



Published in final edited form as:

J Groups Addict Recover. 2013 ; 8(4): . doi:10.1080/1556035X.2013.836867.

Group Process in the single-gender Women's Recovery Group compared with mixed-gender Group Drug Counseling

Shelly F. Greenfield, MD, MPH^{1,2}, Laura E. Kuper, BA³, Amanda M. Cummings, BA⁴, Michael S. Robbins, PhD⁵, and Robert J. Gallop, PhD⁶

¹McLean Hospital, Belmont, MA

²Harvard Medical School, Boston, MA

³University of Illinois at Chicago, Chicago, IL

⁴University of Miami School of Education and Human Development, Coral Gables, FL

⁵University of Miami Miller School of Medicine, Miami, FL

⁶West Chester University, West Chester, PA

Abstract

Enhanced affiliation among members is thought to provide increased support for women in single-gender compared with mixed-gender group therapy for substance use disorders (SUDs) and to provide a potential mechanism of action for its efficacy. In a Stage I trial of single-gender versus mixed-gender group therapy for SUDs we examined affiliative statements made by members in two group treatments, single-gender Women's Recovery Group (WRG) and mixed-gender Group Drug Counseling (GDC). Twenty-eight WRG and 17 GDC group therapy tapes were coded and compared for five types of affiliative statements. Three types of affiliative statements (agreement, supportive, and completing a thought) were highly correlated and were more frequent in WRG than GDC ($D=0.882$, $p=0.27$). In GDC, women were more likely to provide an affiliative statement to a male group member than any other combination of directionality ($p<0.01$). Compared with mixed-gender, single-gender group therapy for SUDs may enhance support through greater frequency of affiliative statements.

Keywords

Treatment for women; Women-only programs; Mixed-gender programs; Treatment outcomes; Group Therapy

1. Introduction

Over the past two decades considerable research has highlighted differences between women and men with substance use disorders (SUDs) as well as their substance abuse treatment needs and participation (Greenfield, Brooks, et al., 2007; Greenfield & Grella, 2009; Prendergast, Messina, Hall, & Warda, 2011). These gender differences include the initiation of substance use, progression to substance use problems and dependence, adverse medical and social consequences, and treatment entry and retention (Greenfield, Brooks, et al., 2007; Randall et al., 1999). In response to these gender differences, women-only

treatment programs as well as women's treatment services within mixed-gender programs have been established. These programs differ from standard mixed-gender treatment programs in providing services to women such as prenatal care; childcare services; transportation; assistance with housing and job or vocational training (Greenfield, Brooks, et al., 2007; Grella, Polinsky, Hser, & Perry, 1999; Hodgins, El-Guebaly, & Addington, 1997; Orwin, Francisco, & Bernichon, 2001; Oser, Knudsen, Staton-Tindall, & Leukefeld, 2009; Volpicelli, Markman, Monterosso, Filing, & O'Brien, 2000). Such services can enhance women's substance abuse treatment outcomes (Oser et al., 2009; Volpicelli et al., 2000). Moreover, studies that have shown that all women's treatment programs may enhance outcomes for women with substance use disorders above and beyond the effects of mixed-gender treatment programs (Grella et al., 1999; Prendergast et al., 2011).

In addition to studies of the effectiveness of gender-specific programs for women's treatment outcomes, investigators have examined manual-based group therapies for women with specific clinical or other characteristics for which a specialized group therapy might enhance outcomes (Hien, Cohen, Miele, Litt, & Capstick, 2004; Hien et al., 2009; Linehan et al., 1999; Luthar & Suchman, 2000). For example, studies have examined the effectiveness of manual-based group therapy for women with co-occurring substance use and post-traumatic stress disorders (Hien et al., 2004; Hien et al., 2009); for methadone-maintained parenting women (Luthar & Suchman, 2000); for women with co-occurring borderline and substance use disorders (Linehan et al., 1999); among others.

However, treatment programs and clinicians in community practice often need to treat women with SUDs who are heterogeneous with respect to substances of abuse, demographic characteristics such as age and stage of life (e.g., transitional age, mid-life, pregnant or parenting, etc.), co-occurring other psychiatric disorders (e.g., depression, anxiety, post-traumatic stress disorders, etc.), as well as with other characteristics. Little is known about the effects of such heterogeneity on clinical processes or outcomes in general and more specifically with respect to women in treatment. Group therapy is the standard format most commonly used in substance abuse treatment programs (Kominars et al., 2004), yet there is a dearth of research assessing effectiveness of group therapy of SUDs, especially with regard to a heterogeneous group of patients with SUDs. Research is needed to elucidate the clinical processes of these treatments in order to help guide clinical practice.

The Women's Recovery Group (WRG) was designed as a manual-based, 12-session, weekly, relapse prevention group that would be tailored for women with SUDs who are heterogeneous with respect to clinical, life-stage, and demographic characteristics (Greenfield, Trucco, McHugh, Lincoln, & Gallop, 2007). In a Stage I therapy development trial, the newly developed WRG was compared to an effective, manual-based, mixed-gender control condition, Group Drug Counseling (GDC). Results from this pilot study showed that women assigned to both WRG and GDC reduced their days of substance use during the 12-week treatment phase with no significant differences between them; however, in the 6 month post-treatment phase, women assigned to WRG continued to reduce their days of use while women assigned to GDC showed no further reductions (Greenfield, Trucco, et al., 2007). Moreover, secondary analyses of these data demonstrated that women with high psychiatric severity at baseline (Greenfield et al., 2008) and low self-efficacy (Cummings, Greenfield, & Gallop, 2010) who were assigned to WRG had better treatment outcomes in the six months following treatment compared with women with high psychiatric severity and low self-efficacy assigned to mixed-gender GDC.

The WRG study hypothesized that both the single-gender group composition and the women-focused content would together result in better substance abuse treatment outcomes than the mixed-gender control condition (Greenfield, Trucco, et al., 2007). Previous

literature on groups suggests that the gender composition of a group may have important effects on group member behavior and group process (Brown & Mistry, 1994; Moss, 2004; Shapiro, 1990). Theorists have regarded groups as a social microcosm of the wider society, in that without proper intervention, they reproduce the dominant status and power relations (Brown & Mistry, 1994). Others have focused on group participants' specific frames of reference; the socially determined lens through which participants read each other as familiar or unfamiliar or higher or lower in status, for example. In this view, the ways in which participants' frames fit together has great influence on the resultant group dynamics, and it becomes the therapist's job to actively construct an alternative environment of empowerment for group members (Shapiro, 1990). In practice, many therapists do not openly discuss gender-related issues within the context of mixed-gender groups, despite the negative consequences such reluctance might cause (Moss, 2004).

Research has also shown that women receiving treatment for SUDs often prefer women-only treatment groups and programs and report perceptions of enhanced comfort and safety in single-gender treatment, as well as reporting negative experiences in mixed-gender treatment (Kauffman, Dore, & Nelson-Zlupko, 1995). The WRG study hypothesized that the all-women group environment would result in not only increased comfort, but also increased supportiveness and openness among members. However, we know of no previous study that has investigated the *in-session differences in process* between single-gender and mixed-gender therapy groups.

While individual and family therapy literature have a rich base of findings regarding processes that are associated with successful and unsuccessful outcomes, less is known regarding the in-session processes in group therapy, and how these might serve as mechanisms of action. Process research has traditionally focused on documenting intervention parameters such as the timing, intensity, duration, and specific content targets of the intervention (Clarke, 1995; Sandler, West, Baca, & Pillow, 1992; Wolchik, West, Westover, & Sandler, 1993). More recent studies of group cognitive behavioral therapy (CBT) have shifted from assessing the effectiveness of CBT to understanding mechanisms of change (Oei & Browne, 2006). Recent studies of group therapy for mood and anxiety disorders have examined group processes such as alliance (Bakali, Wilberg, Hagtvet, & Lorentzen, 2010; Crowe & Grenyer, 2008; Joyce, Piper, & Ogrodniczuk, 2007), group cohesiveness (Crowe & Grenyer, 2008; Joyce et al., 2007; Marmarosh, Holtz, & Schottenbauer, 2005; Oei & Browne, 2006; Schmalisch, Bratitotis, & Muroff, 2010), mutual aid among group members (Schmalisch et al., 2010), and social contact and networking (Schmalisch et al., 2010). However, few of these studies have examined group processes and their relationship to treatment outcome (Crowe & Grenyer, 2008; Joyce et al., 2007; Toren & Shechtman, 2010). Understanding the linkages between in-session clinical interactions and outcomes remains a crucial component to the delivery of effective interventions. To establish these links, research must move beyond a description of program parameters or moderators of treatment outcome to careful investigations of the clinical interior of the group and group processes that are associated with patient outcomes. Such investigations will be critical in developing more effective interventions and in communicating specific recommendations to clinicians (Clarke & Rowan, 2009).

Despite the relative lack of empirical research, theoretical attention has consistently reinforced the significance of in-session mechanisms of actions for group therapy. Of these, the concept of group cohesion is one of the most widely studied to date (Burlingame, McClendon, & Alonso, 2011). In his work on the theory and practice of group psychotherapy, Yalom (Yalom, 1985, 1995) placed special emphasis on group cohesion as a crucial therapeutic factor. He defined members of a cohesive group as those who are "... accepting of one another, supportive, and inclined to form meaningful relationships in the

group” (Yalom, 1985). Subsequently, researchers have described group cohesiveness as a form of connectedness among members, which is evidenced by group members constructively and openly working together toward common therapeutic goals (Budman, Soldz, Demby, & Feldstein, 1989). Clinically, group cohesion is conceptualized as a dynamic process evidenced by high levels of positive and supportive interactions, self-disclosure, and expressed concern among group members (Budman et al., 1989). One recent study supported the theory that group cohesiveness was directly related to collective self-esteem and hope for the self (Marmarosh et al., 2005).

However, the theoretical attention afforded to such within-session mechanisms of action has highlighted not only their importance, but also their relative complexity. Those who have attempted to operationalize or measure group cohesion have done so in a variety of ways, and no clear consensus currently exists. Within psychology, Lewin and Festinger were some of the first to attempt to measure group cohesion, initially describing it as the field of forces which act on members to remain in the group (Festinger, Schachter, & Back, 1950). Early work typically focused on cohesion as interpersonal attraction. At that time, researchers attempted to measure cohesion by attendance (Piper, 1984) or eye contact with other speakers (Flowers, Booraem, & Hartman, 1981).

Despite theoretical advances, much of the research conducted on group cohesion relies on rating scales (Drescher, Burlingame, & Fuhriman, 1985). Such studies have produced mixed findings, likely due in part to the multidimensional nature of cohesion coupled with differences in measurement technique (Burlingame et al., 2011). Nonetheless, they offer important insight into the role cohesion may play. Taube-Schiff and colleagues (2007) found that group cohesion ratings were related to improvement in social anxiety symptoms, generalized anxiety, depression, and functioning among individuals participating in a cognitive-behavioral group for social phobia (Taube-Schiff, Suvak, Antony, Bieling, & McCabe, 2007). Cohesiveness, as well as other group process measures such as self-disclosure and feedback, was also related to outcomes (Tschuschke & Dies, 1994) for patients in inpatient psychotherapy groups. Marziali and colleagues (1997) found that while group cohesion and group therapeutic alliance were correlated, both made independent contributions to outcomes measures, which included psychiatric symptoms, social adaptation, and behavioral dysfunction in patients with borderline personality disorder (Marziali, Munroe-Blum, & McCleary, 1997). However, in their sample of individuals receiving group treatment for SUDs, Gillaspay and colleagues (2002) found that while group cohesion and alliance were correlated, only alliance predicted reductions in self-reported psychological distress, and neither was associated with drug and alcohol use consequences (Gillaspay, Wright, Campbell, Stokes, & Adinoff, 2002). A meta-analysis of 40 studies examined the relationship between treatment outcome and group cohesion and found that cohesion was related to outcome and that the moderator variables such as age, theoretical orientation, length, and size of group were associated with outcome (Burlingame et al., 2011). Other research has found little relationship between cohesion and outcomes yet have highlighted the importance of related constructs such as expressiveness (Oei & Browne, 2006).

While the majority of studies have examined various conceptualizations of group cohesion, researchers have also identified additional therapeutic relationship constructs such as group climate, which represents a sense of constructive and empathic interpersonal interactions among group members (Johnson, Burlingame, Olsen, Davies, & Gleave, 2005). However, much like group cohesion and therapeutic alliance, empirical studies suggest that these various constructs are also overlapping, and all are hypothesized to provide a climate that fosters additional, beneficial therapeutic processes (Johnson et al., 2005). Given that the term cohesiveness has lacked definitional clarity, and measures of cohesion rely on self-

reported ratings by participants and therapists on their perceptions of cohesion, researchers such as Hornsey and colleagues (2007) have argued that that the field should instead focus on identifying more *specific and measurable components of group process* (Hornsey, Dwyer, & Oei, 2007). Some efforts have been made to move the field in this direction. For example, researchers have linked turn-taking patterns to measurements of closeness and conflict among group members (Pincus & Guastello, 2005).

The aim of the current study was to investigate observable, in-session group processes by examining the frequency of a range of statements made by group members that represent supportive, cohesive, or empathic connections among members of the treatment group. We have termed such verbal expressions as "affiliative statements" and we hypothesize that the frequency of giving and receiving such statements by group members is a measurable in-session process that might operationalize the supportive, cohesive, or affiliative experience of a treatment group by participants. We developed a comprehensive coding method to capture the frequency of these statements as one type of marker for therapeutic environment and then examined types and frequency of these affiliative statements in the single-gender WRG group compared with the mixed-gender GDC control group. More specifically, we hypothesized that these types of statements may vary as a function of the gender composition of the groups. We hypothesized that there would be a greater frequency of affiliative statements present in single-gender WRG compared with mixed-gender GDC. In addition, within the mixed-gender group, we hypothesized that women would more frequently verbalize an affiliative statement.

The aims of this study were to (1) define and categorize affiliative statements made in the context of substance abuse group therapy; (2) examine the frequency and type of affiliative statements in single-gender WRG compared with mixed-gender GDC; and (3) examine gender differences in frequency and directionality of affiliative statements in mixed-gender GDC.

2. Materials and Methods

2.1 Group Interventions in the Stage I Trial

The Women's Recovery Group (WRG) is a 12-session manual-based relapse prevention group therapy that utilizes a cognitive-behavioral approach developed as part of a Stage I therapy development trial (Greenfield, Trucco, et al., 2007). The WRG study hypothesized that both the single-gender group composition and women-focused content would result in better outcomes than standard mixed-gender group treatment. A detailed description of the study design and WRG intervention are available (Greenfield, Trucco, et al., 2007). In order to control for patient-therapist gender matching, all therapists in the pilot study were women (Greenfield, Trucco, et al., 2007). Content areas for individual sessions were based on gender-specific substance abuse antecedents, consequences, and treatment outcomes (Greenfield, Brooks, et al., 2007). Group Drug Counseling (GDC; Crits-Christoph et al., 1999) is an empirically supported group treatment delivered typically to mixed-gender groups with no explicit attention to gender-specific issues. Both groups used a 90 minute structured format and covered common recovery topics such as triggers to use, relapse prevention, use of self-help, among others. While the WRG contained women-focused content, GDC gave no explicit attention to gender-specific issues. Groups used a semi-open format and members were added each week until a target enrollment was achieved (approximately 8 members maximum) and continued until the 12 week sequence was complete.

After obtaining approval from the McLean Hospital Institutional Review Board, written informed consent was obtained from all study participants. The study was first examined

feasibility and satisfaction by conducting two pre-pilot WRG groups with a total of 13 women subjects. Then the pilot RCT was conducted in two sequences with a total of 23 women subjects were randomized either to WRG (N = 16) or GDC (N = 7) and 10 male subjects were assigned to GDC (Greenfield, Trucco, et al., 2007). All sessions were videotaped and audiotaped so that therapist fidelity to the treatment models could be assessed (Greenfield, Trucco, et al., 2007) as well as group process. In order to protect confidentiality of the subjects, only the therapist was within the camera's view and group members sat outside camera range.

2.2 Baseline Characteristics of the Study Sample

Baseline characteristics of the female sample have been previously reported (Greenfield et al., 2008; Greenfield, Trucco, et al., 2007). In brief, the female sample was predominantly non-Hispanic White (e.g., one subject endorsed a Native American ethnicity) and alcohol dependent [86.2% (n = 25) of WRG and 100.0% (n = 7) GDC women]. The only statistically significant difference between WRG and GDC was the baseline characteristics was the mean age of the female group members (58.3 years, SD = 7.18 for GDC and 45.0 years, SD = 10.73 for WRG). Over one-third of the women participants were diagnosed with other current co-occurring psychiatric disorders, including mood disorders (i.e., 37.9% WRG and 42.9% GDC), anxiety disorders (i.e., 31.0% WRG and 42.9% GDC), eating disorders (i.e., 3.4% WRG and .0% GDC), other Axis I disorder (i.e., 3.4% WRG and .0% (GDC), and Axis II personality disorders (i.e., 34.5% all WRG and 57.1% GDC). Approximately three-fourths of participants were diagnosed with lifetime co-occurring psychiatric disorders including mood disorders (75.9% WRG and 85.7% GDC), anxiety disorders (i.e., 44.8% WRG and 42.9% GDC), eating disorders (i.e., 20.7% WRG and 0% GDC), and other Axis I disorders (i.e., 3.4% WRG and .0% GDC). None of these differences were statistically significant. We examined the potential impact of baseline differences that had not reached statistical significance on the intervention contrast in our initial analyses in two ways: (a) controlling for each separately as a covariate, and (b) using propensity score described in Rosenbaum (2002) as an additional covariate controlling for all baseline measures for which there was any imbalance. We saw our original results were not sensitive to these baseline differences through both approaches.

The mean age for men enrolled in GDC was 43.7 and all were non-Hispanic White and, as with the female sample, were well educated with 100% reporting some college or completing college or post-graduate studies. The majority had co-occurring psychiatric disorders (i.e., 70% current and lifetime mood disorders, 30% current and lifetime anxiety disorders, and 30% Axis II personality disorders). As with the female sample, the majority were 90% alcohol dependent and 10% cocaine dependent. There were no significant differences between men and women in the GDC groups for these clinical and demographic characteristics.

2.3 Development of Affiliative Statement Coding Manual

We developed a coding manual to capture the range of affiliative statements made by participants in the two group therapies studied, single-gender WRG and mixed-gender GDC. Two researchers (LEK, AMC) watched a randomly selected sample of WRG and GDC tapes, identifying statements that appeared supportive, cohesive, or empathic in nature. These types of statements were referred to as "affiliative statements." Using an iterative method of repeated coding with consultation from a process research expert (MSR) and the study's PI (SFG), these statements were grouped together by degrees of similarity, until several categories of affiliative statements could be operationalized and coded reliably. The two researchers coded 45 tapes during the process to establish inter-rater reliability. The method for coding the tapes as well as the affiliative statement categories were adjusted until

all researchers agreed that the categories of statements formed a cohesive construct and could be measured reliably based on calculations of inter-rater reliability. These strategies have been used to develop inter-coder reliability for identifying therapist and client behaviors in family therapy contexts (Newberry, Alexander, & Turner, 1991; Robbins, Alexander, & Turner, 2000). Ultimately, several categories were collapsed or eliminated which did not best capture the affiliative statement construct, or could not be reliably distinguished from other categories.

2.4 Content of Affiliative Statement Coding Manual

The five types of affiliative statements derived through this iterative coding process included: agreement statements, therapeutic responses, supportive statements, completing another member's thought, and positive statements regarding the group. A statement met criteria as an *agreement statement* when one participant acknowledged an experience of another group member and expressed sharing a similar experience. These had to be positive in nature and demonstrate understanding between two group members. An example of an agreement would be "Like she was saying, I have experienced problems with my marriage as a result of my drinking, too."

Therapeutic responses had to follow or be connected to a statement made by another member, and provide advice, guidance, or insight relevant to the issue presented by the other group member. This type of statement could involve sharing of strategies that a group member found helpful under similar circumstances, or providing a new way of approaching or thinking about a situation. Such a statement had to go beyond simple agreement to presenting a novel thought in response to the other participant. Questions that served the same purpose, including rhetorical questions, were also included in this category. To be included in this category, these statements needed to be directed at a particular group member, and were typically posed in an attempt to gain knowledge or allow the participant to explore a particular idea or perspective. For example, if in response to a group member's difficulty remaining abstinent from alcohol when dining out, another group member responded with "you should try going out to eat at restaurants that do not serve alcohol because you would not be tempted to drink with dinner," then this qualified as a therapeutic statement.

Supportive statements provided encouragement or positive feedback to another group member or the group as a whole. In order to be included in this category, these statements needed to reinforce a positive behavior that a group member had just described. "Great Job! You are making so much improvement in your recovery," is an example of a supportive statement.

A statement that was consistent with the category *Completing Thoughts* involved a group member completing the train of thought of another group member in a supportive manner. Typically this was observed during an interchange between group members who were relating and following along with each other. For example, Participant A said: "I was having a hard time making the commitment to treatment, and in order to get myself to meetings I just found myself having to say..."; then Participant B completed the thought: "this is what's best for my recovery"; and Participant A responded: "yes, it's true." Thus, this is just one example of an affiliative statement that would be categorized here.

Positive statements regarding the group were observed when a participant referenced the group as a whole and talked about what they found helpful or liked about the group. In order to be consistent with this category, the statement needed to reference information presented during sessions or the support received from the group. "I found that last week's topic really

helped me to take a look at my relationships and how they are impacting my recovery,” is an example of a statement in this category.

The specificity of the coding system was necessary for both raters to reliably capture affiliative statements consistent with these categories and to exclude statements that did not fit these criteria. Also, the raters followed a specific set of guidelines for identifying affiliative statements. The raters only coded statements made by participants in the group, not by the therapist leading the session. Statements could only be coded under more than one category if they could be split into phrases that independently met the criteria for each category. Statements were coded as soon as they met criteria for an affiliative statement. Therefore, we coded the incidence rather than duration of the statements. Subsequent continuation of the same thought or idea was coded only if it addressed a different theme and met criteria for another category.

2.5 Procedure for coding Affiliative Statements in group therapy sessions

Two researchers [LEK, AMC] coded group session tapes using Noldus Observer XT (Noldus, Trienes, Hendriksen, Jansen, & Jansen, 2000), a professional software package for the collection, analysis and presentation of observational data. Digitized tapes were analyzed directly on the computer utilizing the software such that affiliative statements were coded in real-time directly into the database. The software enabled the precise time within the session by minute and seconds when the coded statement was made to be entered into the database along with the category of statement. Forty-eight percent of all therapy sessions were included in the analysis; 28 of these sessions were WRG and 17 were GDC sessions. While all sessions were videotaped, certain tapes were excluded from the analysis because of technical recording quality such that they could not reliably be coded. Only those tapes in which the sessions were fully audible for at least 30 minutes of the session were included in the analysis. Sessions were randomly assigned a coding order as well as a designated first coder. The gender composition of the groups prevented the coders from being blind to condition. The coding procedure was as follows:

1. Coder 1 coded the tape initially, indicating the appropriate times at which affiliative statements were heard (according to the manual). In addition, Coder 1 coded 10 random times throughout the session in which a participant made a statement that did not meet criteria for any affiliative statement category. These statements were referred to as non-affiliative statements. The purpose of coding non-affiliative statements was to ensure inter-rater reliability in the accuracy of discriminating between affiliative and non-affiliative statements. For GDC tapes, the coder also indicated the gender of the speaker and the recipient of the affiliative statements.
2. Coder 2 then coded the same tape, using a list of times Coder 1 coded a statement without any other information regarding the category of affiliative statement specified by Coder 1. Coder 2 also coded any additional affiliative statements missed by Coder 1 but did not add any non-affiliative statements. Coder 2 also indicated the gender of the speaker and recipient of the affiliative statements for GDC cases.
3. If Coder 2 identified any statements not initially coded by Coder 1, Coder 1 revisited the tape with the list of times that Coder 2 identified a missed statement. Coder 1 coded these statements, but only if she felt the statement met the specific criteria for an affiliative statement.
4. Inter-rater reliability was calculated for each tape based on the number of agreements and disagreements in coding between Coders 1 and 2. The coders then discussed any discrepancies until they reached 100% agreement and this final

number of affiliative statements was used for analysis. Frequencies of statements for each tape were calculated.

2.6 Data Analysis

The agreement in ratings between the two coders for the number of affiliative and non-affiliative statements within session was assessed with a kappa coefficient, quantifying the agreement between two ratings over the raters, beyond chance (Kraemer & Bloch, 1988). For comparisons of the affiliative statements by categories between the two groups (i.e., GDC AND WRG) we implemented a two-level hierarchical linear model (HLM) (Raudenbush, Bryk, Cheong, & Congdon, 2000). The hierarchical structure of the data consists of affiliative statements coded for the multiple sessions within a specific group. There were four WRG groups each led by four different therapists and two GDC groups each led by a different therapist). Hence, the unit of the analysis is the affiliative statements within one of the six groups for a specific session. The hierarchical levels of clustering (i.e., multiple sessions within therapist within intervention arm) present in this data provide a unique problem for analyses, because: (a) repeated assessments across sessions run by a single therapist may be correlated (i.e. a group run by a particular therapist with high levels of affiliative statements by members participating in sessions early in a sequence could have had similarly high levels of affiliative statements during later sessions, whereas, a group run by a different therapist with low levels of affiliative statements by participating members in sessions early in the sequence may have had low levels of affiliative statements during later sessions; and (b) on-average scores may vary from group to group run by different therapists. When the within therapist correlation session to session, as well as the variability therapist to therapist is properly incorporated through the HLM statistical framework, the approach results in an increase of statistical power (Gibbins & Walker, 1993). We modeled the affiliative category measurement as outcome per session within each therapist's group as the level-1 model, which yields an average estimate per therapist within group. At the level-2 model, we use these therapist-specific estimates as outcome to derive overall averages per intervention arm (WRG versus GDC) as well as contrast the difference between intervention arms. The level-2 model includes a random therapist effect which accounts for the within-therapist correlation across the repeated sessions, as well as the between therapist differences in outcome (i.e., affiliative statements).

In addition, given the limited number of total tapes observed ($n=45$), we may have had inadequate power to find significant differences between groups. We derived Cohen's d effect sizes (Cohen, 1988), to provide an index of clinical significance, which are estimated as the mean difference per group divided by the pooled standard deviation of the outcome, where the pooled standard derivation is calculated as the square-root of the variance component of outcome for each assessment.

For comparisons of the directionality of each statement (GDC only), we implemented a three-level HLM model to account for the structure of the data, that is, directional statement category within session within therapist. In the previous analysis only one outcome was acquired per session-therapist combination. In this setting, 4 measures corresponding to the direction are acquired per session-therapist combination, which required the three-level structure. The level 1 model consisted of the within-session portion, where we observed the rate of the four distinct directions of the various affiliative statements: male to female, female to male, male to male, and female to female (Raudenbush et al., 2000). The level 2 model consists of the within-therapist portion. The level 3 model consists of the between therapist model. Statistical contrast within the three-level structure was performed to derive difference between the difference in directional statements.

Prior to running our two-level HLM model to address the difference between WRG and GDC, we needed to ensure there was no systematic bias in the outcome which may have resulted in differences between groups. Systematic bias might have been introduced because of the recorded duration of each session's tape which was variable due to tape quality, as well as the composition of each session. To address this issue, we ran a two-level HLM as described above, with elapsed session time as outcome. Estimated duration of times per group as a function of session were categorized as early, mid-treatment, and late treatment (indicating whether the group session occurred in the first 4 weeks, second 4 weeks, or final 4 weeks, respectively, of the treatment sequence). Recorded elapsed times appear to be consistent across the sessions, with a shorter recorded elapsed time for GDC compared to WRG that were not statistically different based on the two-level HLM (Overall Contrast $F(1,4)=2.17$, $p=0.22$; Early Contrast $F(1,6)=0.06$, $p=0.81$; Mid Contrast $F(1,6)=0.06$, $p=0.81$; Late Contrast $F(1,6)=2.00$, $p=0.21$). Interaction of the timing of the session in the course of treatment sequence and intervention also yielded a non-significant interaction, $F(2,6)=0.41$, $p=0.68$. In spite of these non-significant differences, we adjusted the rates of affiliative statement category by the elapsed time of each session. We also covaried the effect of session timing within the sequence (e.g., early, middle, and late) in our two-level HLM model.

3. Results

3.1 Inter-rater Reliability

The achieved Kappa coefficient of 0.808 with a range from 0.61–1.00 across the 45 sessions indicates inter-rater reliability that is in the substantial (beyond chance) to near perfect range (Landis & Koch, 1977).

3.2 Comparison of Frequency of Categories of Affiliative Statements between WRG and GDC

Because of the limited number of tapes and small sample size, we were unlikely to see statistically significant results in frequency of affiliative statements between groups; rather, we expected to observe medium to large effect sizes in the differences in frequency of affiliative statements between WRG and GDC groups. Through the two-level HLM, we found no significant difference between groups in the overall rate of *Affiliative Statements* ($F(1,4)=0.72$, $p=0.44$) with an effect size for the difference between groups of $d=0.41$.

However, we did find differences in the rate of some of the five categories of affiliative statements between groups. The effect sizes and p values for the comparison of rates of the five categories of affiliative statements in WRG and GDC are presented in Table 2. Positive effect sizes indicate higher values for WRG compared to GDC. Negative effect sizes indicate higher values for GDC compared to WRG. There was no difference in the frequency of four individual categories of affiliative statements, *Therapeutic Response*, *Positive Statements*, *Completing Thoughts*, and *Support Statements* between WRG and GDC. However, while the group difference in the rate of *Agreement* statements, was not statistically significant ($F(1,4)=5.13$, $p=0.086$), there was a large effect size, ($d=1.08$) in the difference between groups.

Because of the potential overlap among the categories of affiliative statements, we explored a composite scale of affiliative statements that might discriminate between WRG and GDC using two approaches. First, we examined the correlation among the 5 categories and found that three of the categories of affiliative statements, Agreement, Completing Thoughts, and Supportive statements, were highly collinear (Agreement with completing: $r=0.39$, $p<0.01$; Agreement with Supportive: $r=0.51$, $p<0.001$; Completing with supportive: $r=0.38$, $p<0.02$).

All correlation coefficients were above the large effect size threshold ($r=0.371$) (Cohen, 1988) therefore, representing a large inter-relationship. All remaining pairwise correlations were below 0.17. Second, despite the small sample size ($N=45$) we conducted a principal components analysis for the five categories (Ki and Chow, 1995). Kaiser's stopping rule (eigenvalue > 1) and a scree plot were used to determine the number of components (Rummel, 1970), which yielded two components, with Agreement, Completing Thoughts, and Supportive statements forming one dimension with weights of 0.58, 0.57, and 0.52, respectively, and Therapeutic Responses and Positive Statements forming a second dimension with weights of 0.63, and 0.75, respectively. Since our sample size is small, coupled with the dimension formation as exploratory, the scores per dimension were based on surrogate variables formed as the average of the individual categories comprising the respective dimension rather than use of factor scores as recommended by Hair et al. (1995). Analysis of the items comprising dimension two yielded only small effects ($d < 0.20$) for the intervention contrast; therefore, we formed a composite scale of the three components of the first dimension. For this composite scale, items comprising dimension 1, we found group difference was not statistically significant ($F(1,4)=1.61$, $p = 0.27$) but the difference had a large effect size ($d=0.882$). Individual item results for the three components, as well as the two items for dimension two are in Table 2.

3.3 Gender Differences and Directionality of Affiliation Statements in GDC

Prior to running the three-level HLM model, we derived the rate of the directionality of each affiliative statement category. Rates were determined by the gender of the individual talking, the gender of the individual they were talking to, the number of statements, and the elapsed time. Figure 1 shows the results for each individual affiliative statement category and Figure 2 shows the results for the Compositive Affiliative scale. It is evident from the figures that for each category of affiliative statement, as well as the Composite Affiliative Scale, the female to male direction is consistently larger in magnitude. To ensure that this effect was not biased due to session composition of males and females, as above, we covaried the main effect of each session in our three-level model. For the Composite Affiliative Scale, the female to male directional contrast was significantly larger compared to each of the other three directions (male to male: $t(48)=4.03$, $p=0.0002$, $d=1.16$; male to female: $t(48)=2.86$, $p=0.006$, $d=0.83$ female to female: $t(48)=2.85$, $p=0.006$, $d=0.82$). For the categories of therapeutic response and completing thoughts statements we find similar directional differences, where the female to male direction was significantly larger than any other direction. Therapeutic response contrasts are: male to male: $t(48)=4.29$, $p=0.0001$, $d=1.24$; male to female: $t(48)=3.13$, $p=0.003$, $d=0.90$ female to female: $t(48)=3.86$, $p=0.003$, $d=1.11$). Completing thoughts contrasts are: male to male: $t(48)=4.22$, $p=0.0001$, $d=1.22$; male to female: $t(48)=3.50$, $p=0.001$, $d=1.01$ female to female: $t(48)=3.81$, $p=0.004$, $d=1.10$).

4. Discussion

In order to capture verbal statements that are supportive or empathic responses between group members, we developed a taxonomy and method of coding "affiliative statements" to examine observable "in-session" group therapy process in single-gender and mixed-gender substance abuse group therapy conditions. "Affiliative" verbal behaviors between group members can be conceptualized as representing one important aspect of the therapeutic environment within groups that many have described as group cohesion.

We were interested in whether these observable affiliative behaviors among group members would differ between groups that were composed of all women compared with those that were composed of both men and women. While there were no overall differences between groups in the frequency of affiliative statements, we found that certain types of affiliative

statements occurred more frequently in the single-gender WRG than in mixed-gender GDC substance abuse therapy groups. These categories of affiliative statements generally reflected emotionally and empathically supportive statements. For example, the categories *positive statements and therapeutic response* included general positive statements about the group and advice-type statements. However, *agreement, completing thoughts, and supportive statements* included statements of identification and emotional understanding by one group member for another. These identification and emotional understanding affiliative statement types were highly correlated with one another (i.e., *Agreement, Completing Thoughts, and Supportive Statements*). When these statements were combined into one Composite Affiliative Scale, they occurred more frequently in the single-gender WRG than the mixed-gender GDC groups with a large effect size, although this difference was not statistically significant. These differences may reflect the difference between work- or task-oriented relationships and bonding or affective relationships (Burlingame, MacKenzie, & Strauss, 2004). While more structured groups such as cognitive-behavioral, may be more focused on accomplishing specific therapeutic goals, other groups focus more on fostering a sense of supportiveness and closeness among members (Johnson et al., 2005). Such a difference in therapeutic environment may be reflected in the greater frequency in WRG of affiliative statements that reflect identification and emotional support; these may be emblematic of the sense of enhanced support and intimacy reported by women in an all women's substance abuse group therapy format (Kauffman et al., 1995).

Within the mixed-gender GDC, we were interested not only in the rates of the affiliative statements overall, but also the frequency with which affiliative statements were made from women to men, men to women, women to women, and men to men. We found that in the mixed-gender group, the number of affiliative statements made by women to men was significantly greater than in each of the other possible directions. This suggests that men in mixed-gender groups may be more likely to experience a supportive group environment than women in mixed-gender groups. It is also not clear whether the bi-directionality of empathic statements present in the WRG group contributes to an increased sense of support for women in WRG compared with GDC and if this has a role in clinical outcomes.

Both WRG and GDC are manual-based group therapies. However, even when therapists are intensively trained to implement a standardized program using a detailed intervention manual, the specific in-session clinical interactions may have as much to do with outcome as the content of the intervention (Prinz, Miller, Peters, & McMahon, 1996). A review conducted by Burlingame, Fuhrman and Johnson (2001) found that the majority of studies of therapeutic relationships within groups focused on member-group relationships (Burlingame, Fuhrman, & Johnson, 2001). Far fewer investigated member-member or member-leader relationships, and only a few studies have examined more than one of these relationships at a time. Burlingame and colleagues (2001) argue that group cohesion emerges from the context of member-leader, member-member, and member-group relationships (Burlingame et al., 2001). They reviewed 6 empirically supported principles, including pre-group planning, leader interaction and feedback, and member emotional expression, that relate to both aspects of each of these types of relationships as well as to group cohesion. Research using path analysis conducted by Marmarosh and Holtz (2005) supports the theory that group cohesiveness serves as a primary group factor (Marmarosh et al., 2005), and is related to curative group factors identified by Yalom, such as collective self-esteem and hope for the self. Group cohesion ratings have also been shown to correlate with measures of therapeutic change (Yalom, 1995) and well-being (Bliese & Halverson, 1996). It remains unclear whether the affiliative statements defined by this study and coded for each group session represent an aspect of group cohesion or another group process dynamic. Nevertheless, the present study is the first to create a taxonomy of affiliative statements and demonstrate differences in the frequency of such statements made by

members of single-gender compared with mixed-gender treatment groups. In addition, the greater frequency of empathic and supportive affiliative statements in WRG than in mixed-gender GDC may suggest an in-session group process that could be a mechanism of action for the enhanced efficacy of WRG as reported in the main outcome study (Greenfield, Trucco, et al., 2007).

Another striking finding is the significant difference in the directionality of affiliative statements made in the mixed-gender group therapy. Affiliative statements such as supportive, agreement, and completing-a-thought type statements were more frequently said by women group members to male group members than any other potential combination. It is not clear whether being the person who verbalizes an affiliative statement, or being the person who receives the statement is more helpful in engaging members and creating an individual experience of support. However, assuming that both verbalizing and receiving affiliative statements can be helpful to group members, our results would indicate that in a mixed-gender group format, women are more likely to verbalize these statements and less likely to receive them. In the all-women group format, women are equally likely to both verbalize and receive affiliative statements. Again, this may be another in-session process difference that contributes to women's perceptions that a single-gender format can be experienced as more supportive (Kauffman et al., 1995).

It remains unclear whether the content and rate of member-to-member affiliative statements is an essential part of group cohesiveness and whether either, both, or neither of these is an essential mechanism of action for substance abuse group therapy outcomes. However, substance abuse is disruptive to interpersonal and social relationships and some research has demonstrated that individuals with SUDs on average have greater difficulties establishing and maintaining attachments with other people (Flores, 2001). Therefore, researchers and clinicians alike have argued that group therapy plays an important role in recovery, and is often cited as the treatment of choice, as it allows these individuals to develop skills surrounding developing and maintaining relationships. Given such difficulties, combined with the healing benefits of supportive and empathetic peers, group cohesion becomes a crucial foundation for effective treatment (Brook, Galanter, & Kleber, 2008). Attendees of Alcoholics Anonymous (AA) rated group cohesiveness as one of the most therapeutically significant factors of the program, along with acquisition of knowledge surrounding alcoholism and the sense of similarity with other group members working on the same issues (Cierpiatkowska, 1994). However, the specific gender composition of the treatment group is likely to influence the nature of the relationships that form between group members, which has implications for both group cohesiveness and outcome. Our study provides evidence that there are differences in affiliative statement types, frequency, and member-to-member likelihood of verbalizing and receiving these statements of support in single-gender compared with mixed-gender treatment groups.

This study has a number of limitations including the small sample size which may have limited our ability to detect statistically significant differences between groups, as well as the homogeneity of subjects with regard to demographic characteristics and primary substance of abuse. We conducted the groups with female therapists and it is possible that group process would differ with male therapists; however, this was not a focus of the present study. There were also a limited number of tapes available for coding with more tapes available for the WRG than the GDC condition. Coding methods and ratings might have been subject to selection biases due to the limited number of statements within and across taped sessions that were available for coding. Effect sizes for differences between group therapies for each category of affiliative statements were medium in size and it was only by combining the three that exhibited collinearity that we found a large effect size for difference.

CONCLUSIONS

In spite of the limitations, this study provides a taxonomy and method of coding of observable "in-session" group process by examining the frequency of a range of statements made by group members that represent supportive, cohesive, or empathic connections among members of the treatment group. The results may demonstrate an in-session group process that might constitute a mechanism of action for the efficacy of the all women's WRG. In future research, we would hope to replicate these findings in a larger, two-site Stage II trial of the WRG compared with GDC. In addition, similar research studying elements of in-session group process in single-gender men's recovery groups would be informative in defining potential mechanisms of action for these groups as well. While the results of the current study are promising, further research in the area of group process is needed. As the demand for effective community-based group treatments for substance use disorders increases, so does the need to understand the mechanism of action that make treatments effective as well as how this effectiveness might differ by gender. Developing substance abuse group treatments that allow the participants to receive the greatest support from other group members is likely to enhance outcomes. In addition, while we strive to improve the overall quality of group therapy by utilizing treatment manuals, it may also be important to train group therapists to foster an atmosphere in which there is an exchange of affiliative statements and behaviors among group members. Future research investigating the association between in-session group process outcomes with post-treatment clinical outcomes would be useful in expanding the evidence base of mechanisms of action of substance abuse group therapy.

Acknowledgments

Support for this study was provided by the National Institute on Drug Abuse R01 DA015434 and K24 DA019855. The authors would like to gratefully acknowledge Julia S. Kaufman, BA for her help with manuscript preparation.

References

- Bakali JV, Wilberg T, Hagtvet KA, Lorentzen S. Sources accounting for alliance and cohesion at three stages in group psychotherapy: Variance component analyses. *Group Dynamics: Theory, Research, and Practice*. 2010; 14(4):368–383.
- Bliese PD, Halverson RR. Individual and nomothetic models of job stress: An examination of work hours, cohesion, and well-being. *Journal of Applied Social Psychology*. 1996; 26(13):1171–1189.
- Brook, DW.; Galanter, M.; Kleber, HD. Group therapy. In: Galanter, M.; Kleber, HD., editors. *The American Psychiatric Publishing textbook of substance abuse treatment*. 4th ed.. Arlington, VA US: American Psychiatric Publishing, Inc; 2008. p. 413-427.
- Brown A, Mistry T. Group work with 'mixed membership' groups: Issues of race and gender. *Social Work with Groups: A Journal of Community and Clinical Practice*. 1994; 17(3):5–21.
- Budman SH, Soldz S, Demby A, Feldstein M. Cohesion, alliance and outcome in group psychotherapy. *Psychiatry: Interpersonal and Biological Processes*. 1989; 52(3):339–350.
- Burlingame GM, Fuhriman A, Johnson JE. Cohesion in group psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*. 2001; 38(4):373–379.
- Burlingame, GM.; MacKenzie, D.; Strauss, B. Small Group Treatment: Evidence for Effectiveness and Mechanisms of Change. In: Lambert, M., editor. *Handbook of psychotherapy and behavioral change*. 5th ed.. New York: Wiley & Sons; 2004. p. 647-696.
- Burlingame GM, McClendon DT, Alonso J. Cohesion in group therapy. *Psychotherapy*. 2011; 48(1): 34–42. [PubMed: 21401272]
- Cierpiałkowska L. Therapeutic factors in AA and AI-Anon groups. *Polish Psychological Bulletin*. 1994; 25(1):59–73.
- Clarke G, Rowan A. Looking again at the team dimension in systemic psychotherapy: Is attending to group process a critical context for practice? *Journal of Family Therapy*. 2009; 31(1):85–107.

- Clarke GN. Improving the transition from basic efficacy research to effectiveness studies: Methodological issues and procedures. *Journal of Consulting and Clinical Psychology*. 1995; 63(5):718–725. [PubMed: 7593864]
- Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed.. New Jersey: Lawrence Erlbaum Associates; 1988.
- Crits-Christoph P, Siqueland L, Blaine J, Frank A, Luborsky L, Onken LS, Beck AT. Psychosocial treatments for cocaine dependence: National Institute on Drug Abuse Collaborative Cocaine Treatment Study. *Archives of General Psychiatry*. 1999; 56(6):493–502. [PubMed: 10359461]
- Crowe TP, Grenyer BFS. Is therapist alliance or whole group cohesion more influential in group psychotherapy outcomes? *Clinical Psychology & Psychotherapy*. 2008; 15(4):239–246. [PubMed: 19115444]
- Cummings A, Greenfield SF, Gallop R. Self-efficacy and substance use outcomes for women in single gender versus mixed-gender group treatment. *Journal of Groups in Addiction and Recovery*. 2010; 5:4–16. [PubMed: 21753920]
- Drescher S, Burlingame G, Fuhrman A. Cohesion: An odyssey in empirical understanding. *Small Group Behavior*. 1985; 16(1):3–30.
- Festinger, L.; Schachter, S.; Back, K. *Social pressures in informal groups; a study of human factors in housing*. Oxford, England: Harper; 1950.
- Flores PJ. Addiction as an attachment disorder: Implications for group therapy. *International Journal of Group Psychotherapy*. 2001; 51(1):63–81. [PubMed: 11191596]
- Flowers JV, Booraem CD, Hartman KA. Client improvement on higher and lower intensity problems as a function of group cohesiveness. *Psychotherapy: Theory, Research & Practice*. 1981; 18(2): 246–251.
- Gibbins K, Walker I. Multiple interpretations of the Rokeach Value Survey. *The Journal of Social Psychology*. 1993; 133(6):797–805.
- Gillaspy JA Jr, Wright AR, Campbell C, Stokes S, Adinoff B. Group alliance and cohesion as predictors of drug and alcohol abuse treatment outcomes. *Psychotherapy Research*. 2002; 12(2): 213–229.
- Greenfield SF, Brooks AJ, Gordon SM, Green CA, Kropp F, McHugh RK, Miele GM. Substance abuse treatment entry, retention, and outcome in women: a review of the literature. *Drug and Alcohol Dependence*. 2007; 86(1):1–21. [PubMed: 16759822]
- Greenfield SF, Grella CE. What is 'women-focused' treatment for substance use disorders? *Psychiatric Services*. 2009; 60(7):880–882. [PubMed: 19564216]
- Greenfield SF, Potter JS, Lincoln MF, Popuch RE, Kuper LE, Gallop RJ. High psychiatric symptom severity is a moderator of substance abuse treatment outcomes among women in single vs. mixed gender group treatment. *American Journal of Drug and Alcohol Abuse*. 2008; 34(5):594–602. [PubMed: 18821452]
- Greenfield SF, Trucco EM, McHugh RK, Lincoln MF, Gallop R. The Women's Recovery Group Study: A Stage I trial of women-focused group therapy for substance use disorders versus mixed-gender group drug counseling. *Drug and Alcohol Dependence*. 2007; 90:39–47. [PubMed: 17446014]
- Grella CE, Polinsky ML, Hser YI, Perry SM. Characteristics of women-only and mixed-gender drug abuse treatment programs. *Journal of Substance Abuse Treatment*. 1999; 17(1–2):37–44. [PubMed: 10435251]
- Hair, JF.; Anderson, RE.; Tatham, RL.; Black, WC. *Multivariate Data Analysis*. 4th ed.. Englewood Cliffs, NJ: Prentice Hall; 1995. p. 391
- Hien DA, Cohen LR, Miele GM, Litt LC, Capstick C. Promising treatments for women with comorbid PTSD and substance use disorders. *The American Journal of Psychiatry*. 2004; 161(8):1426–1432. [PubMed: 15285969]
- Hien DA, Wells EA, Jiang H, Suarez-Morales L, Campbell ANC, Cohen LR, Nunes EV. Multisite randomized trial of behavioral interventions for women with co-occurring PTSD and substance use disorders. *Journal of Consulting and Clinical Psychology*. 2009; 77(4):607–619. [PubMed: 19634955]

- Hodgins DC, El-Guebaly N, Addington J. Treatment of substance abusers: Single or mixed gender programs? *Addiction*. 1997; 92(7):805–812. [PubMed: 9293039]
- Hornsey MJ, Dwyer L, Oei TPS. Beyond cohesiveness: Reconceptualizing the link between group processes and outcomes in group psychotherapy. *Small Group Research*. 2007; 38(5):567–592.
- Johnson JE, Burlingame GM, Olsen JA, Davies DR, Gleave RL. Group Climate, Cohesion, Alliance, and Empathy in Group Psychotherapy: Multilevel Structural Equation Models. *Journal of Counseling Psychology*. 2005; 52(3):310–321.
- Joyce AS, Piper WE, Ogrodniczuk JS. Therapeutic alliance and cohesion variables as predictors of outcome in short-term group psychotherapy. *International Journal of Group Psychotherapy*. 2007; 57(3):269–296. [PubMed: 17661544]
- Kauffman E, Dore MM, Nelson-Zlupko L. The role of women's therapy groups in the treatment of chemical dependence. *American Journal of Orthopsychiatry*. 1995; 65(3):355–363. [PubMed: 7485421]
- Ki FY, Chow SC. Statistical justification for the use of composite scores in quality of life assessment. *Drug Information Journal*. 1995; 29:715–727.
- Kominars, K.; Dornheim, L.; DeLucia-Waack, JL.; Gerrity, DA.; Kalodner, CR.; Riva, MT. Group Approaches in Substance Abuse Treatment.. In: DeLucia-Waack, J.; Gerrity, D.; Kalodner, C.; Riva, M., editors. *Handbook of group counseling and psychotherapy*. Thousand Oaks, CA: Sage Publications Ltd; 2004. p. 563-575.
- Kraemer HC, Bloch DA. Kappa coefficients in epidemiology: an appraisal of a reappraisal. *Journal of Clinical Epidemiology*. 1988; 41(10):959–968. [PubMed: 3193139]
- Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*. 1977; 33(2):363–374. [PubMed: 884196]
- Linehan MM, Schmidt H III, Dimeff LA, Craft JC, Kanter J, Comtois KA. Dialectical behavior therapy for patients with borderline personality disorder and drug-dependence. *The American Journal on Addictions*. 1999; 8(4):279–292. [PubMed: 10598211]
- Luthar SS, Suchman NE. Relational Psychotherapy Mothers' Group: A developmentally informed intervention for at-risk mothers. *Developmental Psychopathology*. 2000; 12(2):235–253.
- Marmarosh C, Holtz A, Schottenbauer M. Group Cohesiveness, Group-Derived Collective Self-Esteem, Group-Derived Hope, and the Well-Being of Group Therapy Members. *Group Dynamics: Theory, Research, and Practice*. 2005; 9(1):32–44.
- Marziali E, Munroe-Blum H, McCleary L. The contribution of group cohesion and group alliance to the outcome of group psychotherapy. *International Journal of Group Psychotherapy*. 1997; 47(4):475–497. [PubMed: 9314699]
- Moss E. The Gender-Ambivalent Group and the Gender-Blinded Analyst. *Group Analysis*. 2004; 37(2):273–286.
- Newberry AM, Alexander JF, Turner CW. Gender as a process variable in family therapy. *Journal of Family Psychology*. 1991; 5(2):158–175.
- Noldus LPJJ, Trienes RJH, Hendriksen AHM, Jansen H, Jansen RG. The Observer Video-Pro: New software for the collection, management, and presentation of time-structured data from videotapes and digital media files. *Behavior Research Methods, Instruments & Computers*. 2000; 32(1):197–206.
- Oei TPS, Browne A. Components of Group Processes: Have They Contributed to the Outcome of Mood and Anxiety Disorder Patients in a Group Cognitive-Behaviour Therapy Program? *American Journal of Psychotherapy*. 2006; 60(1):53–70. [PubMed: 16770916]
- Orwin, RG.; Francisco, L.; Bernichon, T. Effectiveness of women's substance abuse treatment programs: A meta-analysis. Center for Substance Abuse Treatment. Arlington, Virginia: SAMHSA; 2001.
- Oser C, Knudsen H, Staton-Tindall M, Leukefeld C. The adoption of wraparound services among substance abuse treatment organizations serving criminal offenders: the role of a women-specific program. *Drug and Alcohol Dependence*. 2009; 103(Suppl 1):S82–S90. [PubMed: 19181457]
- Pincus D, Guastello SJ. Nonlinear Dynamics and Interpersonal Correlates of Verbal Turn-Taking Patterns in a Group Therapy Session. *Small Group Research*. 2005; 36(6):635–677.

- Piper WE. Pregroup interactions and bonding in small groups. *Small Group behavior*. 1984; 15(1):51–62.
- Prendergast ML, Messina NP, Hall EA, Warda US. The relative effectiveness of women-only and mixed-gender treatment for substance-abusing women. *Journal of Substance Abuse Treatment*. 2011; 40(4):336–348. [PubMed: 21315540]
- Prinz, RJ.; Miller, GE.; Peters, RD.; McMahon, RJ. Parental engagement in interventions for children at risk for conduct disorder. In: Peters, RD.; McMahon, RJ., editors. *Preventing childhood disorders, substance abuse, and delinquency*. Thousand Oaks, CA: Sage Publications, Inc; 1996. p. 161-183.
- Randall CL, Roberts JS, Del Boca FK, Carroll KM, Connors GJ, Mattson ME. Telescoping of landmark events associated with drinking: a gender comparison. *J Stud Alcohol*. 1999; 60(2):252–260. [PubMed: 10091964]
- Raudenbush, S.; Bryk, A.; Cheong, Y.; Congdon, R. HLM5: Hierarchical linear and nonlinear modeling. Lincoln, IL: Scientific Software International; 2000.
- Robbins MS, Alexander JF, Turner CW. Disrupting defensive family interactions in family therapy with delinquent youth. *Journal of Family Psychology*. 2000; 14(4):688–701. [PubMed: 11132489]
- Rosenbaum, PR. *Observational studies*. New York: Springer; 2002.
- Rummel, RJ. *Applied Factor Analysis*. Evanston, IL: Northwestern University Press; 1970. p. 298
- Sandler IN, West SG, Baca L, Pillow DR. Linking empirically based theory and evaluation: The Family Bereavement Program. *American Journal of Community Psychology*. 1992; 20(4):491–521. [PubMed: 1481786]
- Schmalisch CS, Bratiotis C, Muroff J. Processes in group cognitive and behavioral treatment for hoarding. *Cognitive and Behavioral Practice*. 2010; 17(4):414–425.
- Shapiro BZ. The social work group as social microcosm: 'Frames of references' revisited. *Social Work with Groups: A Journal of Community and Clinical Practice*. 1990; 13(2):5–21.
- Taube-Schiff M, Suvak MK, Antony MM, Bieling PJ, McCabe RE. Group cohesion in cognitive-behavioral group therapy for social phobia. *Behaviour Research and Therapy*. 2007; 45(4):687–698. [PubMed: 16928359]
- Toren Z, Shechtman Z. Association of personal, process, and outcome variables in group counseling: Testing an exploratory model. *Group Dynamics: Theory, Research, and Practice*. 2010; 14(4):292–303.
- Tschuschke V, Dies RR. Intensive analysis of therapeutic factors and outcome in long-term inpatient groups. *International Journal of Group Psychotherapy*. 1994; 44(2):185–208. [PubMed: 8005718]
- Volpicelli JR, Markman I, Monterosso J, Filing J, O'Brien CP. Psychosocially enhanced treatment for cocaine-dependent mothers: Evidence of efficacy. *Journal of Substance Abuse Treatment*. 2000; 18(1):41–49. [PubMed: 10636605]
- Wolchik SA, West SG, Westover S, Sandler IN. The children of divorce parenting intervention: Outcome evaluation of an empirically based program. *American Journal of Community Psychology*. 1993; 21(3):293–331. [PubMed: 8311029]
- Yalom, ID. *The Theory and Practice of Group Psychotherapy*. 3rd ed.. New York: Basic Books; 1985.
- Yalom, ID. *The theory and practice of group psychotherapy*. 4th ed.. New York, NY: US: Basic Books; 1995.

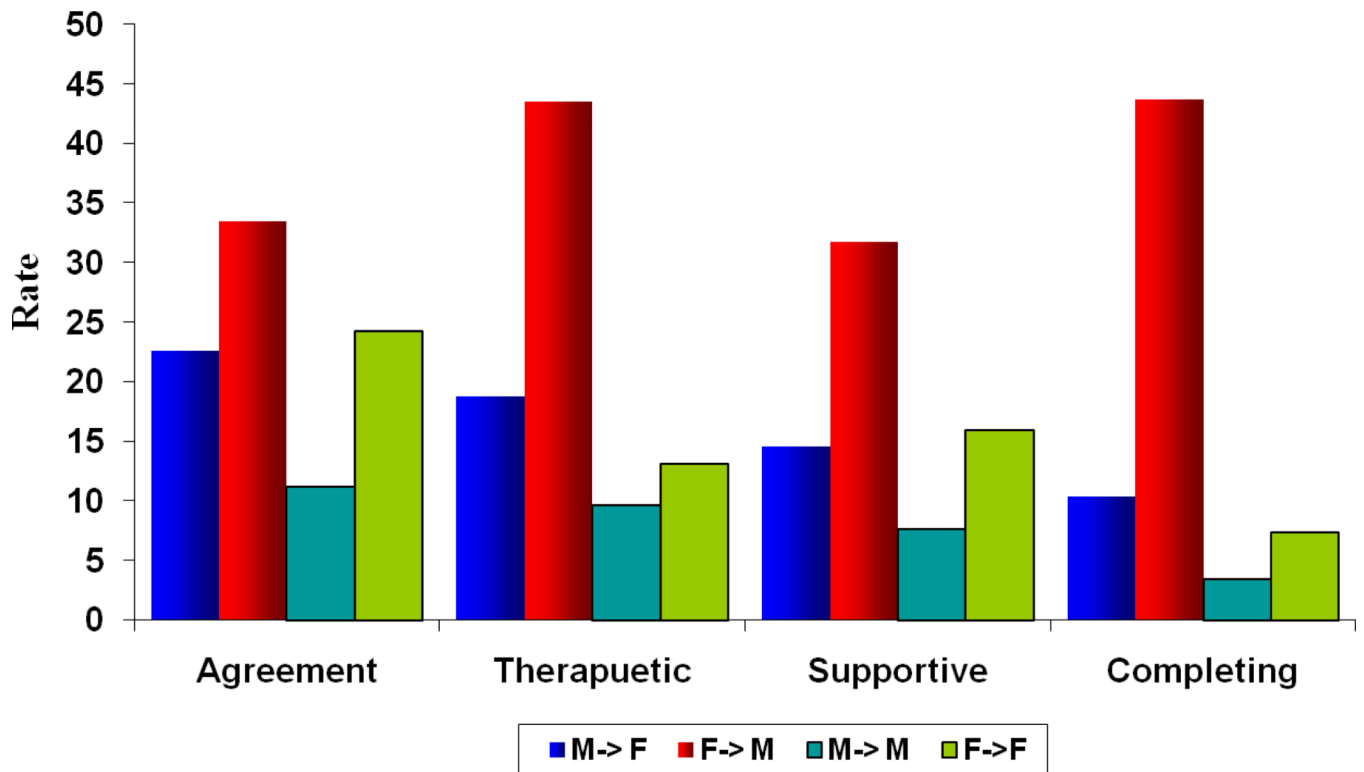


Figure 1.
Directionality rates for the Individual Scales among GDC group participants: Average across sessions

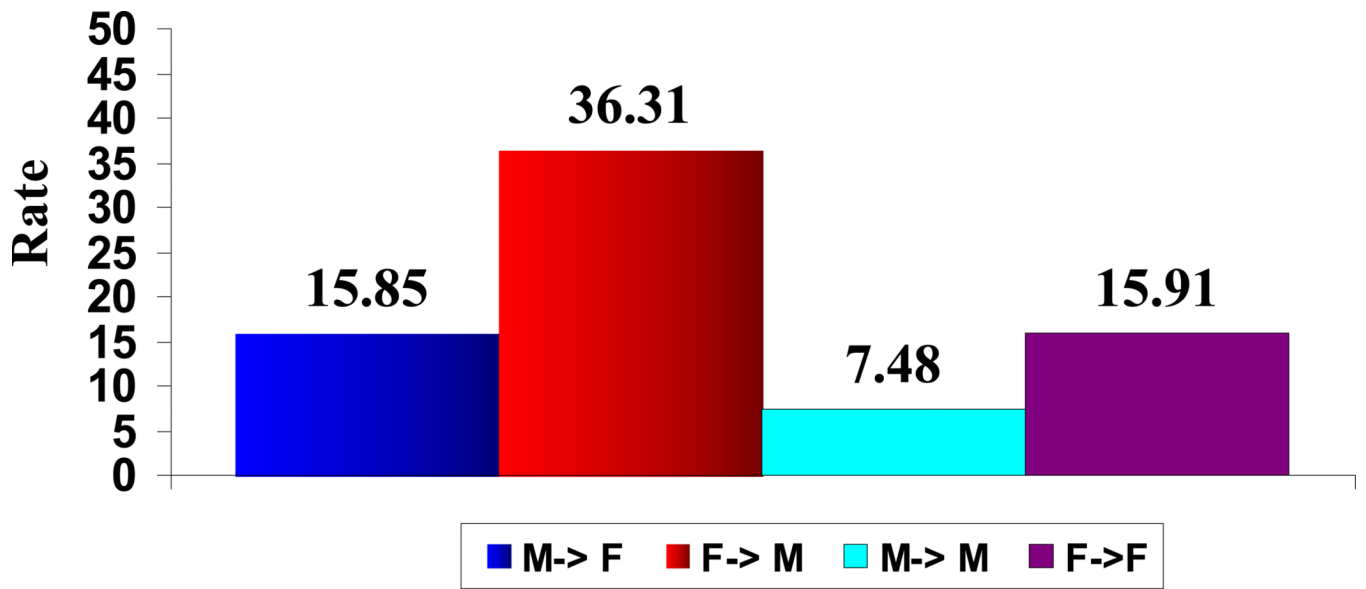


Figure 2.
Directional Rates for the Composite Affiliative Scale among GDC group participants

Table 1

Effect Sizes for Rate of Affiliative Statements by Category Comparing WRG to GDC [Cohen's D (p value)]

*Composite Affiliative Scale: $d=0.882$ ($F(1,4)=1.61$, $p=0.27$) *(Agreement, Completing Thoughts, Supportive Statements Combined)	
Agreement: $d=1.078$ ($F(1,4)=5.13$, $p=0.09$)	Therapeutic Response: $d=-0.095$ ($F(1,4)=0.04$, $p=0.94$)
Completing Thoughts: $d=0.451$ ($F(1,4)=0.31$, $p=0.61$)	Positive Statement: $d=-0.166$ ($F(1,4)=0.29$, $p=0.62$)
Supportive: $d=0.487$ ($F(1,4)=1.37$, $p=0.31$)	