of VA? How would the models do if delivered in group format? Further research is warranted and larger trials with longer followup periods would be optimal.

This research was supported by grant #NEUA-001-08S to Dr. Najavits from the Department of Veterans Affairs, Clinical Sciences Research and Development, Washington, DC. It was registered at www.clinicaltrials.gov (NCT00734903). We thank Joni Utley, PsyD, Kay Johnson, LICSW and Helen Tretiak-Carmichael, MA, for their work on this project and Dan Berlowitz, MD, for his instrumental support of it. We also thank Molly Waring, PhD and Matthew Gallagher, PhD, for their consultation on statistical analyses.

Declaration of Interest

Dr. Najavits is the author of *A Woman's Addiction Workbook,* which served as the basis for the Women's Path

to Recovery model, one of the study conditions, and receives royalties from New Harbinger Press for the book. Her potential conflict of interest for this study was reviewed and cleared in advance of the trial by the Department of Veterans Affairs Office of General Counsel. Drs. Enggasser, Brief, and Federman declare no competing interests.

REFERENCES

1. Greenfield SF, Back SE, Lawson K, et al. Substance abuse in women. *Psychiatr Clin North Am*. 2010;33:339–355.
2. Kendall-Tackett KA, Ruglass LM. *Women's Mental Health Across the Lifespan: Challenges, Vulnerabilities, and Strengths*. New York, NY: Taylor & Francis; 2017.
3. Bradley KA, Bush KR, Epler AJ, et al. Two brief alcohol-screening tests from the alcohol use disorders identification test (AUDIT): validation in a female veterans affairs patient population. *Arch Intern Med*. 2003;163:821–829.
4. IOM (Institute of Medicine). *Substance Use Disorders in the U.S. Armed Forces*. Institute of Medicine of the National Academies; 2012. 9780309262019.
5. Wooten NR, Mohr BA, Lundgren LM, et al. Gender differences in substance use treatment utilization in the year prior to deployment in Army service members. *J Subst Abuse Treat*. 2013;45:257–265.
6. Brady K, Back S, Greenfield S. *Women and Addiction: A Comprehensive Handbook* New York, NY: Guilford Press 2009.
7. Gomberg E, Nirenberg T, eds. *Women and Substance Abuse*. Norwood, New Jersey: Ablex Publishing; 1993.
8. Covington SS, Surrey J. The relational model of women's psychological development: implications for substance abuse. In: Wilsnack R, Wilsnack S, eds. *Gender and Alcohol: Individual and Social Perspectives*. New Brunswick, NJ: Rutgers Center of Alcohol Studies; 1997:335–354.
9. Del Boca FK, Mattson ME. The gender matching hypothesis. *Project MATCH hypotheses: Results and causal chain analyses*. 2001;8:118–131.
10. DeVito EE, Babuscio TA, Nich C, et al. Gender differences in clinical outcomes for cocaine dependence: randomized clinical trials of behavioral therapy and disulfiram. *Drug Alcohol Depend*. 2014;145: 156–167.
11. Graff FS, Morgan TJ, Epstein EE, et al. Engagement and retention in outpatient alcoholism treatment for women. *Am J Addict*. 2009;18: 277–288.
12. Greenfield SF, Brooks AJ, Gordon SM, et al. Substance abuse treatment entry, retention, and outcome in women: a review of the literature. *Drug Alcohol Depend*. 2007;86:1–21.
13. Greenfield SF, Sugarman DE, Freid CM, et al. Group therapy for women with substance use disorders: results from the Women's Recovery Group Study. *Drug Alcohol Depend*. 2014;142:245–253.
14. Messina N, Grella CE, Cartier J, et al. A randomized experimental study of gender-responsive substance abuse treatment for women in prison. *J Subst Abuse Treat*. 2010;38:97–107.
15. Najavits LM, Hien DA. Helping vulnerable populations: a comprehensive review of the treatment outcome literature on Substance Use Disorder and PTSD. *J Clin Psychol*. 2013;69:433–479.
16. Najavits LM. *A Woman's Addiction Workbook*. Oakland, CA: New Harbinger 2002.
17. Najavits LM, Rosier M, Nolan AL, et al. A new gender-based model for women's recovery from substance abuse: results of a pilot outcome study. *Am J Drug Alcohol Abuse*. 2007;33:1–7.
18. Nowinski J, Baker S, Carroll K. *Twelve Step Facilitation Therapy Manual (Project MATCH): A Clinical Research Guide for Therapists Treating Individuals with Alcohol Abuse and Dependence*. Vol 1 Rockville, Maryland: National Institute on Alcohol Abuse and Alcoholism 1995.
19. McLellan T, Cacciola J, Carise D, et al. Addiction Severity Index-Lite-CF. 2005; http://www.tresearch.org/resources/instruments/ASI_Lite.pdf. Accessed (2/6/05).
20. McKay J, Drapkin M, Goodman J, et al. Brief Addiction Monitor. 2010; University of Pennsylvania, unpublished measure.
21. Sheehan D, Lecrubier Y, Harnett Sheehan K, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 1998;59:22–33.
22. Derogatis LR. The brief symptom inventory: an introductory report. *Psychol Med*. 1983;13:595–605.
23. Zemore SE, Kaskutas LA. Development and validation of the alcoholics anonymous intention measure (AAIM). *Drug Alcohol Depend*. 2009;104: 204–211.
24. Najavits LM. Adherence-Competence Scale for A Woman's Path to Recovery. Unpublished manuscript, Harvard Medical School and McLean Hospital 2011.
25. Nuro KF, Maccarelli L, Baker SM, et al. Yale adherence and competence scale (YACSII) guidelines. *West Haven, CT: Yale University Psychotherapy Development Center*. 2005:161.
26. Najavits LM, Liese BS. Seeking Safety Adherence Scale (version 3). Unpublished measure. Harvard Medical School/McLean Hospital, Boston, MA. 2003.
27. Carroll KM, Connors GJ, Cooney NL, et al. Internal validity of project MATCH treatments: discriminability and integrity. *J Consult Clin Psychol*. 1998;66:290–303.
28. Najavits LM, Weiss RD. Cognitive-behavior therapy for dual diagnosis women. NIDA RO1 DA-08631; 1999.
29. Cohen J. Statistical power analysis. *Curr Dir Psychol Sci*. 1992;1:98–101.
30. Röhrig B, du Prel J-B, Wachtlin D, et al. Sample size calculation in clinical trials: part 13 of a series on evaluation of scientific publications. *Dtsch Arztebl Int*. 2010;107:552.
31. Gewandter JS, McDermott MP, McKeown A, et al. Reporting of missing data and methods used to accommodate them in recent analgesic clinical trials: aCTTION systematic review and recommendations. *PAIN*®. 2014; 155:1871–1877.
32. Imel Z, Wampold B, Miller S, et al. Distinctions without a difference: direct comparisons of psychotherapies for alcohol use disorders. *Psychol Addict Behav*. 2008;22:533–543.
33. Lambert MJ. *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change*. Hoboken, NJ: John Wiley & Sons; 2013.
34. Maples S. The inconvenience of being a woman veteran. The Atlantci. https://www.theatlantic.com/politics/archive/2017/11/the-inconvenience-of-being-a-woman-veteran/545987/: downloaded 11/26/17; 2017.
35. Najavits LM. *Recovery from Trauma, Addiction or Both: Finding Your Best Self*. New York, NY: Guilford Press 2017.
36. Schmidt S. Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Rev Gen Psychol*. 2009; 13:90.