2016

Modeling cohesion change in group psychotherapy: the influence of group leader behaviors and client characteristics

Jeritt Ross Tucker
Iowa State University

Modeling cohesion change in group psychotherapy: The influence of group leader behaviors and client characteristics

by

Jeritt Ross Tucker

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Psychology (Counseling Psychology)

Program of Study Committee:

Nathaniel Wade, Major Professor
Lisa Larson
Meifen Wei
Karen Scheel
Craig Anderson

Iowa State University
Ames, Iowa
2016

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The undertaking of a dissertation is no small task; the academic community may be only half-joking when outlining the clinical features of "post-dissertation depression." Looking back on my own experience, there are several people whose technical and emotional support likely prevented the onset of such a condition. I largely credit my major advisor, Dr. Nathaniel Wade, whose unquestioning certainty of my ability gave me the fortitude to weather many setbacks in the writing process. I would also like to thank Dr. Todd Abraham, whose knowledgeability and patience in communicating such knowledge seemed to transcend the limits of human compassion. Unfortunately for them both, the profession of psychology may never command the kind of salary needed to repay such an enormous debt.

I would also like to thank the members of my doctoral committee—Dr. Lisa Larson, Dr. Meifen Wei, Dr. Karen Scheel, and Dr. Craig Anderson—each of whom I admire greatly. Their rigorous yet kind approach while providing feedback during the early phases of writing greatly improved upon the final quality of this manuscript.

Lastly, I would like to express my gratitude to my friends and family for their unerring support. Specifically, I would like to thank my partner Stacy Ko. Many late nights and drawn out weekends were spent cognitively-taxed during the writing of this manuscript, and she stood strongly by my side without hesitation. Her reassurance that no, I am not a thinly-veiled imposter, and that yes, that reasoning does make sense, may have been the largest contributor to my psychological resilience. Perhaps my own need for social and emotional belongingness during the completion of this dissertation is the only evidence needed for the importance of cohesion in one's well-being.

ABSTRACT

Cohesion, the sense of belonging individuals feel toward groups they are a part of, is a well-documented predictor of group psychotherapy outcomes. Meta-analyses reveal a reliable association between cohesion and reductions in psychological distress (r = .25; Burlingame, McClendon, & Alonso, 2011a) as well as between cohesion and task performance (r =.17; Gully, Devine, & Whitney, 2012). Despite this, few studies have attempted to carefully examine predictors of cohesion during the life of a psychotherapy group. Given contradictory findings on the trajectory of cohesion across time (e.g. Kivlighan & Lilly, 1997; Taube-Schiff et al., 2007; Tschuschke & Dies, 1994), as well recent evidence that differences between therapists predict the growth of cohesion (e.g. Bakali, Wilberg, Hagtvet, & Lorentzen, 2010), the present investigation sought to model changes in cohesion by analyzing early leader interventions while accounting for client- and group-level characteristics. For the present investigation, 128 volunteer clients and 14 group therapists participated in 23 separate time-limited psychotherapy groups. Client characteristics (attachment style, self-esteem, and psychological distress), therapist characteristics (counseling self-efficacy), first-session therapist behaviors (structuring, verbal interaction, and emotional facilitation), and group characteristics (number of members, member attendance) were used to predict changes in cohesion across time. For the methodology, a Latent Growth Curve (LGC) Analysis under a Hierarchical Linear Modeling (HLM) framework was used; with client ratings serving as indicators of the outcome variable (cohesion), level 1 representing the effects of time, level 2 representing client characteristics, and level 3 representing group characteristics (including leader behaviors and self-efficacy). Results indicated that a piecewise linear-quadratic model best fit the data, with group membership explaining between 3-20% of the variability in cohesion change. Significant individual level

predictors included gender, race, and anxious and avoidant attachment. Significant group-level predictors included structuring behaviors, which were moderated by the presence of behaviors thought to facilitate an emotional climate. Limitations and possible areas of future research are discussed and implications for the theory and practice of short-term group psychotherapy are provided.

CHAPTER 1. OVERVIEW

Social belonging is a basic human drive, and the maintenance of lasting, positive, and supportive relationships is central to psychological health (Baumeister & Leary, 1995). In parallel fashion, those who suffer from chronic loneliness often experience depression, anxiety (Jones & Carver, 1991), reduced immune functioning, more frequent disease, and lower subjective well-being (Kiecolt-Glaser, et al., 1984). Clearly, social groups are fundamental to both mental and physical health (see Forsyth, 2012 for a review).

To contribute to social belonging, social groups must go beyond simply bringing people together. There is a clear distinction in the social psychology research between what are referred to as social and emotional loneliness (e.g. Weiss, 1973; Russell, Cutrona, Rose, & Yurko, 1984; Van Baarsen, Snijders, Smit, & Van Dujin, 2001). Whereas social loneliness implies a lack of connection to others in general, emotional loneliness implies a lack of intimate relationships (Weiss, 1973). Research suggests that transitory groups, or those that foster little connection between members, do little to prevent deeper forms of emotional loneliness. Reductions in emotional isolation requires more "involving" groups which provide encouragement, information, and support (Forsyth, 2012). As such, psychotherapy groups, especially those which foster belongingness, can provide an important antidote to the negative impacts of social isolation and contribute greatly to psychological health.

Since Moreno's (1932) work on creating healthy subgroups in an in-patient population, *cohesion* has been thought to be a central component of a groups' ability to increase adjustment and decrease interpersonal conflict. Cohesion, although a term not without controversy, captures the sense of belonging individuals feel toward groups they are a part of (Frank, 1957). Cohesion is the force of "we-ness" that binds group members in the pursuit of a common goal (Yalom,

2005). Cohesion is consistently rated as one of the most powerful therapeutic factors in group counseling by clients, and is one of the most well-researched predictors of therapeutic outcomes (Yalom, 2005). Cohesion has been positively associated with the completion of therapeutic tasks, improvements in psychological well-being, and improvements in interpersonal functioning (Joyce, Piper, & Ogrodniczuk, 2007). Despite this research support, however, there is a great deal of uncertainty surrounding the term itself.

Present State of the Literature

In their critique of the cohesion construct, Hornsey, Dweyer, and Oei (2007) raised several important questions regarding the use of the term. They stated that while cohesion has been consistently found to relate to outcomes, the term itself may be too diffuse to identify which specific processes are occurring to promote these outcomes. They stated that possible mediators of the cohesion-outcome relationship have not been empirically examined and that it is unclear as to whether cohesion should be treated as an antecedent, consequent, or covariate with other important group processes (such as self-disclosure or interpersonal feedback; Hornsey, Dweyer, & Oei, 2007). Whereas some of these criticisms have been answered in the cohesion literature (in particular is a common definition of cohesion based on multiple, empirically-derived factors; see Burlingame, McClendon, & Alonso, 2011a), many remain to be addressed.

In line with these shortcomings in the group psychotherapy literature on cohesion, the *2011 Task Force on Evidence-Based Therapy Relationships* identified several primary research aims for the future study of group processes (Norcross & Wampold, 2011). Primary of these were 1) better understand how the observational perspective (e.g. client, counselor, or outside observer) may relate to the conceptualization of relationship variables and their relationship to outcomes, 2) explicate therapist behaviors and qualities that promote facilitative therapy

relationships, and 3) go beyond simple correlational models to parse out the unique contributions of clients and counselors to therapeutic outcomes (Norcross & Wampold, 2011).

The present investigation addresses these three primary research aims as they relate to group psychotherapy. This is done through 1) incorporating client- and leader-rated measures of cohesion in a latent variable model, 2) modelling cohesion change across time, and 3) incorporating client, counselor, and group variables into a nested, multi-level design. In doing so, the present investigation more clearly explicates the antecedents and covariates of cohesion to provide clinical recommendations to group leaders interested in facilitating cohesive and effective psychotherapy groups.

CHAPTER 2. LITERATURE REVIEW

Operationalizing Cohesion

Cohesion is a complex construct with an iterant history; there are perhaps as many conceptualizations of cohesion as there are measures. Various researchers have noted as many as ten separate measures of cohesion regularly used in the analysis of group psychotherapy outcomes (Hornsey, Dwyer, & Oei, 2007). These measures have focused on various constructs, including group attendance, belongingness, group significance, avoidance, the presence or absence of conflict, affiliation, empathy, interpersonal liking, group attractiveness, self-disclosure, and group acceptance among clients (see Burlingame, McClendon, & Alonso, 2011b). Understandably, this has led some researchers to question the study of cohesion altogether; with even early scholars suggesting that there is "[…] little cohesion in the cohesion literature" (Bednar & Kaul, 1978; p. 800).

Still, research over the past decade has added considerable conceptual clarity. Two somewhat orthogonal views of cohesion have emerged; those that focus on the *quality* of group relationships (including such constructs as working alliance, cohesion, and climate), and those that focus on the *structure* or arrangement of these relationships (either horizontal member-to-member interactions or vertical member-to-leader interactions; Burlingame, McClendon, & Alonso, 2011a). In a recent factor analysis of both quality and structural components of cohesion, Johnson, Burlingame, Olsen, Davies, and Gleave (2005) found that the most appropriate model of therapeutic relationships is comprised of three-factors that incorporate components of both quality and structure. These three factors included 1) Positive Bonding Relationships (e.g. member-group cohesion, member-group engagement, and member-leader positive relational aspects), 2) Positive Working Relationships (e.g. member-member and

member-leader agreement on tasks) and 3) Negative Relationships (e.g. member-group conflict and lack of empathy from both leaders and other members). As a part of this research line, these investigators noted that members of psychotherapy groups tend to base their ratings of cohesion primarily on the overall quality of bonding and working relationships, and noted that both positive working and positive bonding may fall within the broader category of *Engagement* (Johnson et al., 2006, p. 137).

Accordingly, the contemporary view suggests that cohesion acts primarily as a group "aggregate" across each of the various member-member, member-leader, and member-group relationships and operates at the "group-as-a-whole level" (Burlingame, Fuhriman, & Johnson, 2002, p. 373). The term "group-as-a-whole" borrows from T. Burrow's early (1928) research into phenomena arising from the summation of interpersonal and intrapersonal interactions. Here, the "collective being" of the group has its own norms, conscious and unconscious goals, and unique patterns of communication and power. It is from this level that feelings of "groupness" emerge (Ettin, Cohen, & Fidler, 1997).

As such, it is appropriate for group researchers, in operationalizing cohesion, to identify which structural and content levels they are assessing. The present study chooses to take one perspective on cohesion; primarily examining the *member-to-member* structural level and the *Engagement* quality level. This is consistent with two of the most long-standing measures of cohesion, the Engagement subscale of the Group Climate Questionnaire (GCQ; MacKenzie, 1983) and the Cohesion subscale of the Group Experiences Scale (GES; Wilson et al., 2008). Questions on the GCQ-E and GES assess member-to-member relationships and have questions such as "members challenged and confronted each other in their efforts to sort things out"

(*positive working*; MacKenzie, 1983) as well as "group members show that they care for one another" (*positive bonding*; Wilson et al., 2008).

Similar to an increasingly clear definition of cohesion, researchers have added some clarity as to whether cohesion should be measured at the individual or group levels. Johnson and colleagues (2006) examined the factor structure of two commonly used measures of group climate (the Group Climate Questionnaire – Short form; GCQ-S; MacKenzie, 1983; and the Curative Climate Instrument; CCI; Fuhriman, Drescher, Hanson, Henrie, & Rybicki, 1986) at both the individual and group levels. Multilevel confirmatory factor analysis revealed that the GCQ subscales failed to provide an adequate fit to the data at both levels, but that the fit was substantially poorer at the group level. Still, much of the difficulty in fit was hypothesized to result from the Conflict subscale of the GCQ; which measures the presence of negative interactions between members and has reportedly low internal reliability. The researchers thus concluded that the appropriateness of analyzing cohesion at the group level is still "open for debate" (Johnson et al., 2006; p. 143). While the jury is thus still out regarding which level is most empirically sound, theory and research suggesting that cohesion emerges across member-to-member, member-to-leader, and member-to-group levels (e.g. Burlingame, McClendon, & Alonso, 2011b; Burlingame, Fuhriman, & Johnson, 2002) may still support modelling cohesion at the group level.

Other researchers have commented on the importance of measuring cohesion from multiple vantage points as a way of unifying the concept and increasing its reliability. MacKenzie and Tschuschke (1993) emphasized using client-rated cohesion; stating that individuals who report higher relatedness to the group also report greater symptomatic improvement. Budman and colleagues (1989) instead emphasized observer-rated cohesion.

They found that group-level cohesion, as rated by an independent observer of both member-member and member-leader interactions, was related to client improvement. They also noted that cohesion within the first 30 minutes of a session had the strongest relationship to outcomes (Budman et al., 1989). Other researchers have noted the importance of therapist ratings of group cohesion. In their 2004 study, Lorentzen, Sexton, and Høglend found that therapist ratings of early alliance correlated significantly with symptom reductions in clients.

Not only is it useful to have multiple vantage points in capturing the construct of cohesion, the amount of agreement *between* these vantage points also appears important to clinical outcomes. McNeil (2005) observed a great deal of variability between client, counselor, and observer ratings of cohesion. Results of this study indicated that clients appeared to have the most accurate perceptions of cohesion, with their ratings being the most highly correlated with working alliance at various time points in group. In another study, a moderate correlation was found between client and counselor views of cohesion (r = .56; Lorentzen, Sexton, & Høglend, 2004), and higher concordance rates between leaders' and members' ratings of cohesion were related to better client improvement. Taken together, the literature in this area points to the importance of multiple perspectives—including members, leaders, and independent observers—in measuring cohesion. Doing so allows for a more nuanced view of cohesion ratings toward improved reliability and an understanding of how disagreements between leaders and members might impact later development of group cohesion.

Cohesion and Consequential Outcomes

Despite increased conceptual clarity, research examining cohesion's relationship to outcomes continues to reveal contradictory findings (e.g. Budman et al., 1989; Alonso, 2012). Although this has led some researchers to be hesitant regarding the use of cohesion as a useful

construct (e.g. Hornsey, Dwyer, & Oei, 2007; Hornsey, Dwyer, Oei, & Dingle, 2009), abandoning cohesion altogether may limit researchers' ability to understand an important group process. Out of the twelve therapeutic factors identified by Yalom (2005), cohesion is virtually unanimous in its ranking as the most important factor to group therapy clients. In psychotherapy groups, cohesion has been linked to increased meaningful self-disclosures among group members and higher frequency and authenticity of feedback (Yalom, 2005). The development of cohesion in groups seems to help members tolerate the stress and negative emotion associated with the "storming" (i.e. conflictual) and working stages of a psychotherapy group. As such, there is strong theoretical evidence for the examination of cohesion in group psychotherapy process and outcomes (Budman et al., 1989; Kivlighan & Lilly, 1997).

There is also a great deal of empirical evidence to support the examination of cohesion. Dinger & Schauenburg (2010) found that higher levels of group cohesion and an increase in cohesion over the course of a group's life are associated with greater symptom improvement. Additionally, groups with markedly higher levels of group cohesion have members who, at termination, experience higher levels of self-esteem and goal attainment (Budman et al., 1989; Tschuschke & Dies, 1994; Alonso et al., 2012). Researchers have also found cohesion to be an important predictor of various group processes such as increases in overall group performance (Evans & Dion, 1991), higher member attendance (Ogrodniczuk, Piper & Joyce, 2006), and higher attraction, bonding, and empathy (Braaten, 1990). Even anticipated levels of cohesion have been linked to positive client improvement (Alonso et al., 2012).

Joyce, Piper, and Ogrodniczuk (2007) provide one clear example of the effect of cohesion on outcome variables. In this study, cohesion was measured from the perspective of group members as well as therapists. Cohesion was highly related to group members'

perceptions of compatibility with fellow members. Therapists perceived that individuals with better group fit were more likely to experience a reduction in symptoms than those individuals with poorer fit. Results revealed that therapists' ratings of clients' compatibility with the group, or their perception of a member's "fit" within the group, was directly associated with improvement in clients' general symptoms (Joyce, Piper, & Ogrodniczuk, 2007).

Perhaps the most compelling evidence for the importance of cohesion in consequential outcomes, however, comes from recent meta-analytic reviews. In non-therapeutic settings, several researchers have observed an important relationship between cohesion and group performance on tasks. Gully, Devine, & Whitney (2012) observed an average effect size of r = .17 [95% CI=.14, .26] between cohesion and objective measures of performance including creative thinking in writing groups, goals scored in ice hockey teams, and combat effectiveness in military groups.

Additionally, Burlingame, McClendon, & Alonso (2011a) found an average effect size of r = .25 [95% CI=.17, .32] between cohesion and reductions in symptom distress and improvements in interpersonal functioning in psychotherapy groups. This was observed across numerous settings and presenting concerns and in groups with leaders of various theoretical orientations. Burlingame and colleagues' work also suggested that when leaders focus specifically on building cohesion, their groups have an enhanced cohesion-outcome relationship (Burlingame, McClendon, & Alonso, 2011a).

Rather than abandon cohesion research altogether, it thus appears wise for group psychotherapy researchers to bring a more nuanced approach to understanding cohesion and its integral role in groups. In particular, researchers may strive to examine how cohesion changes

across the life-span of psychotherapy groups and which client, leader, and group characteristics predict these changes.

Changes in Cohesion Across Time

Although cohesion appears to have an important relationship with group psychotherapy outcomes, it is yet unclear as to how cohesion changes across stages of group development (Kivlighan & Lilly, 1997). There are many reasons to believe that cohesion is not a static variable; the emergence or decrease of anxieties and conflicts can enhance or challenge cohesion across a group's life-span.

Theory and empirical investigation has led to many hypotheses about the development of cohesion across time (see *Appendix A*). Early on, MacKenzie (1983) proposed a group development theory suggesting that cohesion in functioning groups first increases slowly during initial stage of group life, drops during the middle differentiation phase, and finally is restored during the final individuation phase of group. Still, empirical evidence may support a slightly simpler model than the one proposed by MacKenzie. In their study of a twelve-session, manualized cognitive behavioral therapy group for clients with binge-eating disorders, Castongauy, Pincus, Agras, and Hines (1998) found that members with the greatest improvement experienced a negative group climate in the middle of therapy coupled with a supportive (higher cohesion) early and late phase of group. Similarly, Kivlighan and Lilly (1997) used latent growth curve analysis (LGC) to examine the impact of time and group development on cohesion and member outcomes. They examined therapy groups composed of undergraduate and graduate students lasting between 14 and 26 sessions. Their findings suggested a similar pattern to Castonguay and colleagues (1998); namely that cohesion starts high and stays high during the initial stage, drops in the middle stage, and is regained at the end (thus forming a quadratic high-

low-high pattern). Both authors hypothesized that early supportive engagement may set the stage for group members to tolerate the negative emotion involved in making therapeutic changes. This sets the stage for higher cohesion in later stages, but leads to a drop in cohesion in middle stages. Kivlighan and Lilly additionally noted, however, that their group members were relatively "healthy" such that their observed pattern of high-low-high may be an artifact of their particular sample. For example, those with fewer mental health concerns may be able to form close relationships more quickly, whereas those with more pathological concerns may require a slower building of cohesion during initial stages. If this is true, one might expect Mackenzie's (1983) model (a cubic low-high-low-high pattern) to occur in groups with more significant psychopathology.

In one examination of groups composed of individuals with more persistent mental illness, Tschuschke and Dies (1994) examined cohesion in a long-term psychoanalytic treatment group for inpatients. They found that the most successful patients had a high level of cohesion that reached its peak shortly after the beginning of treatment and maintained for the majority of therapy. On the other hand, the least successful patients maintained a lower level of cohesion with a slight tendency to increase at the end of treatment. Still, these were exploratory observations and no modelling or statistical analysis was conducted (Tschuschke & Dies, 1994). Finally, Taube-Schiff et al. (2007) examined the development of group cohesion in a cognitive behavioral treatment group for social phobia. These researchers noted that cohesion ratings significantly improved from the midpoint of group to the end of treatment. Moreover there was little variance in group cohesion change across treatment groups.

Thus, while there is a lack of research on changes in cohesion across time, there appears to be preliminary evidence for Kivlighan and Lilly's assertion that cohesion may take longer to

develop in groups with more severe psychopathology. Still, it may not take Kivlighan & Lilly's hypothesized cubic shape, but instead may take a simpler linear or curvilinear relationship with time. Results in this area foremost point to the need to continue to examine cohesion change in psychotherapy groups.

Client Characteristics Predictive of Cohesion

Along with understanding how cohesion changes it is equally important to determine what predicts such change. There are numerous client, leader, and group variables with potential relationships to cohesion. Those variables addressed in the present investigation are identified in *Appendix B.* At the lowest level, several client variables almost certainly play a role in the development of group cohesion. Given that cohesion is composed of positive relational elements between members (belongingness, engagement, and task alliance), negative relational elements (conflict, avoidance), and positive working elements (task cooperation, goal agreement), client characteristics impacting these variables are likely to influence the development of cohesion across time. In particular, a client's level of psychological distress may impact cohesion as higher levels of distress have been found to encourage attendance of group (Frances, Clarkin, & Perry, 1984; Yalom, 2005).

An empirical examination of the relationship between distress and cohesion appears to partially support this idea. In one study by Ogrodniczuk and colleagues (2004) clients with either moderate or high levels of distress had equivalent attendance, whereas those with low-levels of distress tended to terminate treatment earlier (Ogrodniczuk, Piper, & Joyce, 2004). Ogrodniczuk and colleagues (2004) study was conducted in a sample of those with diagnosed personality disorders, however, which may mean that there was less variation in overall psychological distress in the prediction of attendance. Thus, more research is needed in this area

to examine whether psychological distress is an important client characteristic in the prediction of cohesion development.

Group members' attachment and interpersonal style may also play a role in the development of cohesion. Marmarosh and colleagues (2005) found that clients who dropped out of treatment had greater attachment anxiety than clients who remained in their groups throughout the course of treatment. Other researchers have similarly suggested that attachment anxiety predicts premature termination, especially in early therapy (Huang et al., 2013). Levy, Ellison, Scott, and Bernecker (2011) found anxiously-attached group clients to have poorer outcomes (e.g. depression, anxiety, life satisfaction, psychological symptoms), whereas attachment avoidance was not associated with consequential outcomes. Still, additional research is needed to confirm these findings.

Research in the area of attachment has also demonstrated that attachment style affects members' perceptions of the group's climate. For example, Kivlighan and Angelone (1992) noted that members with avoidant styles are more likely to perceive the group climate as hostile and tense than individuals with anxious attachment styles. These researchers suggested that this was the result of individuals' projections of their own self-perceptions onto the group. Similarly, Chen and Mallinckrodt (2002) found that individuals with insecure attachment styles perceived the group as more hostile and less attractive. Securely attached individuals, on the other hand, appear to be more able to readily attach to group leaders and better cope with positive and negative interactions (Kivlighan, Patton & Foote, 1998). Ambivalent group members have been seen to have difficulty coping with confrontations, and appear to prefer to cope alone (Kivlighan, Patton & Foote, 1998). Shechtman and Rybko (2004) found that a secure attachment style was linked with greater self-disclosure, intimacy, and empathy in the therapy process. Finally, a

study on interpersonal styles indicated that clients described as “too friendly” had improved outcomes when cohesion scores decreased whereas individuals described as “cold” or “distant” demonstrated improved outcomes when cohesion scores increased (Schauenberg, Sammet, Rabung, and Struck, 2001).

Although more work is needed to understand the role of client attachment, the available literature does indicate that attachment plays an important role in group processes. A better understanding of the impact of group member attachment styles on the development of group cohesion over time can assist group therapists in identifying transference reactions that arise throughout the course of the group. A better understanding of the impact of attachment on cohesion can assist in the development of interventions to increase group cohesion and provide corrective experiences that are appropriate for individuals with a wide range of attachment styles and interpersonal problems (Chen & Mallinckrodt, 2002). Differences in levels of insecure versus secure attachment may also begin to explain observed differences in the development of cohesion between inpatient and outpatient psychotherapy groups.

Lastly, member self-esteem has been linked to cohesion in groups. Yalom (2005) noted that the interrelation between group self-esteem and individual self-esteem is foundational for cohesion. While researchers have examined improved self-esteem as an important outcome of cohesive groups (e.g. Marmarosh et al., 2005), other researchers have hypothesized that initial self-esteem begets higher rates of self-disclosure, more empathy toward group members, and more positive responses to feedback (e.g. Shea & Sedlacek, 1997). These factors may mean that group members’ initial self-esteem positively impacts later development of cohesion.

Leader Characteristics and Behaviors Predictive of Cohesion

The potential impact of leadership behaviors in the development of cohesion cannot be understated. Since Lewin, Lippit, and White's classic (1939) work on the influence of different leadership styles on the feelings, thoughts, and behaviors of groups of boys working on hobby projects, it has been evident that leaders can greatly shape the direction and outcomes of groups. Over the course of research on leadership, numerous characteristics of effective leaders in psychotherapy groups have been identified. For example, those leaders who are less controlling, more warm and caring, and who set clear group norms have been found to have more cohesive groups and better outcomes (Antonuccio, Davis, Lewinson, & Breckenridge, 1987).

Although the characteristics of effective leaders have been well-studied, several researchers have commented on the lack of empirical investigation into specific leader *behaviors* (e.g. Morran, Stockton, & Whittingham, 2004; Baldwin et al., 2008; Bakali et al., 2010). For example, Kivlighan and Lilly (1997) specifically recommended that group leaders "look for ways to increase the cohesiveness and engagement of their group" (p. 218), yet are able to point to only a single study outlining how leaders should do so (e.g. Levine & Moreland, 1990). Still, numerous theoretical models for categorizing and describing effective leader behaviors have emerged. An outline of these categories, descriptions of the behaviors, and theoretical support for when interventions are hypothesized to be appropriate is provided in *Appendix B.*

Morran, Stockton, and Whittingham (2004) have helpfully distinguished between two broad types of effective leader behaviors. These include 1) protecting group members and promoting safety and 2) energizing and involving group members. Protecting group members is meant to capture efforts to create a group climate conducive to trust, openness, and closeness. It includes such behaviors as protecting vulnerable members, blocking inappropriate behaviors, and

supporting individual members. The second category, energizing and involving group members, is meant to capture behaviors intended to stimulate forward progress, increase participation, and lead to interpersonal learning. It includes the elements of drawing out hesitant members, modeling appropriate group behaviors, linking members with one another, processing emotional exchanges, providing interpretations of events, disclosing personal information, and providing feedback (Morran, Stockton, & Whittingham, 2004).

Chapman, Baker, Porter, Thayer, and Burlingame (2010) have set out a different categorical approach to leader behaviors, outlining three domains of effective leader interventions in the development of the Group Psychotherapy Intervention Rating Scale (GPIRS). These three broad domains included 1) Group Structuring, 2) Verbal Interaction, and 3) Creating and Maintaining an Emotional Climate. Each item of these domains was derived from previous research on interventions associated with enhancing the therapeutic factors of group psychotherapy. Domain one, Group Structuring, is comprised of interventions meant to establish a framework for members to understand the group's purpose, boundaries, rules, and roles. This facilitates the creation of a safe working environment for engaging in the primary group tasks. It includes the elements of establishing norms, explaining events, identifying critical incidents, and facilitating group member interactions. The second domain, Verbal Interaction, consists of leader interventions meant to model and facilitate appropriate verbal interactions and style. This includes leader modeling, self-disclosure, and feedback. Finally, domain three, Emotional Climate, consists of interventions meant to facilitate clients' emotional expression in a constructive manner. It includes the components of refraining from hostility, eliciting member support, and preventing problem interactions.

An important component of research in the area of cohesion development is identifying the stage(s) in which group leader behaviors are most effective (see *Appendix B*). Leadership behaviors early in the group are hypothesized to be particularly important in the development of cohesion across time. Joining a group is frequently associated with a high level of initial discomfort and anxiety. When early leader behaviors do not effectively deal with this discomfort, low involvement, early drop-out, and reduced cohesion often occurs (Neimeyer & Thomas, 1982; Burlingame et al., 2001).

There is modest empirical support for a relationship between specific leader behaviors and the development of group cohesion. In their validation study of the GPIRS, Chapman and colleagues (2010) found significant correlations between every GPIRS subscale and a proxy for group cohesion (the Group Climate Questionnaire Engagement; GCQ-E; MacKenzie, 1983). Specifically, the Structuring domain correlated with the GCQ-E at r = .42, Verbal Interaction at r = .36, and Emotional Climate at r = .35. Additionally, all of the GPIRS subscales negatively correlated with member-reported conflict (Group Structuring r = -.42, Verbal Interaction r = -.38, Emotional Climate r = -.38, and GPIRS total score r = -.43; Chapman et al., 2010). In further support, other researchers have found a link between ineffective leader behaviors and poor cohesion and outcomes. Fuehrer and Keys (1988) found that too little structuring led to poorer client ratings of member interactions. Smokowski, Rose, Todar, and Reardon (1999) found that observer-rated leader passivity in not protecting or supporting members was linked to clients' premature termination. Together, these findings point to the potential importance of leader behaviors in shaping the development of cohesion across time in psychotherapy groups.

Another potentially important predictor of the development of cohesion is the leader's counseling self-efficacy. Counseling self-efficacy is a concept founded in social cognitive

theory (e.g. Bandura, 1982), and has been defined as "[…] one's beliefs or judgments about her or his capabilities to effectively counsel a client in the near future" (Larson & Daniels, 1998; p. 179). Counseling self-efficacy has been linked to lower anxiety during counseling sessions, greater affective responsiveness to clients, and increased problem-solving during therapy sessions (Al-Darmaki, 2005). Additionally, higher counseling self-efficacy has been related to increased congruence between counselors' and clients' perceptions of session quality, greater alliance building in session, and increased responsiveness to clients (Lent et al., 2006), all of which may be important contributors to the development of cohesion in groups.

Group Characteristics Predictive of Cohesion

Along with client and leader variables, several group-level characteristics have been explored in previous research including the purpose of the group, the group's size, and the average number of sessions attended by members (Burlingame et al., 2011a). In their meta-analytic review, Burlingame and colleagues (2011a) found several group variables that moderated the relationship between cohesion and outcomes. In particular, groups composed of members with a similar diagnosis or common presenting concern showed lower correlations between cohesion and outcomes than less-structured and more interactive groups. Additionally, groups comprised of five to nine members and groups that lasted more than 12 sessions posted higher cohesion-outcome correlations (Burlingame et al., 2011a). These same variables may directly predict the emergence of cohesion across the life-span of a group. Having more members, with similar characteristics, together more consistently, and with a greater focus on intentionally building relationships is likely to augment member and leader perceptions of cohesion.

The Independent Contributions of Client, Therapist, and Group Variables

Few studies have attempted to examine client, therapist, and group contributions to cohesion simultaneously. In a rare case, Bakali et al. (2010) examined patient, therapist, and group variability while accounting for alliance and cohesion over three group stages. They sampled 145 patients in short-term (20 session) and long-term (80 session) psychodynamic therapy groups. As hypothesized, they found that patients explained the largest amount of variability in cohesion (19-25% across three stages). Furthermore, therapist variability explained a small but significant amount of variance in cohesion (3-7% across three stages). They found that therapists explained the greatest variability in cohesion during the second wave of data collection (sessions 10 and 11); suggesting that this is an important stage of differentiation in which therapists may be more influential. Finally, they found that the format of the group (either short or long-term) did not explain variance in cohesion scores. Overall, results of this study suggest that patient variables explain a large degree of variance in alliance and cohesion, and that therapists contribute a small but reliable amount to alliance and cohesion at different stages of group development.

Still, although this work is an important contribution to the literature, these researchers utilized a variance components analysis under the framework of generalizability theory. This format examined each time point separately to determine which variance components contributed most greatly, which did not allow for modelling changes in cohesion across time. Such an analysis does not account for the serial dependency that exists when examining the same individuals across time, and is suspect when comparing time-points to one another or examining longitudinal data. Additionally, utilizing a variance component analysis does not allow for an examination of specific client, leader, and group characteristics that account for the variance

explained, and thus cannot identify underlying mechanisms that may facilitate the development of cohesion.

The Present Study

There exists only a small body of research that examines predictors of cohesion in psychotherapy groups while simultaneously accounting for member, leader, and group contributions. I sought to fill this apparent gap with the present study. Specifically, I sought to examine changes in cohesion across short-term psychotherapy groups by including 1) the client characteristics of self-esteem, psychological distress, and attachment style; 2) the leader characteristic of counseling self-efficacy and the presence of early leadership behaviors thought to promote cohesion; and 3) the group characteristics of size, session attendance, and group format. Finally, I sought to operationalize cohesion in a very methodologically-rigorous way; modeling cohesion as a latent variable composed of member and leader ratings. Consistent with the most frequently used measures of cohesion—Mackenzie's 1983 Group Climate Questionnaire (GCQ) and Wilson and colleagues 2008 Group Experiences Scale (GES)—the present investigation operationalized cohesion as member and leader ratings of the level of member-to-member engagement.

Specific formulations of the hypotheses tested follow below.

Hypotheses

Hypothesis 1: Examining leader- and member-rated cohesion. Primary to this hypothesis was that leader and client-rated measures of cohesion could serve as indicators of a latent cohesion construct using multiple indicator linear growth modeling. This hypothesis involved several steps. First was the investigation of the group-level dependency between member ratings of cohesion to determine if cohesion should be treated as an individual or group-

level variable. If significant member-agreement existed, the concordance between group-level member ratings and leadership ratings of cohesion were to be examined using Cross-Correlation Coefficients (CCF; Lorentzen et al., 2004). Finally, assuming acceptable ICC values, both members and leader ratings of cohesion would be entered as indicators in a multiple-indicator model, allowing the residual variances of both raters to be uncorrelated and estimated across time (see *Methods* section below).

Hypothesis 2: Modeling change in cohesion across time. The second hypothesis involved the prediction that cohesion would reliably change across time (sessions 1-8), that significant variance would exist in this change, and that this change would take on a curvilinear trajectory. The curvilinear trajectory was hypothesized consistent with observations by Taube-Schiff and colleagues (2007) and Tschuschke and Dies (1994). Specifically, cohesion was hypothesized to increase throughout the lifespan of the group, but this increase was suspected to occur more rapidly during the initial sessions (sessions 1-4) while slowing during later sessions (sessions 4-8). In order to test this hypothesis, different measurement models were compared to empirically identify the best-fitting model (e.g. linear, quadratic, or cubic).

Hypothesis 3: Client contributions. First, psychological distress was hypothesized to be linked with the development of cohesion across time, such that early levels of high distress would encourage attendance and increase cohesion. I additionally predicted, however, that continued levels of distress into later stages of group (sessions 4-8) would be negatively related to cohesion as an indication of poor performance on an essential group task (decreasing psychological impairment). I additionally hypothesized that anxious and avoidant attachment styles at time one would predict the development of cohesion. Anxiously-attached individuals were likely to rate cohesion at higher levels in early stages of group, whereas those with

avoidant-attachment styles would rate levels of cohesion as lower during initial stages. I predicted that in later sessions (sessions 4-8), however, both avoidant and anxious attachment, as insecure attachment styles, would show lower levels of cohesion. Finally, I hypothesized that higher client self-esteem at time one would positively predict the development of cohesion across the life-span of the group. Specifically, individual self-esteem would serve as a moderator of the relationship between time and cohesion. That is, for groups with lower cohesion at time 1, those groups with higher individual self-esteem would "recover" more quickly than those with lower self-esteem and show quicker gains in cohesion.

Hypothesis 4: Group-level contributions. Session attendance, group size, and treatment protocol were hypothesized as important group-level variables in the prediction model. Specifically, greater session attendance (as an aggregate, across-individual variable) was predicted to have a positive relationship to the slope of cohesion across time. Group size was hypothesized to positively relate to the development of cohesion as well given empirical evidence that group size is related to greater cohesion (Burlingame et al., 2011a). Finally, treatment protocol (either process-oriented or forgiveness-based), was hypothesized to relate to cohesion. Consistent with past research, process-oriented groups (i.e. those focusing specifically on fostering interpersonal connections) were hypothesized to show a more positive development of cohesion across time.

Hypothesis 5: Leader contributions. The first set of hypotheses related to leader contributions was that changes in cohesion across time could be predicted by the presence and quality of leader behaviors. Specifically, three broad domains of leader interventions with support in the literature (Group structuring, Verbal Interaction, and Emotional Climate; Burlingame et al., 2002; Chapman et al., 2010) were predicted to differentially relate to cohesion

across the lifespan of a group. Overall, the frequency and quality of all three types of interventions were hypothesized to positively relate to the development of cohesion across time. Still, structuring behaviors were hypothesized to be most important during the first session of a psychotherapy group and demonstrate the strongest relationship with both initial cohesion scores as well as changes in cohesion across time. In addition, possible two-way and three-way interaction effects between these behaviors were examined.

CHAPTER 3. METHOD

Use of Existing Data

Existing data came from psychotherapy groups led at a community counseling clinic located in the Midwest. The data were collected during a two and a half year period—September of 2009 to May of 2012—for the purpose of a larger study funded by the Fetzer Institute. Approval for this study was received from Iowa State University's Institutional Review Board (IRB) under IRB number 09-202 (see *Appendix A*). This larger study sought to examine the impact of group counseling interventions in helping individuals overcome interpersonal hurts and increase forgiveness (see Wade et al., 2014). The existing dataset included measures completed after weekly group therapy sessions as well as measures taken at pre, mid, and post treatment stages (See *Appendix D* & *E*). Along with these measures, all psychotherapy sessions were recorded. The present investigation added to the initial dataset through the behavioral coding of leader interventions in the first psychotherapy session (see *Observational Coding* section below).

Participants

Clients. Of 207 participants screened for the study, 186 were deemed eligible based on the criteria described in the *Procedures* section below. Of these, 58 dropped out of the study before attending the first session leaving 128 participants in the present investigation. One-hundred-and-two (79.7%) participants identified as female and Twenty-six (20.3%) were male. The relationship status of participants included single (22.7%), married (34.4%), divorced (32.0%), separated (7.8%), and widowed (.8%). Forty-two participants (32.8%) stated that they were already receiving therapy at the time of the study. Participants' ethnicities were as follows:

European American (94.5%), African American (3.1%), and other (.8%). Ages of participants ranged from 22 to 82 years old, with the mean age being 50.45 years (*SD* = 13.69 years).

Leaders. Group leaders were five male and nine female therapists; either licensed mental health professionals (*n* = 8) or mental health therapist trainees (*n* = 6). Of the licensed mental health professionals, six were licensed psychologists, one was a licensed mental health counselor, and one was a licensed social workers (MSW). Of the mental health therapist trainees, five were doctoral candidates in a counseling psychology program and one was a master's candidate in a social work program. All non-licensed therapists had successfully completed at least one full year of applied counseling practicum and specific group therapy training including coursework and a practicum experience. The trainees also received weekly individual supervision by a licensed psychologist. The average age of leaders was 34.93 years (*SD* = 11.51). The ethnicities of group leaders were as follows: Caucasian (n=12), Asian (n=1), and Mexican-American (n=1). Leaders averaged 6.81 years of therapy experience (*SD* = 6.77) and 3.71 years of group therapy experience (*SD* = 3.77). The primary theoretical orientations of group leaders were Integrative or Eclectic (n=8), Emotion-Focused (n=2), and Humanistic (n=3); one leader did not identify with a theoretical orientation.

These fourteen leaders were randomly assigned to treatment conditions and groups. The average number of groups led across these fourteen therapists was 1.64; with the number of groups led ranging from 1 to 3. The nesting of groups within therapists (that is, dependency that arose from having a single therapist lead more than one group) was examined as a part of the study procedures (see the *Analytic Strategy* section below).

Groups. Eligible participants were assigned to one of two treatment conditions: a manualized forgiveness intervention (*n* = 12) and a manualized process condition (*n* =11)

totaling 23 groups in the present investigation. The forgiveness treatment condition used an intervention designed explicitly to promote forgiveness to assist individuals in overcoming an interpersonal hurt, whereas the general process condition used a parallel treatment manual that included interventions typical of a general process group. All groups were comprised of three to nine individuals (M=6.14, SD=1.35). Group sessions were ninety-minutes once a week and met for a total of eight weeks. The average number of sessions attended by participants was 6.27 (SD=2.21).

During the recording of the first session of these 23 psychotherapy groups, two groups were improperly recorded and resulted in no data for leadership behaviors (the group-level predictor), leaving 21 groups. While all unconditional models and level 1 analyses are unaffected by this (see *Analytic Strategy* section below), conditional models involving leadership behaviors used a reduced sample size of 112 participants across 21 groups. This is further discussed in the *limitations* section below.

Results regarding the treatment efficacy of each condition have been examined elsewhere (see Wade et al., 2014). For the present investigation, differences in the frequency and quality of leadership behaviors between the two treatment conditions as well as in the trajectory of cohesion across time were examined.

Measures

Client characteristics.

Psychological functioning. Psychological functioning was measured using the Clinical Outcomes in Routine Evaluation—Outcome Measure (CORE-OM). The measure consists of 34 items that are rated on a five-point Likert-scale. The CORE-OM assesses four domains: Risk, Problems, Functioning, and Subjective Well-being (Barkham et al., 2001). For the present study,

only the Well-being, Problems, and Functioning subscales were used, resulting in 28 items. The CORE-OM has internal reliability coefficients ranging from .70 to .90 for each of the four domains, and from .93 to .95 for the full scale (Barkham et al., 2001; Barkham et al., 2005). One-week test-retest reliability coefficients are .64 for Risk, .87 for Problems, .87 for Functioning, .88 for Subjective Well-Being, and .90 for the full scale. Construct validity evidence comes from the full scale's correlation with other diagnostic inventories including the Beck Depression Inventory-II (r = .75) and the Hamilton Rating Scale for Depression (r = .67; Barkham et al., 2005).

Attachment style. Attachment style was assessed using the 12-item Experiences in Close Relationships Scale-Short Form (ECR-S; Wei et al., 2007). The ECR-S is a shortened version of the Experiences in Close Relationships Scale (ECR; Brennan, Clark, & Shaver, 1998), which measures the ways in which adults attach to significant others on the two dimensions of anxiety and avoidance. The ERC-S contains six items that assess anxious attachment (e.g., "I worry that romantic partners won't care about me as much as I care about them") and six items that assess avoidant attachment (e.g., "I try to avoid getting too close to my partner"). Respondents rate the degree to which they agree or disagree with each statement on a 7-point Likert scale ranging from 1 = *strongly disagree* to 7 = *strongly agree.* Past research on the construct validity of the ECR-S indicates that the Anxious and Avoidant subscales are related to theoretically-relevant variables such as loneliness and interpersonal distress. Adequate internal consistency reliability (r = .88) and test-retest reliability (r = .80) have also been observed (Wei et al., 2007).

Self-esteem. Self-esteem was assessed using the Rosenberg Self-Esteem scale (RSE; Rosenberg, 1965). The RSE is a 10-item scale used to measure global self-esteem. Items include such statements as "On the whole, I am satisfied with myself." For the present study,

participants were asked to rate their responses using a 4-point scale, with higher scores indicating higher self-esteem. Previous support for the validity of the RSE comes from its correlation with other self-esteem measures as well as negative correlations with measures of anxiety and depression. Evidence of the internal consistency of the RSE is also adequate (Cronbach's α = .87; Rosenberg, 1965).

Group-level variables.

Leader Counseling Self-Efficacy. The Counseling Self-Estimate Inventory (COSE; Larson et al., 1992) is a 37-item measure of counselors' self-perceived ability to counsel clients in the near future. Items are rated on a 6-point Likert-scale and include such statements as "I am confident that I was able to conceptualize my client's problems." The COSE's internal consistency has been rated at α = .93, with 3-week test-retest reliability at *r* = .87 (Larson et al., 1992). Validity evidence for the scale comes from its moderate to high correlations with counselor self-esteem, outcome expectations, and increases in pre-post assessments before and after clinical training (Larson et al., 1999).

***Leader interventions*.** The Group Psychotherapy Intervention Rating Scale (GPIRS; Chapman, Baker, Porter, Thayer, & Burlingame, 2010) was used as an inventory of leader behaviors. The GPIRS involves rating each counseling session according to three domains; group structuring, verbal interactions, and creating and maintaining an emotional climate. Since the 2010 publication of the GPIRS, the authors have proposed several modifications to the original scale. These include reducing the number of items from 48 to 36 based on poor-performing items and reducing the scale options from 5 to 3 based on the reliability of ratings (G. Burlingame, personal communication, Jan. 27, 2014). For the updated GPIRS, items are scored on a 3-point Likert-type scale, where '0' equals *intervention did not occur*, '1' equals *Ambiguous*

- *occurred but clarity could be improved*, and '2' equals *Intervention was performed with clarity* (Burlingame, personal communication, Jan. 27, 2014). The GPIRS is scored on the basis of observable leader behaviors by independent raters (see *Training Research Assistants* section below). Concurrent validity evidence for the updated GPIRS comes from its correlations with Hill's (1971) Hill Interaction Matrix (HIM; r = .69) and the Group Climate Questionnaire Engaged subscale (r =.40, Chapman et al., 2010). The GPIRS Group Structuring domain has a reported internal consistency score of α = .75, the Verbal Interaction domain an internal reliability of α = .89, and the Emotional Climate domain an internal reliability of α = .81. The overall internal consistency of the GPIRS has been reported as α = .93 (Chapman et al., 2010).

Consistent with the initial authors' recommendations, the GPIRS was scored at identified intervals (in the present study these were set at 5-minute intervals) across each session using the proposed 3-point scale. Items were then averaged across each of these intervals, and then items were averaged across each of the GPIRS domains to create subscales. In this way items and scale scores reflected both the quality of an intervention as well as the quantity of that intervention. Scores close to 0 represented the behavior not occurring at all during the session, scores of 1 represented a frequently-occurring poor-quality behavior, and scores of 2 represented a high frequency of high quality behavior.

As the initial authors note, the GPIRS thus uses a two-tiered scoring system, where scores reflect both quality and quantity (Chapman et al., 2010). When a scoring system reflects two types of scales (quality being a continuous variable and frequency being a count or categorical variable), Cronbach's α and other methods relying on parametric distribution assumptions may not be appropriate (see *Leadership Behaviors* section below). As such, while Cronbach's α is listed here, I also propose an alternative measure of internal reliability using a Latent

Confirmatory Factor Analysis (LCFA) model. Using Cronbach's α, reliability of the GPIRS scales in the present study were α = .55 for the Group Structuring domain, α = .68 for the Verbal Interaction domain, α = .01 for the Emotional Climate domain, and α = .63 for the overall internal reliability of the GPIRS. CFA analysis resulted in alpha estimates of α = .63 for the Group Structuring domain, α = .77 for the Verbal Interaction domain, no model fit for the Emotional Climate domain and thus no alpha estimate, and an overall GPIRS of α = .82. The CFA alpha estimate and decisions regarding the final GPIRS measure are discussed in the *Leadership Behaviors* as well as the *Limitations and Future Directions* sections below.

Session Attendance, Membership, and Treatment Condition. Qualities of each psychotherapy group were also examined in the present investigation. These were examined as simple count variables, with attendance entered as a group aggregate (average) of the number of sessions attended by each member, where the minimum number of sessions allowed for inclusion in the present study was one and the maximum number of possible sessions attended was eight. Membership was defined as the total number of members within a group and treated also as a group-level variable. Finally, treatment condition was entered as a dichotomous variable (0 or 1) reflecting whether a group was a manualized forgiveness intervention (0) or manualized process-oriented group (1).

Client- and leader-rated cohesion. Cohesion was assessed with two of the most frequently cited cohesiveness measures in the literature; the Engagement subscale of the Group Climate Questionnaire – Short Form (GCQ-S; MacKenzie, 1983) and the Cohesiveness subscale of the Group Environment Scale (GES; Wilson, 2008). These two measures were used as indicators The Engagement subscale of the GCQ-S consists of 5 items which describe a positive group climate such as "The group members liked and cared about each other" (MacKenzie,

1983). Items are rated on a 7-point Likert scale (from 1 to 7, where 1 = *Not at all*, 4 = *Moderately*, and 7 = *Extensively*). The internal consistency of the Engaged subscale has been reported as α = .92 (Kivlighan & Goldfine, 1991). The external validity of the Engaged subscale has been demonstrated through moderate correlations with measures of group cohesion and member bonds (Johnson, et al., 2005).

The intervention Group Experiences Scale (GES; Wilson et al., 2008) was developed based on the original Group Environment Scale (Moos; 1994) as a brief measure of group environments. The GES contains 3 subscales; Implementation & Preparedness, Counterproductive Activity, and Cohesiveness. Only the Cohesiveness subscale was used in the present investigation. In the present investigation, items were rated on a 6-point Likert scale with 1 = *strongly disagree* and 7 = *strongly agree*. Internal reliability the Cohesiveness subscale has been reported as *a* = .87. Evidence for the scale's construct validity comes from its association with higher attendance and improved social well-being (Wilson et al., 2008). Because scores used the same likert scaling and represent the same construct, scores were summed across the 5-item GCQ and the 8-item GES to aid in interpretability. This resulted in a possible scale range of 13 to 91.

Procedures

Participants were recruited through advertisements in their local community for a group counseling intervention to promote coping with interpersonal hurts. Interested individuals contacted the research clinic conducting the study and received initial eligibility criteria. Individuals who contacted the clinic attended a screening appointment with a counseling psychology doctoral student, at which time they signed an informed consent document and completed a questionnaire packet. The doctoral student then confirmed that the participant met

inclusion criteria, which included experiencing a significant interpersonal hurt which they (a) had not been able to overcome and (b) wanted to work on in therapy. The screeners also determined that participants did not meet exclusion criteria, which included (a) suicidal or homicidal plan or intent over the previous year and (b) a lifetime history of a psychotic disorder. Eligible participants were then randomly assigned to one of three treatment conditions; a manualized process group, a manualized forgiveness-oriented group, or a waitlist control (see *Treatment Groups* above). Once their treatment group was formed, participants received notification and then started treatment. If participants did not meet the criteria, they were provided a list of alternative treatment options in the local community.

Members were asked to complete a detailed questionnaire (See *Appendix D*) at the beginning of the first session, after the fourth session, and after the final (eighth) group session. At the end of each group session the group leader and clients were asked to complete a short, ten-minute questionnaire (see *Appendix E*). Upon completion and submission of group questionnaires, group members were thanked and reminded of the following week's appointment. Participants were provided monetary compensation for their time at the end of the pre-, mid-, post-treatment, and follow-up questionnaires for a combined total of forty dollars.

Training Research Assistants

There is strong evidence that even undergraduate trainees with little to no clinical experience can be trained to achieve good to excellent reliability (e.g. r = .84-.90) while rating counseling sessions (Honos-Webb, Stiles, & Greenberg, 2003). For the present investigation, eight undergraduate research assistants, who had completed foundational courses in psychological research methods, assessment, and statistics, were trained in observational coding. The research assistants were trained in use of the GPIRS (Chapman et al., 2010) across a four-

week training period. During this time, and consistent with the recommendations of Hill and Lambert (2004), the author was highly involved in the training of the raters to ensure the measures were used as intended.

Several procedures were used in order to ensure rater validity. Before the first meeting, research assistants read the training manual for the study (*Appendix H*) which included coding guidelines for each measure (*Appendix G*). During the first training session the clinical premises for the scoring criteria, use of the scales, and possible biases (e.g. halo effects) were reviewed and questions regarding scoring were addressed. During the second training session the principal investigator reviewed a videotaped psychotherapy session together with the research assistants, stopping the tape to discuss rated items. Finally, during the third and fourth sessions, research assistants were asked to review two additional training videos of a different counseling session and rate them using the scales. The principal investigator and trainees met to discuss any discrepancies between independent raters and the principal investigator's scoring. Raters were given individualized feedback regarding the validity of their ratings during each of these four training sessions.

RAs were trained for reliability using Stein and colleagues (2010) discrepancy analysis procedure. Here, 30-minute segments of recorded group therapy sessions from the project were divided into 5-minute intervals. All research assistants involved in the coding process rated segments. RAs were given individual feedback including their rating and the group average rating for that item. Items in which more than 75% of RAs (6 or more members) agreed on were used to provide feedback to the one or two RAs who were inconsistent with the majority. RAs reviewed these sections of tape in which there was disagreement and informed the principal investigator if they either agreed or disagreed with the ratings of the majority. If they disagreed,

RAs provided written feedback to the principal investigator and this was reviewed anonymously by the larger group. Any items on which there was only 33-66% agreement were discussed as a group. During this process, item wording and anchors on the GPIRS were revised to clarify wording and increase reliability of coding while preserving the initial intent of the items (see *Appendix H*). This feedback process was continued until interrater reliability exceeded 60% for a given training video (defined as 'good' interrater reliability, Shrout & Fleiss, 1979).

Research Assistant Interrater Agreement

After week two, three, and four of the training period, an Intraclass Correlation Coefficient (ICC) using SPSS software was calculated. Several facets of the present investigation helped determine the most appropriate calculation of ICC. First, coders were assigned to rate each 5-minute interval on the GPIRS 3-point rating scale. This led to agreement being calculated based on when raters agreed both when a behavior *did* and when it *did not* occur. As such, agreement on ratings of zero was used in the final calculation of interrater reliability. Second, in order to ensure the reliability of coded psychotherapy sessions, two raters of each psychotherapy session were used (see Hallgren, 2012 for a discussion). Here, each psychotherapy session was randomly assigned to two of the nine RAs involved in the study, leading to each RA coding two to three sessions. In cases where the set of coders is randomly selected from a larger population of possible coders, a one-way random effects model is appropriate. Additionally, because the average of the two raters was used for every coded session, averages served as the unit of analysis in the ICC results. In such cases an average measures ICC is considered the best measure of reliability (Hallgren, 2012). Still, in cases where there is a large discrepancy between single measures and average measures ICC, both should be reported (Shrout & Fleiss, 1979). Finally, interrater reliability analysis should be performed on

the variables as they will be used in the final study. In the present investigation, the GPIRS items were rated at five-minute intervals but averaged across the length of a session before summing to create each subscale and full scale. As such, reliability analysis was conducted on the averaged ratings across all 5-minute intervals for a given session.

An average measures ICC under a one-way random effects model (as specified above) was used to determine adequate rater agreement. After the second training session the nine research assistants received an average measures ICC of .92 (95% CI = .90, .94) and a single measures ICC of .57 (95% CI = .51, .64). After the third training session, RA's received an average measures ICC of .95 (95% CI = .94, .96) and a single measures ICC of .70 (95% CI = .64, .75). After the fourth training session, RAs received an average measures ICC of .96 (95% CI = .95, .96) with a single measures ICC of .71 (95% CI = .69, .74). At this point both single measures and average measures were considered within the 'good' range for interrater reliability (Shrout & Fliess, 1979), and RAs were randomly assigned to code group psychotherapy sessions for leadership behaviors, with once monthly check-ins to correct for rater drift.

Analytic Strategy

Prior to examining the development of cohesion through Latent Growth Curve analysis (LGC), several preliminary examinations of the data were necessary to determine the model. First was an examination of patterns of missing data. Second was an examination of the effects of treatment condition to determine whether significant differences existed between treatments and if condition should be included in the final LGC model. Third was an examination of the leadership behaviors present, if research assistants agreed to a satisfactory degree on these observed behaviors, and if the quality of such behaviors were appropriate for examination through LGC. Finally was an examination of the outcome variable, cohesion, for levels of group

dependency. This was to help empirically determine if member-rated cohesion should be treated as an individual-level (level-two) or group-level (level-three) outcome variable in the final model.

After examination of these criteria, Latent Growth Curve (LGC) analysis under a Hierarchical Linear Modeling (HLM) framework was used. LGC is a form of HLM analysis that accounts for dependencies in nested data in modeling slopes and intercepts. LGC provides a means for analyzing repeated measures and multilevel processes across groups and multiple time points (Kivlighan & Lilly, 1997). Nested data violates the assumption of independence that is integral to most linear models such as Ordinary Least Squares (OLS) which may otherwise generate spuriously significant findings through OLS regression (Guo, 2005). LGC conducts analyses at both individual and group levels and can also account for repeated measures by nesting time-points within individuals, or, alternatively, by modeling time as a multivariate outcome vector (Muthén & Muthén, 1998-2010). LGC determines how much of the variation in the data can be explained by differences between individuals within groups as well as between groups (which, when examined together, is known as "intraclass correlations" or ICC). Researchers can use the ICC to determine whether a nested model is necessary. This is done by determining if there is enough variance between groups, and not just within groups, to examine the effects of group membership (and group variables) on the outcome.

In addition to statistically allowing for an examination of the process of change in repeated measure designs, another benefit of using LGC under an HLM framework is that it can accommodate unequal group sizes (Kivlighan & Lilly, 1997). This is of particular appeal to researchers in counseling psychology due to the sometimes inconsistent attendance of group members or early drop-out of certain members. As such, researchers in counseling psychology

have utilized LGC to investigate changes in group phenomena such as working alliance (Kivlighan & Shaughnessy, 1995) and climate development (Raudenbush, Rowan & Kan, 1991).

For the present study, a three-level linear growth model for continuous outcomes was used (Muthén & Muthén, 1998-2010). Initially, an unconditional model, or a model in which no covariates are included, is specified. Here, only the growth factors of the outcome variable (cohesion) were modelled. In the present study, the slope growth factor (*s*) was fixed at zero, one, and two in the case of a linear growth model, and the growth factor at time one defined the intercept growth factor (*i*) as the initial status factor. A linear model (specifying intercept and slope) as well other models hypothesized by the literature were empirically examined (e.g. a quadratic model in which an intercept, slope, and a quadratic slope factor are specified; Muthén & Muthén, 1998-2010). M-Plus software and Maximum-Likelihood estimation were used to determine how well the growth factors' covariance was estimated and to determine the best fitting model at level-1. This level-1 produced an estimated intercept, regression coefficient, and error term for the simple main effect of time on cohesion. The intercept and regression coefficients of this level-1 time variable then served as a fixed effect (e.g. any given time point will have a determinate cohesion value) for further levels of analysis (Muthén & Muthén, 1998-2010).

After determining the best fitting unconditional model, a conditional model was examined in which covariates were used to determine the values of the intercepts and slopes at levels 2 and 3. Here, the outcome variable, cohesion, was examined at the level of the individual. Fixed effects resulting from modelling a latent cohesion construct at each time point served as dependent variables and were regressed on time-one client characteristics (psychological functioning, attachment style, and self-esteem). A regression equation at level 2

then estimated the intercept, regression coefficient (slope), and an error term for each of the three client characteristics. In such models, error terms are referred to as "random effects" because the model estimates a specific error term value for each individual in the model. Specifying these as random effects corrects for the autocorrelation that occurs as a result of having multiple individuals within groups at level 3.

At level 3, a regression equation that estimates the intercept, regression coefficient, and error term for each of the group characteristics and leader behaviors was conducted. In the present investigation, the leader characteristic of counseling self-efficacy and leader behaviors were modelled as time-invariant, measured only at session 1. Additionally, hypotheses involving different proposed relationships between covariates and phases of group were modelled through "piecewise growth modelling". This involved using more than one slope growth factor and including different time points (e.g. sessions 1-4 and sessions 4-8) for each (Muthén & Muthén, 1998-2010, pg. 115). Effect sizes and significance of these parameter estimates was examined to determine factors influencing change in cohesion across time.

CHAPTER 4. RESULTS

Missing Data

Three components of growth-curve analysis utilizing MPLUS software are relevant to a discussion of missing data. First, MPLUS utilizes a random-effects model to capture individual differences in development in the outcome variable (Muthén & Muthén, 1998-2010). An important part of random-effects models is that individuals are not assumed to be measured at all time-points. A random effects model estimates trend lines on the basis of whatever data a participant has. Thus, all data is included (rather than excluding any participant who does not have data at all time-points; Hedeker & Gibbons, 1997). Secondly, MPLUS utilizes Maximum Likelihood (ML) estimation for all outcome variables. In full likelihood-based procedures, data are allowed to be missing completely at random (MCAR) or missing at random (MAR). In MCAR, levels of missingness are independent of both observed covariates and observed levels of the dependent variable. In MAR, levels of missingness in the dependent variable are related to observed covariates but not to observed previous levels of the dependent variable (Hedeker & Gibbons, 1997). If data have non-ignorable missing data (NMAR), meaning levels of missingness depend both on observed levels of the dependent variable, growth-curve modelling is still possible as long as categorical outcomes serving as indicators of missingness can be predicted by continuous and categorical latent variables (Muthén & Muthén, 1998-2010). Finally, missingness is not allowed for observed covariates as these are not a part of the growth curve model. Still, multiple imputation of missing data using Bayesian analyses can be used for covariates to improve model estimation (Muthén & Muthén, 1998-2010). Multiple imputation was used for any participant that had at least 80% completion of a given questionnaire. For all

those questionnaires with less than 80% completion, pairwise deletion was used consistent with the recommendations of Schlomer, Bauman, and Card (2010).

With these factors in mind, items in the outcome variable, cohesion, were first examined for levels and patterns missingness. Levels of missingness were found to increase across time consistent with the attrition of study participants. At week one the amount of missing data in the outcome variable was 7.6%. This went up to 20.0% at week two, 19.5% at week 3, 21.5% at week 4, 24.6% at week 5, 25.9% at week 6, 25.2% at week seven, and 26.7% at week 8; suggesting that the largest drop-out was between weeks one and two. Levels of missingness in this range is not uncommon in longitudinal research (Peng et al., 2006), however it can still pose problems for the power of an analysis to detect an effect (Schlomer, Bauman, & Card, 2010). As such, an examination of actual power is presented in the *Discussion* section below.

Time-related missingness in the outcome variable can still fall within the MCAR definition in any model in which time is included (Hedeker & Gibbons, 1997), as is the case in the present LGC, as long as levels of missingness do not depend on early levels of the dependent variable. As a first step in examining this, Little's MCAR test was used. Little's MCAR test splits the sample into two groups, one in which a missing data point from a given variable is included and one in which a present data point from that same variable is included. The means of the observed value and the missing value groups for all variables are then compared simultaneously under an asymptotic chi-square distribution (Little, 1988). If the test is non-significant, there is evidence that missingness in any one item in the outcome variable is not dependent on a previously observed value. Little's MCAR test was found to be non-significant ($\chi^2_{3905} = 3814.544$, $p = .847$), indicating non-dependence in the outcome measure. Still, because of the large degrees of freedom in this test ($df = 3905$), there were concerns of Little's MCAR

test being overpowered in the rejection of the null hypothesis of independence. As such, levels of missingness of the outcome were examined through binary logistic regression using first-session cohesion as a predictor.

To do so, clients with missing data points were coded as either monotone (meaning all subsequent longitudinal data was missing) or non-monotone (meaning at least one subsequent session was present). This procedure revealed that 31.3% of participants had complete data, 44.5% had non-monotone missing data, and 24.2% of participants had monotone missing data (meaning these participants dropped-out of the study). A multinomial logistic regression was then conducted to examine if early levels of cohesion predicted monotone and non-monotone missingness. This analysis revealed that cohesion at session 1 predicted monotone missingness (B = -.042, Wald χ^2= 4.491, p = .034) but not non-monotone missingness (B = -.022, Wald χ^2 = 1.77, p =.183). This indicated that lower cohesion at session 1 predicted later drop-out, and thus that the data was NMAR.

Given the NMAR quality of missingness, in order for the assumptions of growth-curve modelling to be met, categorical outcomes serving as indicators of missingness must be predictable from other variables in the model (Muthén & Muthén, 1998-2010). To examine this, covariates in the dataset, both those in the present analysis and others from the larger Fetzer Institute study, were examined as predictors of missingness. Again, a multinomial logistic regression was conducted to examine non-missing, monotone missing, and non-monotone missing data. Bivariate correlations were inspected for variables including 1) client predictors in the present investigation (psychological distress, attachment style, and self-esteem); 2) demographic variables including age, gender, past therapy experience, relationship status, and race; 3) study variables including facilitator and treatment condition; and 4) other client variables

from the larger Fetzer study including variables such as religious commitment, trait rumination, and trait forgiveness. These variables were first examined independently for a relationship with missingness. No variables were found to predict non-monotone missingness, while four variables were found to predict monotone missingness (i.e. dropout). These included age, psychological distress, religious commitment, and race. Entering these into a single multinomial regression equation significantly predicted missingness (-2LL = 242.04, $\chi^2_{10} = 31.15$, $p = .001$). Only one of the covariates remained a significant predictor in the final model, however; initial psychological distress (B = 1.585, Wald $\chi^2 = 6.63$, $p = .01$). To further examine this relationship, the effects of psychological distress and early cohesion were examined for those who dropped out of the study after session 1 (the largest increase in drop-out rates between the time points). Here, a bivariate logistic regression with initial psychological distress and time 1 cohesion predicting drop-out (Yes or No) was conducted. This revealed that while psychological distress predicted drop-out ($\beta = 1.02$, Wald $\chi^2 = 5.82$ $p = .02$) wave 1 cohesion was no longer a significant predictor ($\beta = -.012$, Wald $\chi^2 = .333$, $p = .56$). This provided some evidence that persons did not drop out of the study due to low cohesion values but rather due to higher levels of psychological distress, indicating that an increase of one standard deviation in psychological distress increased the likelihood of drop-out by 2.7 to 1 (this finding is further discussed in the *Discussion* section below). Thus, because psychological distress was included as a covariate in the model, the assumptions of growth curve modelling were considered to be met (Muthén & Muthén, 1998-2010).

To examine the effects of missingness on the trajectory of cohesion across time, a dichotomous drop-out indicator (DROP) was created that reflected whether respondents 1) provided complete or only partially-missing data (non-monotone missingness) or 2) dropped out of the study at any time. A significant effect of time on cohesion indicated that responses increased on average

over time. The drop-out indicator did not significantly influence the trajectory of cohesion across time. The results of this analysis indicated that study attrition did not relate to changes in cohesion. Taken together, these results indicated that study attrition did not relate to one's level of cohesion with the group.

Leadership Behaviors

As a first step in the examination of leadership behaviors, interrater agreement across the 21 coded therapy sessions was first examined. Intraclass correlates (ICC) values for the two-raters of each session ranged from .68 to .99, with the average ICC being .91 [Average 95% CI = .83, .96] across the 21 sessions. This indicated strong interrater agreement and supported both the reliability of ratings as well as the appropriateness of averaging coder ratings of each session as a measure of leader behaviors.

Using average rater agreement values, the GPIRS items were first examined for item-level and scale-level characteristics. First, the item-level means and standard deviations were examined (see *Table 1*). Two items never occurred in the 21 coded sessions; item 25 ('leader reframed injurious member-member feedback') and item 26 ('leader restated corrective member-member feedback'). Importantly, these behaviors, which focus on providing interpretive and corrective feedback, are theorized to occur in the later stages of group (e.g. working and ending phases; Morran, Stockton, & Whittingham, 2004). Because these behaviors did not occur, these items provided no variability in the prediction of outcomes and were excluded from further analysis.

What follows is a description of the most frequent and least frequent behaviors across the 21-coded first-session psychotherapy groups. To aid in the interpretability of this section, item means can be understood as the average occurrence of a given behavior across all 5-minute

sections. So, during a 90-minute session for example, an item mean of .10 would equate to the behavior occurring one time with good quality (.10 * 18 = 1.8). At the middle of the scale (values close to 1), the "two-tiered" scoring system of the GPIRS makes interpretability less clear. For example, a mean of 1.0 could indicate a behavior occurring within every five-minute interval with poor clarity or approximately half of the time with good clarity. Values close to 2, however, represent the behavior occurring at nearly every five minute interval with high clarity.

The four most frequently occurring leadership behaviors were all within the Emotional Climate domain. These included "maintained an active style of engagement with the group and its work" (item 14; *M* = 1.91, *SD* =.10), "used nonjudgmental language or non-verbals with members" (item 15; *M* = 1.77, *SD* =.34), "modeled expressions of open and genuine warmth to all members who share throughout section" (item 16; *M* = 1.54, *SD* =.29), and "recognized and responded to the meaning of group member's comments" (item 17; *M* = 1.21, *SD* =.26). The next most frequently occurring behaviors were "provided structure that facilitates overall member interaction" (Structuring domain; item 4; *M* = .87, *SD* =.31), "encouraged self-disclosure relevant to the current group agenda without 'forcing it'" (Emotional Climate domain; item 9; *M* = .83, *SD* =.30), and "set group agenda" (Structuring domain, item 1; *M* = .56, *SD* =.18). The two least-frequent behaviors were within the Emotional Climate domain and included "stopped attacking or judgmental situations and expressions between members" (item 18; *M* = .00, *SD* =.01) and "refrained from conveying personal feelings of hostility and anger in response to disruptive member behavior" (item 36; *M* = .00, *SD* =.01). The next least frequent leader behaviors were "attended to a balance between positive and corrective member-member feedback" (Verbal Interaction, item 28; *M* = .01, *SD* =.03) and "helped members apply in-group

member-member feedback to out-of-group situations" (Verbal Interaction, item 30; *M* = .01, *SD* =.02).

Table 1. GPIRS Item-level Statistics

Item	Item wording	Mean	SD	Range	Skew	Kurtosis
1	Set group agenda	.56	.18	.30-.94	.38	-.75
2	Described rationale for treatment	.30	.15	.05-.63	.33	-.60
3	Discussed group rules	.30	.13	.05- .50	-.17	-.80
4	Provided structure for member interaction	.87	.31	.21- 1.40	-.34	-.54
5	Interacted in here-and-now manner	.21	.25	0-1.09	2.37	7.37
6	Modeled information self-disclosure	.37	.18	.11-.74	.37	-.94
7	Modeled feeling self-disclosure	.12	.10	0-.31	.75	-.56
8	Managed group when "out of control"	.04	.06	0-.24	2.58	8.20
9	Encouraged self-disclosure	.83	.30	.28-1.44	.19	-.49
10	Encouraged here-and-now exchange	.03	.05	0-.67	1.59	2.13
11	Interrupted dominant member	.02	.05	0-.21	3.68	14.77
12	Shared brief personal experience	.10	.11	0-.42	1.79	3.39
13	Not defensive when made mistake	.32	.45	0-1.21	1.23	-.31
14	Actively engaged with group	1.91	.10	1.64-2.00	-1.23	1.36
15	Used nonjudgmental language	1.77	.34	.93-2.00	-1.86	2.18
16	Modeled expressions of warmth	1.54	.29	.91-2.00	-.06	.04
17	Responded to meaning of member comment	1.21	.26	.82-1.65	.24	-1.22
18	Stopped a judgmental exchange	.00	.01	0-.03	4.58	21.00
19	Discussed fears/concerns about group	.22	.20	0-.68	.97	.25
20	Discussed roles and responsibilities	.17	.16	0-.62	1.47	2.15
21	Modeled member-member behavior	.02	.04	0-.15	2.21	4.33
22	Facilitated member-member interaction	.11	.13	0-.59	2.52	8.08
23	Educated members on self-disclosure	.15	.17	0-.59	1.43	1.44
24	Elicited member-feeling disclosure	.03	.05	0-.20	2.02	4.02
27	Used consensus to reinforce feedback	.04	.06	0-.23	1.90	3.02
28	Balanced positive and corrective feedback	.01	.03	0-.12	2.98	7.62
29	Encouraged positive feedback	.02	.04	0-.16	2.59	6.00
30	Apply in-group feedback to out-group situation	.01	.02	0-.06	2.98	7.61
31	Helped members identify underlying concerns	.15	.13	0-.44	.87	-.21
32	Encouraged engagement between members	.09	.15	0-.65	2.80	9.05
33	Fostered climate of supportive challenge	.07	.10	0-.39	2.11	4.92
34	Assisted members in describing emotions	.12	.11	0-.38	.91	-.25
35	Elicited verbal expressions connecting members	.09	.12	0-.46	1.88	3.40
36	Did not convey personal feelings of anger	.00	.01	0-.03	2.97	7.56

Note: n = 112, j = 21.

A notable feature of the GPIRS scale in the present investigation is the high frequency of zeros and resultantly skewed data (see *Figure 2*). As with Chapman and colleagues' (2010) validation of the GPIRS, scores represented both frequency (a count or categorical variable) and quality (a continuous variable). In such instances, items can take on discriminal distributions, or independent rating anchors with distributions about each point on the rating scale (i.e. distributions about values of 0, 1, and 2; Hayes & Embretson, 2012). While discriminal distributions are common in stimulus-centered variables where observers use anchor points, they can pose problems for traditional parametric statistics which rely on assumptions of normality and homogeneity of variance (Hayes & Embretson, 2012; Sheng & Sheng, 2012; Kendall, 1958).

In cases where skew comes from a high percentage of zeros or discriminal distributions, transformations offer little help because they do not change the fact that so many scores have the same value. In such instances, it is recommended that researchers use non-parametric statistics (i.e. distribution-free methods) to describe and test data. To examine if non-parametric tests are appropriate for a given item, skew and kurtosis can be evaluated using a z-score distribution (i.e. $Z_k = \frac{K-0}{SE_K}$; Corder & Foreman, 2014). Here, values of skew and kurtosis for individual items can be compared to an assumed normal distribution (Corder & Foreman, 2014). Using this method, the following GPIRS items were found to deviate significantly from normal at a Bonferonni-adjusted $p \leq .001$: Items 5, 8 , 11, 12, 15, 18, 21, 22, 24, 27, 28, 29, 30, 32, 33, 35, and 36 (the distributions for these items are in *Figure 1*). This provided strong evidence for the necessity of using non-parametric methods and measures to make decisions about internal consistency of the GPIRS as well in examining the GPIRS relationship to other variables in the model.

Figure 1. Distributions of the Scored GPIRS Items across 21 first-session Psychotherapy Groups

Note: n = 112, j = 21. X-axis reflects GPIRS anchors; 0 = intervention did not occur, 1 = intervention ambiguously performed, 2= intervention performed with clarity. Y-axis (frequency) is scaled 0-20. Highly skewed/leptokurtic items include item 5, 8, 11, 12, 15, 18, 21, 22, 24, 27, 28, 29, 30, 32, 33, 35, and 36. Vertical lines depict discriminal distributions (Hayes & Embretson, 2012) of many of the GPIRS items, with distributions at each of the three GPIRS anchors of 0, 1, and 2.

Given the discriminal distributions of the GPIRS items, Cronbach's α coefficients may not be the most appropriate test statistic (see Sheng & Sheng, 2012). As an alternative to Cronbach's α, Latent Confirmatory Factor Analysis (LCFA), which optimizes the weighting of items based on the covariance between items (Muthén & Muthén, 2012), can be used. Using LCFA, an internal consistency coefficient can be computed, where the variability due to the factor (a true score) is compared to the variability due to error (see Suhr, 2003). Using this method, an alternative alpha using a LCFA framework was computed for each GPIRS domains and the GPIRS total score. This resulted in alpha estimates of $\alpha = .63$ for the Group Structuring domain, $\alpha = .77$ for the Verbal Interaction domain, and no model fit for the Emotional Climate domain. The overall GPIRS alpha was found to be $\alpha = .82$. The item-total and item-domain factor loadings for each of the GPIRS items are reported below in *Table 2.*

Table 2. Group Psychotherapy Intervention Rating Scale (GPIRS) – Scale-level Statistics

Item #	Domain	Item wording	Item-Total Factor Loading	Item-Domain Factor Loading
1	Strct	Set group agenda	$-.16^{\dagger}$	$-.07^{\dagger}$
2	Strct	Described rationale for treatment	.48	.43
3	Strct	Discussed group rules	.73	.72
4	Strct	Provided structure for member interaction	$-.16^{\dagger}$	$.02^{\dagger}$
5	Vrbl	Interacted in here-and-now manner	.69	.67
6	Vrbl	Modeled information self-disclosure	.59	.65
7	Vrbl	Modeled feeling self-disclosure	$.26^{\dagger}$	$.34^{\dagger}$
8	Vrbl	Managed group when "out of control"	.58	.57
9	Vrbl	Encouraged self-disclosure	$.31^{\dagger}$	$.21^{\dagger}$
10	Vrbl	Encouraged here-and-now exchange	$.30^{\dagger}$	$.18^{\dagger}$
11	Vrbl	Interrupted dominant member	$-.16^{\dagger}$	$-.14^{\dagger}$
12	Vrbl	Shared brief personal experience	.65	.78
13	Emtl	Not defensive when made mistake	$-.01^{\dagger}$	N/A
14	Emtl	Actively engaged with group	$-.12^{\dagger}$	N/A
15	Emtl	Used nonjudgmental language	$.08^{\dagger}$	N/A
16	Emtl	Modeled expressions of warmth	$-.26^{\dagger}$	N/A
17	Emtl	Responded to meaning of member comment	$.33^{\dagger}$	N/A
18	Emtl	Stopped a judgmental exchange	$-.04^{\dagger}$	N/A

Table 2 continued

19	Strct	Discussed fears/concerns about group	.81	.90
20	Strct	Discussed roles and responsibilities	.50	.60
21	Vrbl	Modeled member-member behavior	.55	.48
22	Vrbl	Facilitated member-member interaction	-.10†	-.21†
23	Vrbl	Educated members on self-disclosure	.89	.86
24	Vrbl	Elicited member-feeling disclosure	.22†	.07†
27	Vrbl	Used consensus to reinforce feedback	.00†	-.17†
28	Vrbl	Balanced positive and corrective feedback	.59	.74
29	Vrbl	Encouraged positive feedback	.45	.30†
30	Vrbl	Helped apply in-group feedback to out-group situation	.72	.75
31	Emtl	Helped members identify underlying concerns	-.01†	N/A
32	Emtl	Encouraged active engagement between members	.17†	N/A
33	Emtl	Fostered climate of supportive challenge	.41	N/A
34	Emtl	Assisted members in describing emotions	.52	N/A
35	Emtl	Elicited verbal expressions connecting members	.49	N/A
36	Emtl	Did not convey personal feelings of anger	.43	N/A

Note. n = 112, j = 21. N/A = item-domain correlation not applicable due to poor model fit of Emotional Climate domain. † = Item-domain correlation ≤ .40

Fortunately, because leadership behaviors serve as level-two covariates in the present LGC, assumptions on the distributions of the predictors are not implied (Curran, Obeidat, & Losardo, 2010) and thus skewed variables do not strictly violate the assumptions of the model used in the present investigation. Still, highly skewed variables can attenuate any linear relationship in a latent growth curve (LGC) analysis and inflate standard errors (Duncan, Duncan, & Strycker, 2006). In order to improve the internal consistency of the GPIRS scale and reduce attenuation in the final model, item-factor loading of the GPIRS items were used to remove those with low commonalities. Items sharing little common variance with the GPIRS domains were considered to be those with factor loadings of less than .40, consistent with the recommendations of Worthington and Whittaker (2006).

This procedure resulted in the removal of two items from the Structuring Domain (items 1 and 4), eight items from the Verbal Interaction Domain (items 7, 9, 10, 11, 22, 24, 27, & 29),

and six items from the Emotional Climate Domain (items 13, 14, 15, 16, 32, and 35). The final GPIRS scale thus consisted of a total of 16 items; four Group Structuring Items, eight Verbal Interaction items, and four Emotional Climate items. These items, as well as their final item-total and item-domain factor loadings, are presented in *Table 3.* Subscale level statistics indicated improvements in the normality of distributions, where a corrected test of kurtosis and skew ($Z_k = \frac{K-0}{SE_K}$; Corder & Foreman, 2014) revealed a non-significant skew/kurtosis in the Structuring and Emotional Climate domains at a Bonferonni-adjusted $p \leq .01$. The overall GPIRS and the Verbal Interaction domain were still significantly skewed, however (Z = 2.82, p = .002 and Z = 4.88, $p < .001$; respectively).

As a final analysis of the revised GPIRS's model fit, a Chi-square difference test was conducted. Here, factor loadings of the 16 GPIRS items from the revised scale were estimated while the 22 poor-fitting items were fixed to zero, thus creating a nested model for comparison. Conducting a chi-square difference test between the two nested models revealed no difference in overall fit to the data ($\Delta\chi^2_{20} = 25.46$, p = .18), indicating that the more parsimonious model (the revised scale) explained the data equally well to the full 36-item scale. As such, the 16-item revised GPIRS scale was used in all further analysis. A discussion of the appropriateness and limitations of the removal of items for the present investigation is noted in the *Discussion* section below.

Using the 16-item revised GPIRS, the internal consistency of the GPIRS domains was α = .78 for the Group Structuring Domain, α = .89 for the Verbal Interaction Domain, α = .76 for the Emotional Climate Domain, and α =.82 for the full GPIRS scale. This demonstrated considerable improvement in internal reliability compared to the full GPIRS scale in the present study. As such, the revised 16-item GPIRS was used in all further investigations.

Table 3. Revised GPIRS items – Item & Scale-level Statistics

Item	Domain	Item wording	Mean	SD	Item-Total Loading	Item-Domain Loading
2	Strct	Described rationale for treatment	.30	.15	.49	.44
3	Strct	Discussed group rules	.30	.13	.72	.72
5	Vrbl	Interacted in here-and-now manner	.21	.25	.64	.71
6	Vrbl	Modeled information self-disclosure	.37	.18	.62	.60
8	Vrbl	Managed group when "out of control"	.04	.06	.58	.61
12	Vrbl	Shared brief personal experience	.10	.11	.66	.80
17	Emtl	Responded to meaning of comments	1.20	.26	.84	.86
18	Emtl	Stopped a judgmental exchange	.001	.005	.47	.47
19	Strct	Discussed fears/concerns about group	.22	.20	.77	.89
20	Strct	Discussed roles and responsibilities	.17	.16	.47	.61
21	Vrbl	Modeled member-member behavior	.02	.04	.48	.52
23	Vrbl	Educated members on self-disclosure	.15	.17	.92	.80
28	Vrbl	Balanced positive/corrective feedback	.01	.03	.60	.80
30	Vrbl	Helped apply feedback to out-group	.01	.02	.76	.72
31	Emtl	Helped members identify concerns	.15	.12	.95	.51
34	Emtl	Assisted members describing emotions	.12	.11	.53	.73

Note. $n = 112$, $j = 21$.

Descriptive Statistics

Overall means, possible scale ranges, actual scale ranges, standard deviations, and bivariate correlations for the final variables are presented in *Table 4* and *Table 5*. Client-level variables are presented in *Table 4*, with group-level variables shown in *Table 5*. As a representation of possible predictors of the intended outcome of cohesion across time, week one, four, and eight were chosen as a cross-section of the development of cohesion.

Psychological distress was found to significantly correlate with all other client-level predictors, including a moderate, positive correlation with anxious attachment and avoidant attachment ($r = .40$ and $r = .40$, respectively) and a strong, negative correlation with client self-esteem ($r = -.68$). Notably, at the client level, session one and session eight client-rated cohesion were not significantly correlated with any of the client-level predictors to be entered in the

model, while session four was significantly negatively correlated only with avoidant attachment (r = -.36).

Table 4. Descriptive Statistics for Client-level Variables

Variables	1	2	3	4	5	6	7
1. CORE-OM_{PRE}	—						
2. ECR-ANX_{PRE}	.40***	—					
3. ECR-AVD_{PRE}	.40***	.02	—				
4. RSE_{PRE}	-.68***	-.30***	-.51***	—			
5. COH_1	-.16	-.09	-.17	.06	—		
6. COH_4	-.13	.05	-.36***	.14	.63***	—	
7. COH_8	-.01	-.04	-.21	.02	.49***	.68***	—
M	50.19	24.05	21.72	27.40	66.77	79.33	84.05
SD	20.79	6.68	7.25	5.29	12.13	8.82	6.80
Possible Range	0-112	6-72	6-72	10-40	13-91	13-91	13-91
Sample Range	4-100	6-39	7-41	11-39	37-91	57-91	67-91

Note: N = 72-128 .CORE-OM_{PRE}: Clinical Outcomes in Routine Evaluation (Pre-treatment), ECR-ANX_{PRE}: Experiences in Close Relationships - Anxiety scale (Pre-treatment), ECR-AVD_{PRE}: Experiences in Close Relationships - Avoidance scale (Pre-treatment), RSE_{PRE}: Rosenberg Self-Esteem scale (Pre-treatment), COH_1: Client-rated cohesion measure (Week 1), COH_4: Client-rated cohesion measure (Week 4), COH_8: Client-rated cohesion measure (Week 8).
*p < . 05. ** p < .01. *** p < .001

At the group-level, leadership characteristics and leader-rated cohesion were examined and reported in *Table 5*. For the revised GPIRS, Spearman's Rank-Ordered coefficients were used as a nonparametric replacement for Pearson's r given the highly skewed data (Corder & Foreman, 2014). Spearman's Rank-Ordered coefficient compares ranked pairs in the data to examine a positive or negative relationship between two variables. As a nonparametric test, Spearman's Rank-Ordered coefficient is appropriate for non-normal data in samples where $n \geq 4$, and the method has strong support in the literature (Corder & Foreman, 2014).

Utilizing Spearman's Rank-Ordered coefficient for correlations involving the GPIRS scale and Pearson correlations for all other group level variables, several observations can be made. First, the only significant correlations between the predictor variables to be entered into the model and leader-rated cohesion was for session eight, with counselor self-efficacy showing a moderate, positive correlation with session eight cohesion (r = .48). This indicated that leaders

with greater self-efficacy were more likely to rate cohesion more highly, but only in the final session of group. Additionally, there were high correlations between the GPIRS subscales, indicating that leaders who engaged in any of the behavior domains were also more likely to engage in behaviors across domains.

Table 5. Descriptive Statistics for Group-Level Variables

Group Variables	$COSE_{PRE}$	$GPIRS_1$	GPIRS-STR_1	GPIRS-$VRBL_1$	GPIRS-$EMTL_1$	COH_{1L}	COH_{4L}	COH_{8L}
1. $COSE_{PRE}$	—							
2. $GPIRS_1$	.04†	—						
3. GPIRS-STR_1	-.04†	.96***†	—					
4. GPIRS-$VRBL_1$	.13†	.86***†	.75***†	—				
5. GPIRS-$EMTL_1$	.07†	.79***†	.70***†	.54*†	—			
6. COH_{1L}	-.05	.10†	-.04†	.19†	.30†	—		
7. COH_{4L}	.13	-.22†	-.20†	-.15†	-.37†	.37	—	
8. COH_{8L}	.48*	-.26†	-.22†	-.20†	-.31†	.34	.49*	—
M	174.09	.15	.25	.11	.10	65.41	77.30	81.75
SD	16.78	.09	.12	.08	.23	9.18	8.70	9.20
Possible Range	37-222	0-2	0-2	0-2	0-2	13-91	13-91	13-91
Sample Range	130-194	.04-.40	.05-.46	.04-.41	.00-.35	48-84	51-91	52-91

Note: j = 14-23. $COSE_{PRE:}$ Counseling Self-Estimate Inventory (pre-treatment), GPIRS: Group Psychotherapy Intervention Rating Scale (week 1), GPIRS-STR_1: Group Psychotherapy Intervention Rating Scale – Group Structuring (week 1), GPIRS-$VRBL_1$: Group Psychotherapy Intervention Rating Scale – Verbal Interaction (week 1), GPIRS-$EMTL_1$: Group Psychotherapy Intervention Rating Scale – Emotional Climate (week 1), COH_{1L}: Composite cohesion measure (week 1, leader rating), COH_{4L}: Composite cohesion measure (week 4, leader rating), COH_{8L}: Composite cohesion measure (week 8, leader rating).
*p < . 05. ** p < .01. *** p < .001
†Values represent Spearman-Rho Correlations for the revised GPIRS scale.

Effects of Treatment Condition

Treatment conditions were examined for differences in both leadership behaviors as well as the development of cohesion across time. To examine differences in leader interventions, a Bonferonni-corrected Independent Samples t-test was conducted across the GPIRS full-scale and each of the domains. As noted above, several of the GPIRS items violated the assumption of normality that underlies parametric tests such as t-tests. In such instances, Mann-Whitney *U*-tests serve as a nonparametric test to compare the rank-ordering of values and determine if the distributions of two independent samples are statistically different (Corder & Foreman, 2014). A

Mann-Whitney *U*-test essentially examines if there is a better than chance likelihood of a randomly selected value from the population with the larger mean rank being greater than a randomly selected value from the other. Both t-test and Mann-Whitney *U*-tests were performed and are presented in *Table 6*. Using a Bonferonni-adjusted alpha of .008 (.05/6), the Mann-Whitney *U*-test indicated significant differences for the full scale and all of the subscales, while mean-comparison tests indicated significant differences for the full-scale as well as the Structuring and Verbal Interaction domains, with non-significant differences in the facilitating an Emotional Climate domain. For all of these significant differences, leaders in the process condition were found to perform significantly more behaviors than leaders in the forgiveness condition.

To aid in interpretation, a difference of .10 between the groups on a GPIRS domain equates to *each* of the behaviors within that domain occurring approximately one more time. This is because scores on the GPIRS are averaged across domains across each of the scored intervals. So, for example, a mean difference of .16 on the Structuring domain equates to 2-3 more behaviors for each item scored within this domain, or approximately 8-12 more structuring behaviors in a 90 minute session. Thus, while the mean differences are small, they equate to significant behavioral differences between leaders in the two treatment conditions.

Table 6. Independent Samples t-test and Mann-Whitney U-test of between-treatment differences in Leadership Behaviors

	Process Condition		Forgiveness Condition		Mean			Mann-Whitney U-test
GPIRS Domain	M	SD	M	SD	Diff.	t-test	95% CI	
Full-Scale	.21	.09	.10	.03	.11	3.86*	.05-.17	9.00*
Structuring	.32	.11	.17	.07	.16	3.66*	.07-.24	18.50*
Verbal Int.	.16	.10	.07	.02	.09	2.98*	.03-.16	7.00*
Emotional Cli.	.15	.10	.06	.06	.09	2.67	.02-.17	18.00*

Note. $n = 112$, $j = 21$. * = Bonferonni-adjusted $p < .008$.

To give a better sense of how the behaviors compare between conditions, specifically on variables where distributions were skewed and means may not provide the best estimate, Independent Samples Mann-Whitney U-test results are depicted in *Figure 2*. Here, all values are

ranked from highest to lowest, and the mean ranking of these distributions is compared between groups. Results from this and the independent samples t-test provide preliminary support for the hypothesis that differences existed between the two treatment conditions with respect to leadership behaviors. Specifically, leaders in the process condition engaged in more Structuring and Verbal Interaction behaviors, which resulted in more behaviors overall, while mean differences in emotional climate behaviors were non-significant.

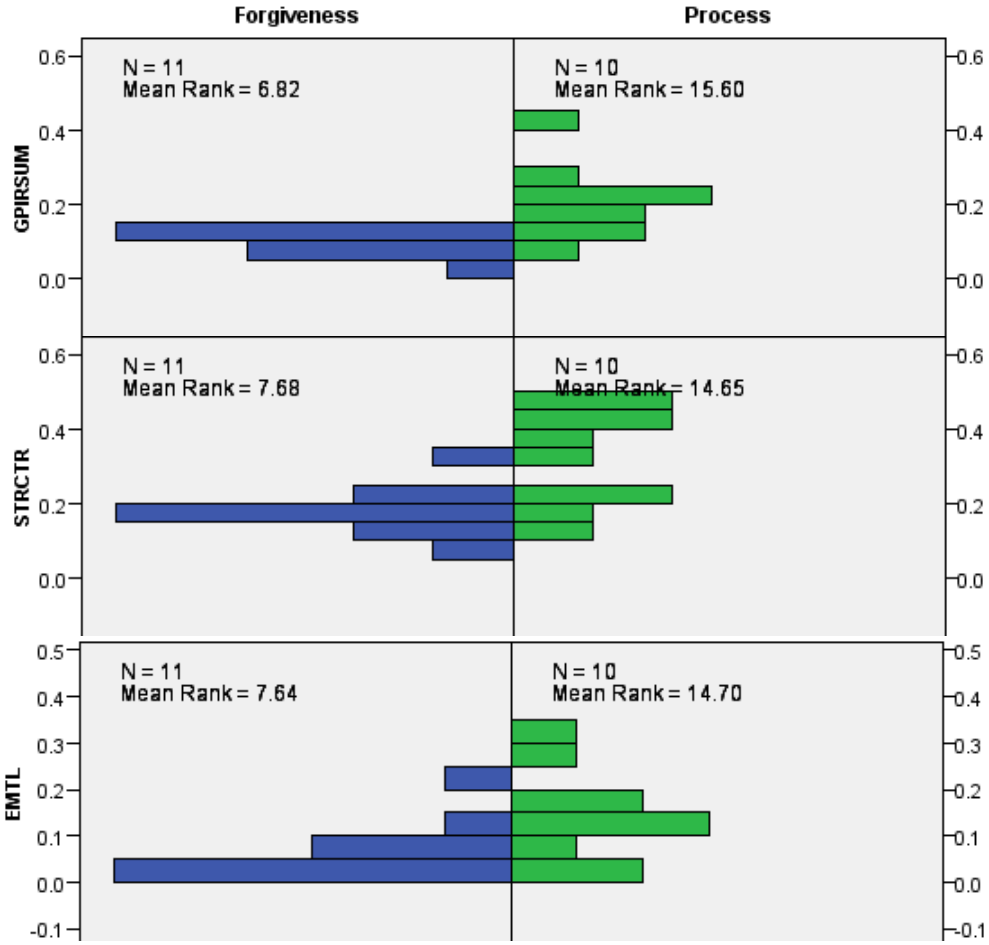


Figure 2. Independent-Samples Mann-Whitney U Test of GPIRS domains by treatment condition

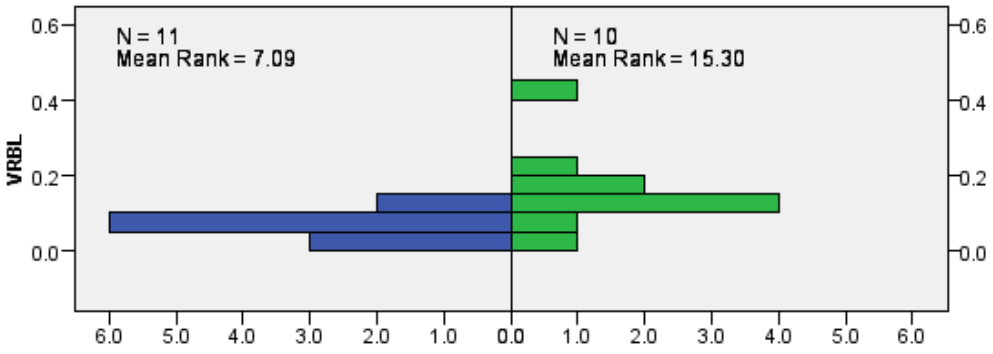


Figure 2 continued

Note. n = 112, j = 21.

To examine the effects of treatment condition on the trajectory of cohesion across time, a dichotomous between-group predictor variable (condition) was created. *Figure 3* plots the cohesion trajectory for the two treatment conditions across time. Results of the growth curve model indicated that responses were increasing on average over time. The condition indicator did not significantly influence the cohesion intercept or slope (β = -.22; p = .42 and β = .24; p = .52, respectively), suggesting that the development of cohesion was not significantly influenced by treatment condition. As such, treatment condition was collapsed for further analysis, with significant differences between conditions thought to be captured by differences in leadership behaviors as discussed above.

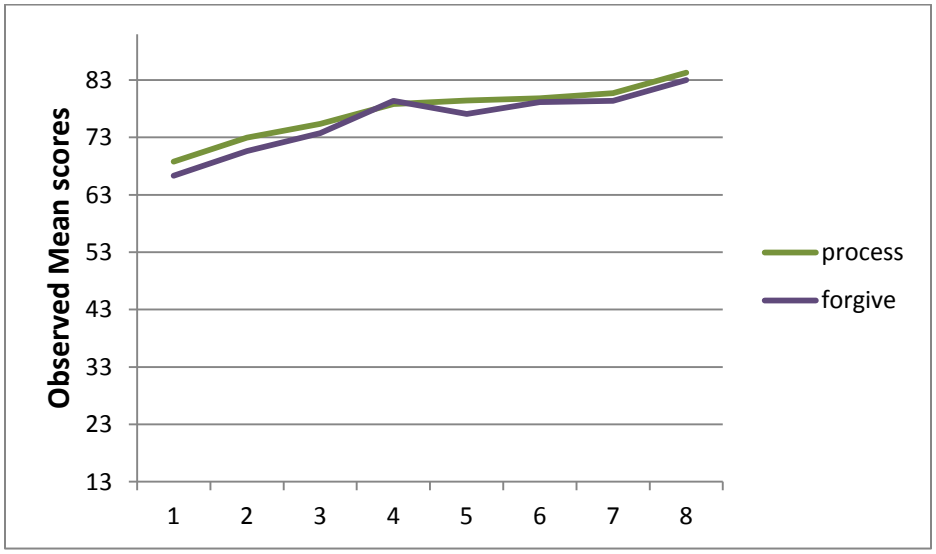


Figure 3. Cohesion Trajectories for the Two Treatment Conditions

Note. n = 128, j = 23.

Examining Assumptions

Assumptions of multilevel regression analyses such as latent growth curve analysis (LGC) are similar to those of Ordinary Least Squares regression—with a few additions. First, LGC requires an adequate sample size and number of repeated measures in order to have sufficient power to detect effects. Typical minimum requirements are a sample size of at least 100 and at least three repeated measures per individual (Curran, Obeidat, & Losardo, 2010). While both of these basic criteria were met in the present study, power can still be impacted by such factors as amount of missing data and the effect size of the predictor-outcome relationship. In order to more fully examine power issues in the present study, a power analysis was conducted and is presented in the *Limitations* section below.

Second, the typical method of estimation, maximum likelihood (ML), requires that the repeated measures are continuous and normally distributed, although this same property is not required of predictor variables. To examine this, the distribution of the estimated latent linear cohesion variable across individuals and groups was examined through MPLUS. Visual inspection of these histograms revealed that individual repeated measures were continuous and normally distributed. Inspection of the group-level (level two) histogram, however, revealed that repeated measures assumed a bi-modal distribution. This suggested that while the repeated measures were normally-distributed at the individual level, between group differences existed when a linear relationship was assumed. This suggested the possibility of a non-linear relationship between time and cohesion, which was examined as a function of the study hypotheses.

Next, level-one residuals must be constant at level two and must be uncorrelated with residuals at level two. These criteria were both satisfied by the modelling program used,

MPLUS, which sets level-one residuals to be constant at level two and allows residuals to be estimated freely between levels. In addition, multicollinearity (or highly correlated variables) can introduce error, attenuate effects, and, if high enough, lead to poor estimates of unstandardized and standardized regression coefficients. This was of particular concern in the leadership behaviors measure, which included three separate measures that were found to be highly correlated (r = .54 – .86). To examine potential problems with multicollinearity in this measure, collinearity diagnostics were examined. Variance Inflation Factors (VIF) were between 1.83 and 3.45; below the typical cutoff value of 5 considered to be cause for concern (Menard, 1995). Additionally, no condition index was above 30 (ranging from 1.00 – 3.5), consistent with the data screening procedures recommended by Tabachnick and Fidell (2001). As such, multicollinearity was not hypothesized to be a problem in the present investigation.

Finally, observations at the highest level (in this case the group level) must be independent of each other (Snijders & Bosker, 1999). While observations at the highest level (group) were necessarily independent by their quality of each individual being in only one group, there was the possibility of violating this assumption by virtue of the fact that the same leader led multiple groups. To examine this, ICC values between therapists in the outcome variable (cohesion) were examined. ICCs across the eight-weeks ranged between .009 to .058; far below the frequently cited cut-off of .10 that indicates non-trivial amounts of intragroup dependency (Johnson et al., 2006). This indicated that leader was likely a nuisance variable that did not need to be accounted for in the final model, lending support for the present investigation meeting assumptions of independence at the group level.

Testing Study Hypotheses

What is the agreement between member and leader ratings of cohesion? First, in order to determine empirically whether member ratings of cohesion could be viewed as a group-level variable, interrater agreement between members of the same group was examined. This was done by estimating ICC(1) and ICC(2) values. ICC(1) reflects the amount of dependency between ratings by members of the same group, and can be used as an estimate of how well a single member's rating estimates the average rating of an outcome within a group (Bleise, 1998). One difficulty with ICC(1), however, is that its magnitude is largely unaffected by group size (Bleise, 1998). As such, ICC(2) is often calculated in addition to capitalize on the fact that with increased group size comes an increase in the reliability of the estimated group means (Gold, Kivlighan, & Patton, 2013).

ICC(2) can be used to determine whether responses should be aggregated at the group level, as it serves as an estimate of the magnitude of dependence among members' session-level responses. Baumgartner, Jackson, Mahar, and Rowe (2003) have developed a decision rule for determining whether responses should be aggregated, and has been used by group therapy researchers in the past (e.g. Gold, Kivlighan, & Patton, 2013). Baumgartner and colleagues (2003) suggest a decision rule that ICC(2) values should exceed a minimum of .70, where values between .70 and .79 are considered below-average acceptable, values between .80 and .89 are considered average acceptable, and values between .90 and 1.00 are considered above-average acceptable.

ICC(1) and ICC(2) values are presented in *Table 7* below. In order to account for increasing missingness across time due to group drop-out and study attrition, Estimation Maximization (EM) was used to replace missing values. EM, a maximum-likelihood (ML)

procedure, uses a two-step iterative procedure in which missing values are imputed and a covariance matrix is estimated and used to conduct a series of regression equations to estimate missing values (Enders, 2003). EM has been found through Monte Carlo simulation to consistently yield accurate reliability estimates (Enders, 2003). Given the previous finding that pre-test levels of psychological distress significantly predicted drop-out, this was used along with early levels of cohesion and group membership in the EM procedure.

Results indicated that while group member agreement was acceptable at week one (ICC(2) = .70); all subsequent sessions demonstrated lower estimates of member agreement. In particular, member agreement within groups became very low toward the end of group (weeks seven and eight; ICC(2) = .42). As such, it was determined that client-rated cohesion should not be aggregated at the group level. As a result, a multiple-indicator model, where clients and leaders serve as indicators of an underlying cohesion construct, requires that the indicators be at the same level (in this case at the group level). Because member ratings were found to be divergent enough within groups to avoid aggregation to the group level, a multiple-indicator growth model was considered inappropriate for the present data.

Table 7. ICC(1) and ICC(2) Values for Cohesion Across Eight Weeks

Week	1	2	3	4	5	6	7	8
ICC(1)	.30	.15	.18	.23	.25	.21	.11	.11
ICC(2)	.70	.50	.55	.62	.64	.60	.42	.42

Note. k = 5.565; j = 23. N = 128.

As an alternative to a multiple-indicator model, a latent variable modeling framework was used. First, in order to examine session-level agreement, a two-level regression analysis was conducted. Here, agreement between client ratings was modelled at the group level (i.e. not simply aggregated ratings but latent, fixed effects accounting for residual errors due to differences in member ratings) and leader ratings for each group. This procedure fixes all of the

means, variances, and covariances to the observed sample values while freely estimating the covariance between client and leader ratings. This provided an estimate of the correlation between client and leader ratings and a test of this value against zero. Latent means for members, leaders, and Pearson correlations (*r)* across each of the eight weeks of treatment are presented in *Table 8.*

Table 8. Member-Leader Session-Level Agreement Under a Latent Variable Framework

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
Member *M*	67.41	71.22	75.21	79.45	78.75	79.68	80.49	83.87
Leader *M*	65.38	69.96	73.96	77.30	77.10	78.60	78.78	81.75
Client-Leader *r*	.59***	.58***	.83***	.61***	.48***	.45***	.59***	.38***

Note. $k = 5.565$; $j = 23$. $N = 128$. *= $p < .05$. ** = $p < .01$. *** = $p < .001$. Both Maximum Likelihood (ML) and Maximum Likelihood with standard errors approximated by First-order derivatives (MLF) were computed. MLF was indicated due to saddle-point problem (a point in the function between two variables where both partial derivatives are zero), but both MLF and ML estimation resulted in the same estimates (reported above).

Notably, there was greatest agreement between leaders and clients after the third week of treatment. Agreement decreased steadily through week 6, and increased again after week 7 before reaching its lowest point at the final week of treatment. Still, there was a strong overall degree of agreement between clients and leaders, with 14%-69% of variance shared by the two raters (coefficient of determination (r^2) = .14 - .69). This was also observed in the similar trajectory of both patients and therapists across the group sessions, as depicted in *Figure 4.*

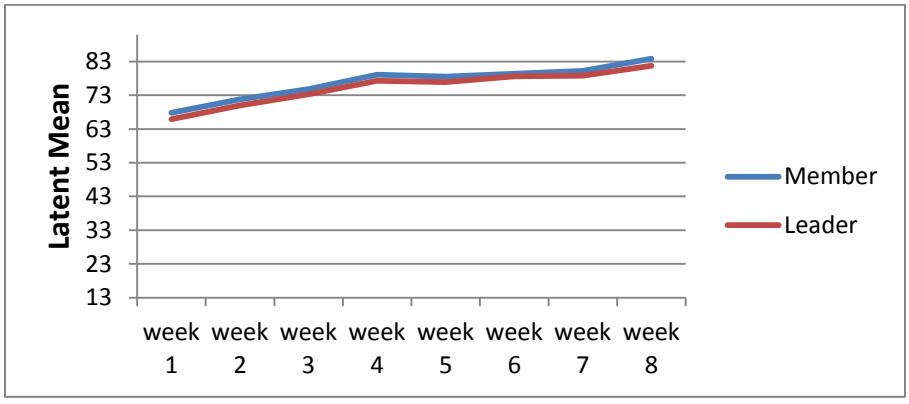


Figure 4. Estimated Group Means for Member and Leader Ratings of Cohesion Across Time
Note. $n = 128$, $j = 23$.

How does cohesion change across time? The first step in this hypothesis was to examine whether group membership explained significant variance in the outcome variable (cohesion) and thus whether a multilevel model was necessary. To do so, an Intraclass Correlation Coefficient (ICC) was calculated for the unconditional growth curve model. Here, group membership was included as a clustering variable in a two-level growth model, with level-one representing change in cohesion and level-two representing group membership alone. This resulted in the calculation of an ICC at each time-point while accounting for growth model parameters (Duncan, Duncan, & Strycker, 2006).

ICC results for each of the eight time points are presented in *Table* 7. In multilevel modelling, ICC values of .10 or more indicate non-trivial amounts of intragroup dependency (Johnson et al., 2006), while values of .25 and above indicate high levels of intragroup dependency (Heinrich & Lynn, 2001; Guo, 2005). ICC results in the present study ranged from .13 to .31, with an average ICC of .21. In other words, the average amount of overall variance attributable to between group variance was 21%. These results indicated non-trivial amounts of intragroup dependency at all time-points, with the highest agreement between members of the same group occurring at session one, session five, and session eight. This provided strong evidence that group membership impacted member-ratings of cohesion, and thus that multilevel modeling was necessary.

Table 9. Intraclass Correlation Coefficients (ICCs) for Cohesion Modelled Across Time

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
ICC	.31	.23	.13	.16	.27	.13	.19	.24

Note. n = 128, j = 23.

The second part of this hypothesis involved examining the shape of overall cohesion change across time. Data were examined through a hierarchical linear model to determine whether a linear, quadratic, cubic, or piecewise model best fit the data. To determine which of

these to examine initially, individual cohesion change along with ordinary least squares (OLS) growth plots were examined (Tasca et al., 2009; Singer & Willett, 2003). Simple individual change plots are presented in *Figure 5* below.

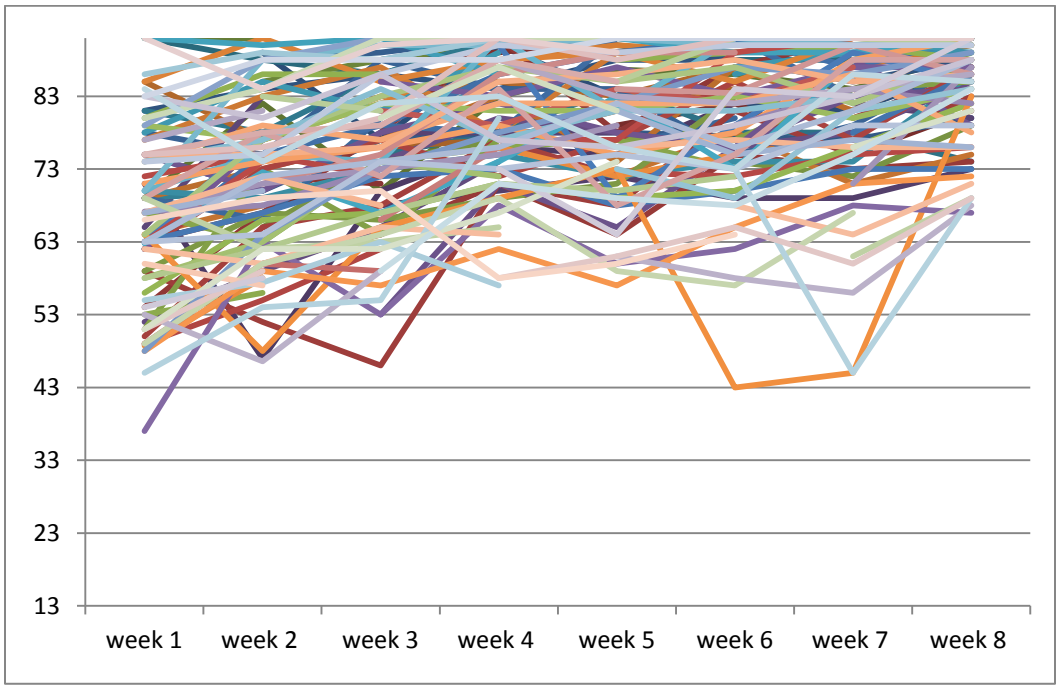


Figure 5. Trajectories of Client-rated Cohesion Across Time
Note. n = 128, j = 23.

Looking at individual trajectories of cohesion across time, one can observe a generally high level of cohesion, even at time one. On a scale that ranged from 13-91, average cohesion at time one was 66.77 with a standard deviation of 12.13. This indicated that, on a qualitative scale where 1= "not at all", 4 = "moderately", and 7 = "extensively", groups began at an average of 5.13, with 95% of the population falling between "moderate" and "extensive" levels of cohesion (3.26 to 7.00). Examining the shape of the cohesion trajectories across time indicated a general, positive linear trend. OLS regression Q-Q plots examining time by cohesion, however, indicated that a linear trajectory underestimated cohesion at early stages and overestimated cohesion at later stages, indicating the possibility of a quadratic, cubic, or piece-wise growth model.

To examine these various growth possibilities, linear, quadratic, cubic, and piecewise growth models were examined and compared for fit. Recommendations for model fit have been reported as a CFI/TLI of above .95, an SRMR of below .08; and an RMSEA of below .06 (Hu & Bentler, 1999). To determine the best fitting unconditional model, chi-square difference testing under a multilevel modeling framework using a weighted likelihood function was used. Here, parameters are estimated using a weighting function to account for the non-independence of observations (Muthén & Muthén, 1998-2012).

All single slope models (linear, quadratic, and cubic) resulted in a generally poor fit to the data; Linear $\chi^2_{31} = 98.63$, $p < .0001$; RMSEA = .13; CFI = .82; TLI = .84; SRMR = .31; Quadratic $\chi^2_{27} = 87.08$, $p < .0001$; RMSEA = .13; CFI = .84; TLI = .84; SRMR = .28; Cubic $\chi^2_{22}=63.05$, $p < .001$; RMSEA = .12; CFI = .89; TLI = .86; SRMR = .22. Of note, the cubic model resulted in a non-positive variable covariance matrix, resulting in negative variances for the quadratic and cubic terms. This indicated poor model convergence and an inadmissible solution. The estimated group means for these models is presented in *Figure 6.*

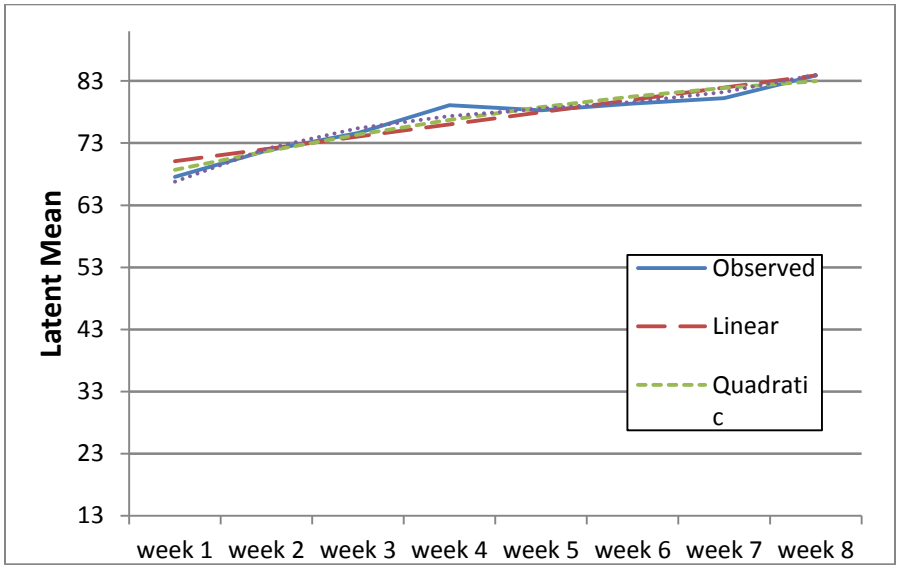


Figure 6. Observed Group Means and Linear, Quadratic, and Cubic Model Group Estimates
Note. $n = 128$, $j = 23$.

To further examine the best-fitting models, modifications to the linear model were examined along with piecewise growth models. First, to improve fit of the linear model, modification indices and sample and estimated group mean plots were used to hypothesize points of non-linearity of change. There is general consensus in the structural equation modelling literature that modification indices can be used to improve model fit, but should be used with caution. In particular, Kline (2011) warns that any modifications based on such indices should be (a) theoretically-justifiable, (b) few in number, and (c) minor, meaning they should not largely impact other parameters in the model. Using these guidelines, there was theoretical justification for reducing the requirement of strict linearity. In particular was the large number of time-points being modelled, which increased power to detect any misspecification from perfect linearity even though a linear trend may have existed.

Examination of modification indices for the linear model indicated specific problems with time 1 and time 4. These were allowed to be freely estimated by the model rather than fixed at a specified time value. Allowing week 1 and week 4 to be freely estimated resulted in moderate fit indices ($\chi^2_{29} = 59.09$, $p < .001$; RMSEA = .09; CFI = .92; TLI = .92; SRMR = .29) and significant improvement in fit over the quadratic and linear models ($\Delta\chi^2_2 = 39.54$, $p < .001$, $\Delta\chi^2_2 = 27.99$, $p < .001$; respectively).

Piecewise growth models are those in which segmented growth results in two or more growth trajectories across time (see Kohli et al., 2015 for a review). To examine the fit of a piecewise growth model, the observed group means were used to hypothesize two separate trajectories, with a significant split (or 'knot', Kohli et al., 2015, p. 260) between weeks four and five. This was hypothesized based on the possibility of early and late-phase cohesion development, as discussed by several authors in the group therapy literature (see Appendix A).

To examine a piecewise trajectory, two slope growth factors with a single growth intercept were specified. Consistent with the observed group means, a linear-linear piecewise and linear-quadratic piecewise model was specified. The fit-indices of the linear-linear piecewise model indicated a moderate fit to the data; χ^2_{27}= 74.96, $p < .001$; RMSEA = .12; CFI = .88; TLI = .87; SRMR = .29. The linear-quadratic piecewise model indicated acceptable model fit estimates (χ^2_{22} = 44.06, p = .004; RMSEA = .09; CFI = .94; TLI = .93; SRMR = .14) and improved fit over the linear-linear model ($\Delta\chi^2_5$ = 30.9, p = .024). The linear-quadratic model was then compared to the best-fitting linear slope model (with freely estimated time 1 and time 4 values). Chi-square difference testing indicated that the linear-quadratic piecewise model significantly improved upon the fit of the modified linear model ($\Delta\chi^2_7$ = 15.03, p = .035)[1]. The group-mean estimates for the modified linear, piecewise linear-linear, and piecewise linear-quadratic models are depicted in *Figure 7*.

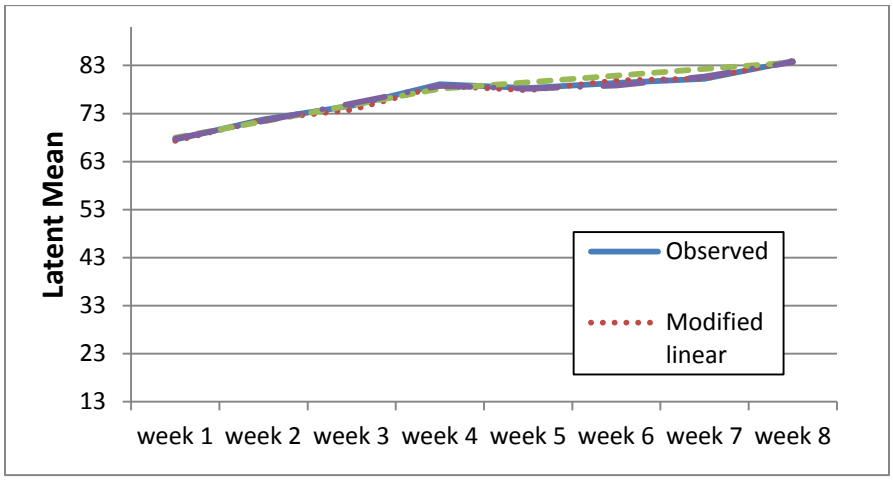


Figure 7. Observed group means and the piece-wise linear-linear and piecewise linear-quadratic model estimates

Note. n = 128, j = 23.

[1] Whether the piecewise models could be considered a "nested" model with respect to the single slope models is not immediately clear. While the piecewise models are certainly more restricted with respect to the number of free parameters, covariance-matrix and moment-matrix nesting is much more difficult to determine. While much has been written on this topic (e.g. Bentler & Satorra, 2010), the magnitude of the difference in fit between the simple quadratic and linear-quadratic models was moderate (Cohen's w = .13), supporting the use of the linear-quadratic model over the simple quadratic slope model.

To examine variability in growth parameters attributable to individuals and group-membership, a two-level HLM was conducted, specifying individual and group-level weighting (Muthén & Muthén, 1998-2010). This resulted in increased parameter specification due to estimating both group and individual levels ($df = 40$), resulting in more free parameters than number of clusters ($j = 23$). The two-level specification thereby demonstrated a non-positive definite first-order derivative product matrix, a result of a negative residual variance at time 8 at the individual level and time 2 and 7 at the between-group level. This was hypothesized to result from the small amount of variability in these time scores (see Figure 8 above). The negative residual variance indicated a possible overfit to the data. "Overfit" occurs when a model contains more parameters than data and may excessively exaggerate minor fluctuations across time (Muthén & Muthén, 1998-2010). In such instances, parameter estimates are considered interpretable due to using a maximizing weighted loglikelihood function, while interpretation of standard errors and resulting significance tests are suspect. Thus, while the results of this model are reported, standard errors and p-values should be interpreted with caution.

Parameter estimates reflecting the best-fitting linear-quadratic piecewise growth model are reported in *Table 10*, consistent with the recommendations of Duncan, Duncan, & Strycker (2006). To aid in interpretation, second-order coefficients represent a group-specific growth factor, to which is added the first-order individual parameter in the prediction of a person's overall change across time. Thus, first-order coefficients represent between-person change across a given time series, while the second-order coefficients capture the variation between groups in that same term.

Table 10. Parameter Estimates for the piecewise linear-quadratic LGC

	Estimate	
Coefficient	Second-Order	First-Order
Means		
Cohesion Intercept	67.75***	
Cohesion Slope$_{1\text{-}4}$	3.72***	
Cohesion Slope$_{4\text{-}8}$	-1.161	
Cohesion Quadratic$_{4\text{-}8}$	.60***	
Variances		
Cohesion Intercept	25.72	88.62***
Cohesion Slope$_{1\text{-}4}$	.60	2.40*
Cohesion Slope$_{4\text{-}8}$	.38	2.41
Cohesion Quadratic$_{4\text{-}8}$	.01	.35

Note. $n = 128$, $j = 23$. *$p < .05$. ** $p < .01$. *** $p < .001$. Results are in raw scale units.

Parameter estimates indicated significant second-order mean levels of the intercept Mi = 67.81, $p < .001$, session 1-4 cohesion slope $Ms_{1\text{-}4} = 3.64$, $p < .001$, and a non-significant late-stage linear slope $Ms_{4\text{-}8} = -.367$, $p = .695$ mitigated by a significant quadratic trend $Mq_{4\text{-}8} = .70$, $p = .024$. This indicated that initial average levels of cohesion differed significantly from zero, as did early linear change in cohesion (3.72, $p < .001$) and session 4-8 quadratic change (.60, $p < .001$).

Given the significant session 1-4 and session 4-8 slope terms, an effect size of cohesion change across time could be calculated. Average change in cohesion across the first four sessions was 3.64 points on the cohesion scale per session, the average change under the last four sessions (quadratic slope) was 1.24 across each of the four sessions. Using standard deviations from the mean of initial session cohesion ($SD = 12.21$), an estimated effect size of ($d = .20$) can be calculated. This indicated that, on average, client-rated cohesion increased by one fifth of a standard deviation for each session of group counseling. This suggests an overall increase in cohesion of approximately 1.59 SD across the lifespan of these psychotherapy groups.

Individual (first order) variability around the average intercept and early phase slope parameters was also significant ($Di = 88.62$, $p < .001$, and $Ds_{1\text{-}4} = 2.40$, $p = .038$) while non-significant individual differences existed in the late phase slope estimates ($Ds_{4\text{-}8} = 2.41$, $p = .61$

and $Dq_{4\text{-}8}$ = .35, p = .23). At the second-order (between) level, variability around all average intercept and slope parameters was not significantly different from 0 (Di = 25.72, p = .254, and $Ds_{1\text{-}4}$ = .60, p = .568; $Ds_{4\text{-}8}$ = .38, p = .904; $Dq_{4\text{-}8}$ = .005, p = .98).

Taken together, the above findings indicated that while significant change was observed at each time point, significant individual variability existed only around initial levels and early stage change, and average cohesion did not significantly differ from one group to another at initial cohesion nor growth in cohesion across time. This is further supported by visual inspection of the latent means and standard deviations for the between and individual level cohesion scores, as depicted in *Figures 8a and 8b*. Here, noticeably more variability existed at the individual level than at the group level. Also noticeable is a trend toward decreasing variability in average group-level cohesion across time.

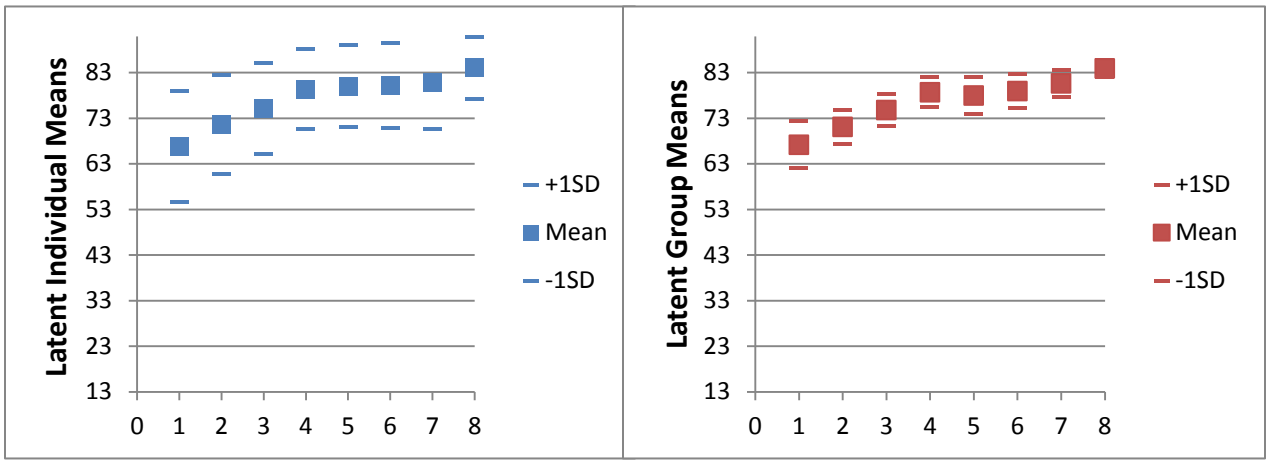


Figure 8a & 8b. Estimated Individual and Group Means for the Cohesion Time-scores
Note. n = 128, j = 23.

Results from the LGC with prediction paths between growth parameters indicated a strong negative association between initial individual levels of cohesion and session 1-4 slope (β = -.69, p < .001), and negative but non-significant association with session 4-8 quadratic growth (β = -.38, p = .096), with a positive but non-significant association with session 4-8 linear growth (β = .27, p = .55). Beyond this, a strong negative association was found between session 4-8 quadratic growth and

session 4-8 linear growth (β = -.91, $p < .001$). These results are depicted in *Figure 9* below. At the between-group level, non-significant associations were found between all growth parameters. Taken together, this indicated that individuals who rated cohesion as higher initially saw a slower rate of change then individuals rating initial cohesion as lower, but that there was little effect at the group level of initially rating cohesion as high or low on later changes across time. The significant association at the individual level may reflect a ceiling effect, such that those who rated initial cohesion as very high had less room to grow across the period of the study. Average initial levels of cohesion were 67.56 on a scale of 1-91, meaning that individuals were on average rating groups at a value of 4.48 on a 6 point scale. Even initial levels of cohesion were thus above "moderate", with less room for growth overall. As such, predicting initial session cohesion represented an important component of later cohesion growth, with first session cohesion contributing greatly to the ultimate cohesion trajectory across time.

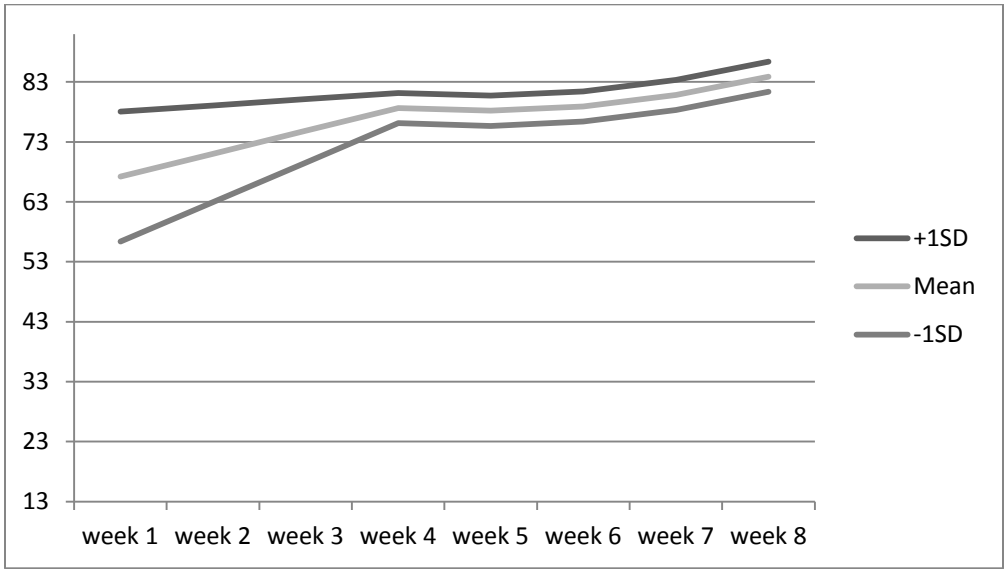


Figure 9. Individual Rates of Change in Cohesion as a Function of Initial Level
Note. $n = 128$, $j = 23$.

To examine the contributions of individual versus group membership to variability in scores across time, second-order-to-total factor variances can be used (Duncan, Duncan, & Strycker, 2006). For the intercept, the second and first order factor variances were estimated as

25.72 and 88.62, respectively, with a second-order-to-total factor variance ratio of .22. For the session 1-4 slope, the factor variance ratio was .20, for the late-stage slope, the factor ratio was .14, and for the late-stage quadratic trend the factor ratio was .03. Taken together, this indicated that 22% of the total variation in intercept scores, 20% of the total variation in early-stage slope, 14% of the late-stage slope, and 2.7% of the late-stage quadratic trend was accounted for by group membership.

Do client, leader, and group characteristics predict changes in cohesion? The statistically significant variance in first-order intercept and slope terms from the above LGC allowed for an examination of individual characteristics as potential predictors of both initial levels of cohesion and changes over time. Non-significant variance at the between-level, however, prevented an examination of predictors of variability in group-level intercept and slope cohesion terms. As such, two models were examined. First, the piecewise linear-quadratic model was examined for individual-level predictors of initial cohesion and cohesion change. Second, the simpler, linear model (with time points 1 and 4 freely estimated) was used to examine group-level contributors to growth, controlling for level-one covariates.

To examine individual-level predictors of initial cohesion and changes in cohesion across time, the best-fitting piecewise linear-quadratic model was examined under a complex LGC model. Consistent with the study hypotheses, individual (first-order) variables included psychological distress, self-esteem, and attachment style (anxious and avoidant), while demographic characteristics of age, gender, and race were controlled for. All model results are reported as beta-weights (β), or the predicted difference in the outcome variable in standard units for a one standard deviation increase on the given predictor variable holding all other predictors constant. Unstandardized regression coefficients were used for null-hypothesis testing and

reporting of statistical significance, consistent with the recommendations of Muthén & Muthén, (1998-2010). In addition, parameter estimates indicating the percentage of variance accounted for by the variables (r^2) are presented.

To simplify interpretation of race and gender these variables were dummy coded. 121 persons identified as "white" or "Caucasian", while 5 identified as a person of color (two did not respond). Thus, race was dummy coded to reflect white (POC=0) or person of color (POC=1). All participants in the present study identified as male (n=26) or female (n=102). Accordingly, gender was dummy coded to reflect either female-identified (FEMALE=1) or male-identified (FEMALE=0). Using these variables, gender was found to significantly relate to first session cohesion, with women being more likely to rate first session cohesion higher than men (β = .17, p = .041), while gender did not significantly relate to changes in cohesion across time. Being a person of color was found to have a moderate, negative relationship with session 4-8 quadratic cohesion change (β = -.39, p = .009) but no other parameters. Age was not found to significantly relate to initial cohesion or cohesion change at any time point.

Controlling for the effects of gender, race, and age, only avoidant attachment significantly predicted initial level of cohesion, with a moderate negative relationship (β = -.26, p = .02). Avoidant attachment accounted for about 6.7% of the total individual variance in cohesion scores. In other words, for each standard deviation increase in avoidant attachment initial levels of cohesion dropped by 2.44 points on a 78-point scale. Anxious attachment was the only variable found to correlate with cohesion slope in sessions 1-4. This relationship was moderate and positive (β = .35, p = .02), and accounted for about 12.2% of individual variance. Phrased differently, for every one standard deviation increase in anxious attachment the linear slope in cohesion increased by .54 points per session.

As a set, anxious attachment, avoidant attachment, psychological distress, self-esteem, and the included demographic variables accounted for very little of the individual variability in initial cohesion and session 1-4 slope, about 11% and 15%, respectively. While these rates were greater in later sessions (68% of session 4-8 linear slope and 40% of session 4-8 quadratic slope), the variability in session 4-8 growth parameters was not significantly different from 0 and thus these terms were non-significant. The results of this conditional LGC analysis are reported in *Table 11* below.

Table 11. Individual-level covariate effects on complex LGC model parameters

Coefficient	Estimate (β)			
First-Order Covariate	Intercept	Linear Slope$_{1\text{-}4}$	Linear slope$_{4\text{-}8}$	Quadratic Slope$_{4\text{-}8}$
Anxious Attachment	-.18	.35*	-.62	.38
Avoidant Attachment	-.26*	.10	.39	-.16
Initial distress	.06	-.20	.27	-.14
Self-esteem	-.04	.01	.19	-.11
Total r^2	.11	.15	.68	.40
Total explained r^2	.09	.12	.58	.39

Note. $n = 128$, $j = 23$. *$p < .05$. ** $p < .01$. *** $p < .001$. Total r^2 represents the total explained variance of the variance attributable to individual differences. So, for example, with a total r^2 value of .11 and .78 of the variance attributable to individuals, the total explained variance of the included individual-level predictors was .11*.78, or approximately .09. While Beta-weights are reported, null-hypothesis testing was conducted on the unstandardized data.

As noted above, several limitations prevented an analysis of group-level covariates under the piecewise linear-quadratic slope, primary of which was the number of parameters specified in such a model (*df*=40) and an insufficient number of groups (*j*=23). In addition, because of the non-significant variance at the between-level intercept and slope terms in this model, all group-level predictors would be rendered uninterpretable. Although representing a poorer-fit to the data, the modified linear model did provide modest model fit parameters ($\chi^2_{29} = 59.09$, $p < .001$; RMSEA = .09; CFI = .92; TLI = .92; SRMR = .29) and was significantly simpler (*df* = 29). This afforded enough free parameters to examine group-level covariates and still capture the general,

positive linear trend that occurred across sessions 1-8. This model was thus used for an examination of group-level predictors.

For reference, parameter estimates of the unconditional linear model indicated significant second-order mean levels of the intercept Mi = 70.43, $p < .001$, and cohesion slope $Ms_{1\text{-}8} = 1.86$, $p < .001$. This indicated that initial average levels of cohesion and average change in cohesion differed significantly from zero. Individual (first order) variability around the average intercept and slope parameters was also significant ($Di = 91.65$, $p < .001$, and $Ds_{1\text{-}8} = .99$, $p = .009$). At the second-order (between) level, variability in average intercept scores was significant ($Di = 25.84$, $p = .032$) while variability in slope scores was not significantly different from 0 ($Ds_{1\text{-}8} = .23$, $p = .326$). Results of the level-one covariate predictors were consistent with those of the quadratic model, with only avoidant attachment predicting intercept cohesion scores ($\beta = -.31$, $p = .014$). No significant predictors of the cohesion slope across the eight sessions were found.

Results of the unconditional linear LGC indicated that it would be possible to examine group-level predictors of initial cohesion given the significant variance in this term. Non-significant variance in the slope term at the group-level, however, indicated that hypotheses could not be tested around contributors to cohesion change across time. Still, because of the strong relationship between intercept and slope scores, as well as the variability in intercept scores, hypotheses regarding predictors of initial cohesion were considered important. The second-order-to-total factor variance ratio was .21 for the intercept score and .19 for the slope score, indicating that 21% of the variability in intercept and 19% of the variability in slope terms was attributable to group membership.

The two-level LGC analysis included the demographic covariates specified in level-one (age, gender, race) as well as the level-two predictors. Level two predictors included the number

of members in the group, the average of number of sessions attended in the group, and the three categories of leadership behavior (Structuring, Verbal Interaction, and Emotional Climate). Leadership behaviors were standardized prior to conducting the analyses to address issues of scaling.

Three iterations of the two-level model were examined, consistent with the recommendations of Howell & Lacroix (2012) for the examination of interaction effects. First, in accordance with study hypotheses, all of the predictors, all possible two-way interaction terms (structuring * verbal interaction; structuring * emotional climate, and verbal interaction * emotional climate), and a three-way interaction term were included. Next, the three-way interaction term was fixed to 0 to examine significant two-way interactions. Finally, the three-way and two-way interaction terms were fixed to 0 to examine possible main effects. Because two of the

Modelling using the above procedure resulted in three nested models, which could be compared using Log-Likelihood testing. Non-significant improvements are evidence for the parsimony principle, such that even though an interaction term may be significant it may not drastically improve prediction of the latent cohesion variable. In order to control for the influence of demographic variables—age, gender, and race were included in all iterations of the model. In addition, although the slope term is not depicted below given the non-significant group-level variance in slope, a slope term was included in all models for model testing. Results of these analyses are presented in *Table 12* below.

Table 12. Group-level Predictors of the Latent Cohesion Intercept across Two-level Linear LGC models

	Model			
Second-Order Covariate	Three-way Interaction (β)	Two-way Interaction (β)	Minimally Adequate (β)	Main Effects (β)
Group Attendance	-.06	.06	.11	-.12
Group Membership	-.25	-.27	-.26	-.23
Counselor Self-Efficacy	-.01	-.13		-.05
Leadership Behaviors				
Structuring	-.63	-.53	-.20	-.54
Verbal Interaction	2.22*	1.28*	.46	.71
Emotional Climate	.05	-.49	-.38	.12
Structuring*Emotional	.36	.92*	.65**	
Verbal*Emotional	.25	-1.71		
Structuring*Verbal	.49	.87		
Structuring*Verbal*Emotional	-2.63			
Total r^2	.88	.70	.44	.33
Total explained r^2	.18	.14	.09	.05

Note. Note. $n = 112$, $j = 21$. *$p < .05$. ** $p < .01$. *** $p < .001$. Total r^2 represents the total explained variance of the variance attributable to group membership. So, for example, in the case of the minimally adequate model, a total r^2 value of .44 with .21 of the variance in intercept scores attributable to group membership, the total explained variance of the included group-level predictors was .44*.21, or approximately .09.

Under the three-way interaction model, nearly all predictors, including the three-way interaction term, were non-significant. The only predictor found to significantly differ from 0 in this model was the Verbal Interaction category ($\beta = 2.22$, $p = .025$)[2]. Fixing the three-way interaction term to 0 resulted in a significant two-way interaction between Structuring Behaviors and Emotional Climate behaviors ($\beta = .92$, $p = .035$), as well as a significant main-effect of Verbal Interaction ($\beta = 1.28$, $p = .031$). In the *minimally-adequate* model, described below, only the interaction between Structuring Behaviors and Emotional Climate was significant ($\beta = .65$, p

[2] Because beta-weights are the relationship between predictor and outcome holding all other predictors constant, it is possible for this value to be greater than 1. Such instances may result when there is a variable in the model which has little relationship to the outcome but a high relationship to other variables. This was thought to be the case for the emotional climate domain, which showed non-significant relationships to cohesion but was highly correlated with the other leadership behaviors in the model ($r = .70$ and .54 for structuring and verbal behaviors, respectively). Beta-weights greater than one can also result from multi-collinearity between predictors. However, as discussed in the assumptions section above, collinearity diagnostics were non-problematic for the GPIRS measure.

= .007). In the final, main-effects only model, none of the predictors were found to differ significantly from 0.

Log-Likelihood testing (*-2LL*) comparing the fit of these models was then conducted. *-2LL* is defined as -2*(log-likelihood$_0$ - log-likelihood$_1$) where the degrees of freedom for the distribution test equal to the difference in parameters between the nested models, or $p_0 - p_1$. Because of the estimation method used in the LGC (MLR), a correction factor (*cd*) must be applied, which is defined as $(p_0 * c_0 - p_1 * c_1)/(p_0 - p_1)$ where c_0 and c_1 are the H0 scaling correction factors. The scaling correction factor is applied by the formula *-2LL/cd* with df of $p_0 - p_1$.

The log-likelihood for the three-way interaction model was -2153.53 with an H0 scaling correction factor of 1.123 and 54 parameters. The two-way interaction model was Log-Likelihood was -2155.077 with an H0 scaling correction factor of 1.184 and 52 parameters. The main effects only model had a log-likelihood of -2160.00 with an H0 scaling correction factor of 1.153 and 46 parameters. Comparing the three-way interaction model to the two-way interaction model resulted in a negative corrected chi-square test statistic (Uncorrected $\Delta\chi^2_2 = 3.08$, p = .214), indicating no improvement in fit through addition of the three-way interaction. Comparing the two-way interaction model to the main effects only model resulted in a non-significant difference as well ($\Delta\chi^2_6 = 6.94$, p = .32), supporting the parsimony of the main-effects only model.

To examine the *minimally-adequate* model (Bingham & Fry, 2010), only the significant two-way interaction term was estimated while all other two-way interaction effects were fixed at 0. The log-likelihood value for this model was -2157.714, with an H0 scaling correction factor of 1.156 and 47 degrees of freedom. Comparing this to the main effects model also resulted in a

near but non-significant improvement in fit ($\Delta\chi^2_1 = 3.55$, $p = .059$). Given all of this, the interaction effect presented below should be interpreted as preliminary evidence for such an effect. This is especially true given that the current investigation had limited power to detect effects at the group level, and the standard errors of estimates may have been inflated.

The non-significant main effects in the two-level model of both emotional climate and structuring behaviors indicated that, on average, neither facilitating an emotional climate nor engaging in structuring behaviors significantly predicted cohesion scores. A significant interaction between these two terms, however, suggested that the effect of structuring behaviors on cohesion differed depending on how many emotional climate behaviors the leader engaged in, and indicated that structuring behaviors may be importantly related to cohesion at varying levels of emotional climate behaviors.

To examine this, the group-level simple slope of structuring behaviors at various levels of emotional climate were examined. This was done using the recommendations of Aiken and West (1991). Here, which cohesion was regressed on Structuring behaviors (x), differing levels of Emotional Climate – (1 SD/-1 SD; z_s), and the interaction of these two variables ($x*z_s$), along with the standardized covariates. Results indicated that the relationship between the number of structuring behaviors and cohesion was negative but non-significant when leaders also engaged in a high number Emotional Climate behaviors, ($\beta = -.32$, $p = .548$). When leaders engaged in few Emotional Climate behaviors, however, the relationship between Structuring Behaviors and cohesion was more strongly negative and became significant ($\beta = -1.11$, $p = .039$). This indicated that engaging in a high number of Emotional Climate behaviors "buffered" the impact of structuring behaviors on cohesion, which otherwise negatively related to first session cohesion

levels. The size of this effect was quite large, with a difference of d = .79 between high and low levels of emotional climate behaviors. This ordinal interaction is depicted in *Figure 10* below.

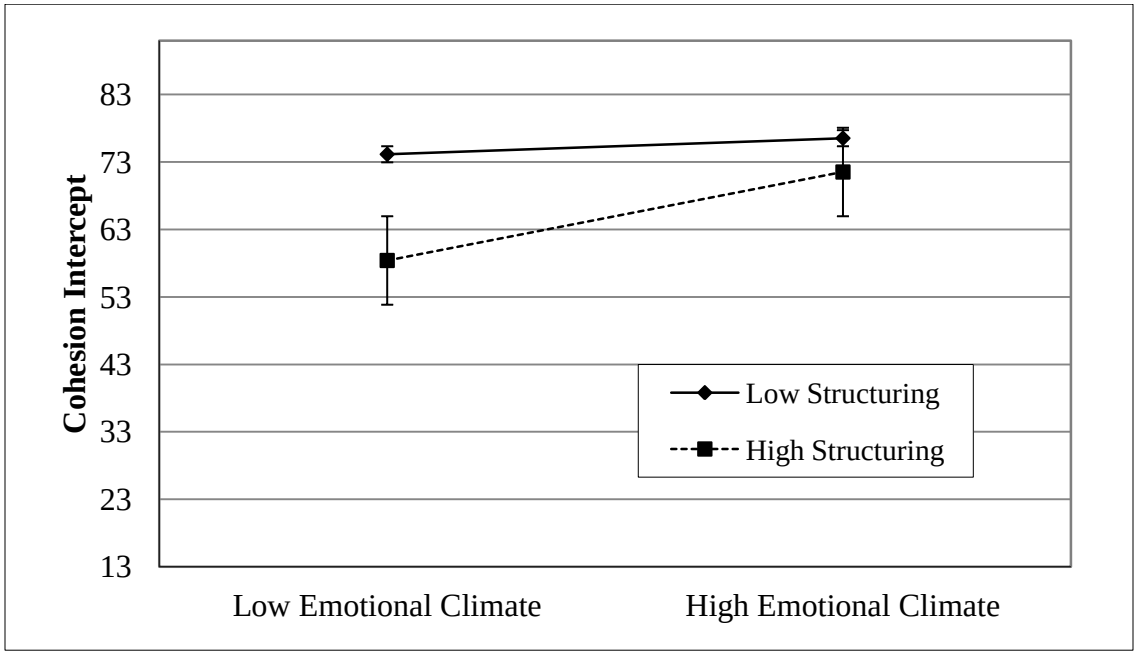


Figure 10. Moderating Effects of Emotional Climate Behaviors on the Relationship Between Structuring Behaviors and Initial Cohesion

With respect to the other group-level predictors in the model, verbal interaction behaviors were found to significantly relate to intercept cohesion scores when the three-way interaction and two-way interaction terms were included in the model, but not in the minimally-adequate or main effects model. As discussed by Aiken and West (1991), main effect terms do not have the same interpretation when an interaction is included. The main effects of a variable without an interaction term in the model examines the average effect of that variable ignoring the impact of the other variables. However, once an interaction is included, these terms represent the simple slope for variables, or the strength of the relationship between that predictor and the outcome at a particular level of the other variables. This suggested a possible moderating relationship but one that was not included in the analysis. As such, a possible positive relationship between verbal interaction and cohesion was implicated but not observed in the present study.

CHAPTER 5. DISCUSSION

This study provided a new understanding of the trajectory of cohesion across the life-span of short-term psychotherapy groups; offering support for individual and group factors that may influence such change. Results indicated that while clients and leaders have a high degree of agreement regarding cohesion (r^2 = .14 – .69), there is low agreement amongst members within groups (ICC(2) = .42 – .70), suggesting that cohesion may be best measured at the individual level. Still, non-trivial amounts of intragroup dependency were found across the eight weeks (ICC = .13 - .31), demonstrating that a multilevel model is necessary in examining changes in cohesion across time.

Modelling cohesion as a client-level outcome while accounting for the nested structure of the data resulted in a best-fitting piecewise linear-quadratic growth model. Here, early stage (session 1-4) cohesion increased at a constant, linear rate while late stage (session 4-8) increased in a convex, quadratic fashion. On average, groups began with 'moderate' to 'high' levels of cohesion at the first session while group cohesion continued to grow across time. Overall cohesion increased a total of 3.85 standard deviations across the eight weeks. A strong, negative association between initial cohesion scores and early stage growth was found, such that a one standard deviation increase in cohesion at time one led to a .69 standard deviation decrease in slope (or a drop in growth of 1.06 points per session). This was hypothesized to result from the initially high levels of cohesion in some groups which left little room for positive growth.

Individual-level predictors of intercept and slope terms included attachment style, psychological distress, and self-esteem. Demographic variables included gender, age, sex, and race. Of the included demographic variables, gender (identifying as female) was found to have a small, positive association with week 1 cohesion while being a person of color was found to have

a moderate, negative association with session 4-8 quadratic cohesion change. Controlling for these demographic variables, only avoidant attachment was found to have a moderate, negative relationship with first session cohesion, while only anxious attachment was found to have a moderate, positive relationship with cohesion change during sessions 1-4. Avoidant attachment explained 6.7% of the total individual variance in initial cohesion while anxious attachment explained 12.2% of the variance in cohesion slope in sessions 1-4. Together, the included individual-level predictors explained 11% of the individual variance in intercept scores and 15% of the individual variance in cohesion slope scores for sessions 1-4.

Undergraduate RAs were found to be reliable raters of first-session leadership behaviors, with a high degree of interrater agreement under an average measures framework. Several of the leadership behaviors hypothesized by the literature to be importantly-related to the development of cohesion occurred in nearly all groups (e.g. modelling expressions of open and genuine warmth) while many occurred in none or very few of the groups (e.g. reframing member-to-member feedback). Use of latent CFA methods removed poor performing items and resulted in improved internal consistency and more normally-distributed scales; with possible implications for a revised, 16-item first-session GPIRS scale for future research.

Group membership ultimately accounted for a moderate amount of variation in initial cohesion and cohesion change, with a decreasing trend across time. For the piecewise linear-quadratic model, group membership accounted for 22% of variation in intercept values, 20% of early-stage linear slope, 14% of session 4-8 linear slope, and 2.7% of variation in session 4-8 quadratic slope. For the linear model, group membership accounted for 21% of the variability in intercept scores and 19% of the variability in slope scores. Of the group-level variance terms,

only the linear model group-level intercept variance differed significantly from zero. This permitted an examination of group-level contributors to initial cohesion scores only.

Using the linear model to examine group-level predictors, non-significant relationships were observed between cohesion intercept scores and group attendance, group membership, and leader self-efficacy. With respect to leadership behaviors, however, an interaction was observed between structuring behaviors and emotional climate. An examination of the simple effects in this interaction revealed a buffering effect of emotional climate on structuring behaviors such that when emotional climate behaviors were frequent, structuring behaviors did not impact initial cohesion. When emotional climate behaviors were infrequent, however, structuring behaviors negatively impacted initial levels of cohesion. Finally, preliminary support for a positive association between verbal interaction behaviors and initial cohesion was observed.

Implications

Leader- and member-rated cohesion. One important finding in the present study is in regards to how best measure the construct of cohesion. This includes both at which level (individual or group) as well from which perspective (member, observer, leader) cohesion should be measured. The present investigation finds that these two questions are importantly related given that a multiple indicator model which allows for latent agreement and residual error necessitates that indicators are at the same level. Aggregating client-level ratings to the group level should only be done when ICC(2), a measure of dependency that is scaled for group size, is at or above .70 (Baumgartner, Jackson, Mahar, & Rowe, 2003; Gold, Kivlighan, & Patton, 2013). While early session agreement between group members was sufficient at .70, no subsequent session met this criterion. As such, a multiple-indicator model was not indicated, and

may not be a good methodological approach for group therapy research should this finding hold in future studies.

Many alternate models have been proposed in group-therapy research, but none address the unique situation of having multiple group member ratings and one leader rating. One area with emerging support in the literature is the One-With-Many reciprocal design, outlined by Marcus, Kashy, and Baldwin (2009). Still, such designs necessitate that not only does every member rate cohesion with their group but that each group therapist rates their perceptions of each clients' cohesion with the group, forming a dyadic data pattern. Cohesion, as a group-level phenomenon, however, leads to only a single rating by a group leader. This presents a difficulty in coding for group membership while at the same time accounting for the role of each participant (leader and client). While other researchers have begun to apply the One-With-Many design to group therapy (e.g. Garcia, Meagher, & Kenny, 2014), these often account for group roles by including the number of persons occupying a given role (termed actor similarity). While useful in examining how having more members with similar characteristics in groups may relate to outcomes, it is not applicable to a leadership role in which there is exactly one leader per group and a varying numbers of clients. Beyond this, neither perspective can examine longitudinal concordance (how a previous session's rating impacts continued agreement), but rather must examine session-level concordance alone.

As such, it may not be advisable for group researchers to examine cohesion as a "group-as-whole" phenomenon, at least when it is the *member-to-member* structural level within groups that is being assessed as is the case in the most frequently cited group therapy measures (i.e. the GCQ, MacKenzie, 1983; and the GES, Wilson et al., 2008). This is consistent with recent research which has found a poor fit of the GCQ at the group level (Johnson et al., 2006).

Instead, cohesion can be assessed at the individual level while accounting for group-level dependencies through HLM methods, as used in the present study. This is especially the case given that leadership behaviors as rated by independent observers were found to relate to client ratings of cohesion, indicating that clients are sensitive to leadership behaviors occurring in the group.

With respect to agreement between client and leader ratings, the results of the present investigation showed levels of agreement that trended consistently with previous research in the area. Lorentzen, Sexton, and Høglend (2004) found moderate to strong agreement between clients and leaders (mean CCF=.47, range .31-.80), with levels being the highest at weeks 16-30 of a 90-week treatment group. These researchers hypothesized that this trend may relate to leaders learning how to perceive client's view of the relationship, and that the drop in mid-sessions could result from leader interpretations of client defenses.

In the present investigation involving short-term therapy groups, agreement between clients and leaders on levels of cohesion was also moderate to strong across the groups (latent mean r = .56, range .38-.83), demonstrating a trajectory mirroring Lorentzen and colleagues' (2004) findings. In addition, similar to Lortenzen and colleagues' study, agreement was highest shortly after the beginning of group (week 3), after which point agreement steadily decreased. Future research may wish to examine if the general agreement between leaders and clients follows a trajectory of greatest agreement shortly after the beginning of group with a subsequent decreasing trend in both short and long-term treatments. Given findings that symptom reduction follows a similar but condensed trajectory in short-term versus long-term treatment (e.g. Baldwin, Berkeljon, Atkins, Olsen, & Nielsen, 2009), this may be theoretically plausible. Such a finding would be theoretically and empirically important given that a higher degree of

concordance between client and leader ratings has been related to changes in symptomatic outcome (Lorentzen, Sexton, and Høglend, 2004), and could have implications for how to increase client and leader agreement.

Cohesion change across time. A core contribution of the present study is the examination of multiple possible trajectories of cohesion change while accounting for group dependencies in the data. Through multiple decades of research, several possible trajectories of cohesion have been hypothesized. These vary from a cubic trend of low-high-low-high (MacKenzie, 1983), a quadratic trend of high-low-high (Castongauy, et al., 1998; Kivlighan & Lilly, 1997), a linear trend of low to high (Tschuschke & Dies, 1994), and a linear trend of moderate to high (Taube-Schiff et al., 2007). Of these, the present study most closely aligns with Taube-Schiff and colleagues' 2007 model, with important differences. While cohesion started at a moderate to high level initially, it showed a positive, linear trend during the initial half of group (sessions 1-4) followed by an increasing quadratic trend in the latter half of group (sessions 4-8). An important 'knot' (Kohli et al., 2015, p. 260) emerged after session 4, at which point the shape of cohesion change took on a different trajectory. This may suggest important differences between the development of early and late cohesion. Consistent with past theorists on short-term group psychotherapy (namely Shapiro, Peltz, & Bernadett-Shapiro, 1997), an early linear trajectory may result from the initial level of cohesion building that occurs early on as members increasingly self-disclose. Similar to theorists who have proposed that the "storming" stage of groups leads to conflict as group members work through interpersonal patterns (e.g. Tuckman, 1965; Tuckman, & Jensen, 1977), the present research suggests a slight dip in cohesion between sessions 4 and 5, a trend that emerged across groups. This dip quickly recovers, however, as quadratic growth takes over and cohesion increases quickly in the last sessions. This may be a

result of what Shapiro, Peltz, & Bernadett-Shapiro (1997) refer to as the "trust boost", which occurs at the end of group as members prepare to say goodbye.

Although a piecewise linear-quadratic trend was the best fitting model to the data, the complexity of such a model proposes difficulty in examining predictor variables under a LGC format with groups fewer than 40. As such, a linear growth model that does not impose strict linearity at all time-points can also be used. Under both models, cohesion grows considerably throughout the life-span of short-term psychotherapy groups, with an effect size of approximately $d = 1.59$ across the eight weeks, or an increase in nearly 4 standard deviations across the life-span of a short-term psychotherapy group. Thus, while cohesion starts at a moderate level, substantial changes still occur across time. This promotes the importance of investigating individual and group-level contributors to such change.

Client contributions. The first notable component in the present study was the large degree of variance explained by individual-level variables in the development of cohesion across time; with individual differences accounting for between 80-97% of the variance in cohesion. This is consistent with much of the research suggesting that clients are the single most potent contributor to outcomes in psychotherapy (e.g. Tallman & Bohart, 1999; Ahn & Wampold, 2001). Thus, although the present study sought to emphasize the importance of leadership behaviors, it is evident that client predictors ultimately explain the greater degree of variance.

With regard to specific individual characteristics, the present findings included demographic and personality variables in the development of cohesion. Several findings here are noteworthy. First, gender (identifying as female) was found to have a small, positive association with initial cohesion. This may be the result of the large number of women participants in the present study, with women composing 79.6% of the sample. A similar finding may be in the

moderate, negative relationship between being a person of color and late stage cohesion growth, where only 5 persons of 128 (4%) identified as a non-white.

Both of these findings are consistent with group identification theory (Henry, Arrow, & Carini, 1999) and social identity theory (SIT; Tajfel & Turner, 1979). First, group identification theory suggests that one's perceptions of similarity to a group relates strongly to their perceptions of belongingness. In one recent example, actor-partner modelling was used to find that the more similar in gender a person was to other members of a discussion group, the greater was their perceived satisfaction with that group (Garcia, Meagher, and Kenny, 2014). This suggests that the great degree of similarity between persons with respect to gender and racial identity in the present investigation may have in part contributed to the high levels of cohesion observed.

Similarly, SIT examines the conditions under which successful versus competitive intergroup contact may occur. It suggests that, when threatened, persons tend to accentuate between-group differences, promote their own group's value, and devalue the worth of outgroup identities (Hogg, 2006). On the other hand, there are many factors that can lead to positive intergroup contact, and many of these are in line with general processes in group psychotherapy. According to SIT, the development of cohesion requires equal status, where differences are discussed but put on equal rank; a focus on common, superordinate goals; intergroup cooperation; personal interaction; and inclusion of an authority figure who encourages egalitarian attitudes (Dovidio, Gaertner, & Kawakami, 2003). While these features may be common to most groups, the present study suggests that group leadership may wish to pay specific attention to social identity and intergroup contact theories in the promotion of cohesion

in groups with diverse composition given that minority identities appear to result in decreased perceptions of cohesion.

With respect to client-level characteristics, some interesting findings were observed between psychological distress and attendance as well as between attachment style and the development of cohesion. First, psychological distress was a strong predictor of attendance in therapy, with high levels of initial distress predicting subsequent drop-out (a one standard deviation increase in distress led to an increase in the odds-ratio of drop-out by 2.7). This is opposite to what has been suggested by previous research, which has found that those with low levels of distress tend to terminate treatment earlier (e.g. Ogrodniczuk, Piper, & Joyce, 2006).

What may differentiate the present study from those such as Ogrodniczuk and colleagues' (2004) study is the composition of the group. Ogrodniczuk and colleagues (2004) were examining those with personality disorder diagnoses, while the present study examined those who experienced interpersonal hurts in a general outpatient setting. In the present study, the mean level of psychological distress was 1.57, a level below the clinical cut-off for women but above the cut-off for men. It is likely that Ogrodniczuk and colleagues' participants were at higher levels of distress. In that case, attendance may be more related to identifying as a 'group isolate" (Yalom, 2005), as discussed above, rather than one's distress level. Future research will need to clarify the relationship between distress and attendance in order to identify which factors may moderate this relationship.

With respect to attachment style, there was clear support for a relationship between cohesion and anxious and avoidant attachment. The present study found support for a negative relationship between avoidant attachment and initial levels of psychological distress. This is in line with previous research suggesting that those with avoidant attachment styles are more likely

to perceive the group climate as hostile and less attractive (e.g. Kivlighan & Angelone, 1992; Chen & Mallinckrodt, 2002), adding to this literature by indicating that these members are also less likely to view their groups as cohesive. While avoidant attachment was not related to cohesion change across time, because initial cohesion varied quite drastically and impacted the later development of cohesion, this is an important finding.

On the other hand, anxious attachment was found to positively relate to early stage change but not initial levels. This is inconsistent with past literature which suggests that greater attachment anxiety is related to early termination (Marmarosh et al., 2005; Huang et al., 2013). While inconsistent with past research, this finding is consistent with theory on attachment style. Attachment styles are behaviors that reflect the expectations and beliefs persons have about themselves and their relationships based on their attachment histories. Those with anxious attachment styles are more likely to excessively seek attention and care, looking to others to help manage their anxiety (Fraley & Shaver, 2000). As such, it is not surprising that those with anxious attachment styles are more likely to express a closer bond to their therapy groups, especially early on in therapy. Although it is interesting to note that neither attachment style was related to later cohesion in group, the lack of significant variance in session 4-8 cohesion change does not permit conclusions about this time-frame.

These findings may support the importance of assessing attachment style early in psychotherapy groups. This may help inform researchers about potential later developments in cohesion, potentially useful interventions, and the possibility of needing to balance attachment styles. More research is needed to examine whether this early relationship between attachment style and cohesion relates to differences in outcomes and what different attachment styles mean at the group level.

Group and leader contributions. In the present investigation, non-significant between-group variance in the cohesion slope term prevented an examination of variables predictive of cohesion change. However, significant variance existed in initial cohesion levels after session 1, permitting an exploration of group-level predictors. Many of the hypothesized variables were not significantly related to initial cohesion, including group attendance, the number of group members, and the leaders' counseling self-efficacy. Instead, leader behaviors were the only group-level variable found to have a significant impact on client-rated cohesion. The overall variance explained by the included group-level variables was between 33% and 44%, of which the included leadership behaviors explained approximately 25%.

This has profound implications for leaders of psychotherapy groups. First, notable about the present investigation was that the observed relationship was between client-rated cohesion and observer-rated leader behaviors, meaning that the observed relationship is between independent ratings. Because of this observational method variance, the present findings are more likely to represent a true relationship between leadership behaviors and cohesion (Campbell & Fiske, 1959). Secondly, the findings are very similar to past research suggesting that therapists explain a small but significant amount variance in cohesion (3-7%, Bakali et al., 2010). In the present investigation, group differences accounted for 20% of the variance in intercept and slope terms, with leaders predicting about 25% of that variance, or roughly 5% of the total variance.

Another finding of the present investigation was that leadership behaviors were found to differ between the two treatment conditions in the present study, with more structuring and verbal interaction behaviors in the process treatment as compared to the forgiveness treatment condition. These differences were found to be meaningfully large, with process groups

involving, on average, 5-18 more structuring behaviors and 4-23 more verbal interaction behaviors in a 90-minute session. Still, treatment condition did not significantly influence initial cohesion scores or the slope of cohesion across time. Primarily, these results seem to suggest that differences between treatment conditions can be effectively captured by differences in leadership behaviors; a finding which may be capitalized on by future research into group processes.

The present study also finds some support for the leadership categories as outlined by the authors of the GPIRS, and specifically provides support for measuring each of the GPIRS domains given the observed interaction effects between the three domains. This is to suggest that it may be important to take a more nuanced study of leadership behaviors than previously anticipated. Dividing leadership behaviors into two categories based primarily on a distinction between protecting members and involving members (such as in models by Morran, Stockton, and Whittingham, 2004), while helpful in conceptualizing group leadership, may not allow researchers to test and observe important moderating relationships.

More specifically, the primary finding in the present investigation was that structuring behaviors have a significant, negative relationship with initial cohesion, but only at low levels of emotional climate behaviors. In other words, structuring behaviors actually resulted in lower cohesion ratings at time one, but if leaders also engaged in emotional climate behaviors, no such negative relationship is observed. In addition, some preliminary evidence suggested a potential positive relationship between verbal interaction behaviors and initial cohesion.

Foremost, the findings from the present study suggest that, contrary to the study hypotheses, leaders should be aware of the potentially negative impact that structuring behaviors might have on cohesion change. Such behaviors include things like discussing fears and

concerns about group, discussing roles and responsibilities, going over the rationale for treatment, and discussing group rules. While such things might be important for other curative therapeutic factors hypothesized to relate to client outcomes (for example, teaching about psychological problems and recovery, developing social skills, enforcing group rules, Yalom, 2005), such behaviors may interfere with the development of cohesion. If, however, group leaders are also attentive to a group's emotional climate, for example helping members identify their feelings and concerns and responding to the meaning of member comments, this negative relationship is not observed. As such, insofar as structuring behaviors are considered necessary for therapeutic gains, leaders should be careful to attend to the emotional components of their group in order to avoid negatively impacting their group's cohesiveness. This is consistent with previous findings suggesting that the most effective leaders are those who attend to the relationship (Lambert & Barley, 2001; Vocisano et al., 2004).

Limitations and Future Research

The largest limitations in the present research are those relating to the measurement of leadership behaviors in psychotherapy groups. As a relatively new area of research within the field of counseling psychology, approaches to measurement will likely need continued refinement. One limitation in the present study was difficulties and subsequent revisions of the Group Psychotherapy Intervention Rating Scale (GPIRS; Chapman, Baker, Porter, Thayer, & Burlingame, 2010). Specifically, at the level of individual behaviors, many items of the GPIRS did not occur while some occurred in nearly every group. The high frequency of a single response for a given behavior led to highly leptokurtic and skewed distributions; resulting in low internal consistency under traditional estimates (e.g. Cronbach's alpha). Still, as a measure of individual behaviors which may occur independently of one another (e.g. a leader who restates

corrective member feedback may not also share a brief personal experience, despite the fact that these items are within the same domain), the GPIRS is perhaps best thought of as an *index* rather than a *scale* (Strieder, 2012).

Indices are not effectively evaluated through internal reliability measures (Strieder, 2012), and some modifications were made in order to allow for averaging within domains and using these as predictors of cohesion outcomes. Specifically, those items that were removed were those that almost all or nearly no leaders utilized. Items in which a single value occurs in nearly all groups have low covariance with other variables because they are primarily static and therefore cannot aid in the prediction of differences in outcome variables. The resulting scale, a 16-item revised GPIRS, was thus based primarily on the specific behaviors exhibited in the present study. As such, the revised scale might be thought of as a "first session leadership behaviors scale." This scale may be highly useful in future research, however the generalizability of this scale to other psychotherapy groups may be limited to the extent that the behaviors of leaders in these groups does or does not reflect the behaviors of leaders in other settings.

Secondly, the scaling method of the GPIRS was difficulty in interpretation at the mid-point of the scale. That is, because the GPIRS combines both quality and frequency in its ratings, it is difficult to decipher whether a score at the mid-point reflects a behavior occurring infrequently with high quality or frequently with low quality. As such, the present investigation could not examine the effect of doing interventions poorly, but rather only less frequently. In addition, because the GPIRS only examines positive leadership behaviors (i.e. those thought to promote cohesion), the present study did not assess the effects of negative interventions, including such things as acting defensively or judgmentally in response to member behaviors (Morran, Stockton, & Whittingham, 2004).

Another limitation of the present study was the homogeneity of the study sample. Specifically, the sample was comprised mostly of white female clients in their 40s to 60s who had experienced interpersonal hurts. In addition, there was a relationship between high levels of psychological distress and drop-out in the present study, indicating that participants were likely at lower levels of psychological distress. Moreover, the marked similarity of demographics and presenting concerns may have directly impacted the cohesion in the groups, contributing to the initially high levels of cohesion and its consistent growth across time.

The high level of homogeneity across groups may pose another significant limitation to the present study, which is reduced power. Because cohesion began at a relatively high level and because between group variance decreased as time went on, there was less variance available for prediction by the included covariates. Although there was significant change across groups (an effect size of nearly 3.84 SD across the eight sessions), the overall power of the present study to find an effect was reduced.

As such, although the present study utilized an existing dataset, a power analysis was conducted to determine how much power the model possessed (given the design characteristics and parameters of interest) to detect an effect. For this, Montecarlo simulation utilizing MPLUS software was utilized. To determine power to detect model misspecification, parameter estimates were varied and then examined for the corresponding proportion of times the simulation resulted in a chi-square above the critical value. For the present investigation, several simulations were run. The first represented the unconditional model; examining the power to detect mean change in cohesion across time when the population parameter was set as linear, quadratic, or cubic. The second model represented the conditional model, which included leadership behaviors as a predictor of the intercept and slope cohesion terms.

Parameter estimates from the present dataset were used for the power analyses. These included mean change in cohesion (the model effect size), the variance of each time-point, and the test-retest reliability of cohesion across this same period (i.e. the model's covariance structure). Item means and variances were considered standard normal for the power analysis (~N(0,1)). An average of the correlations in cohesion scores across time was set as the degradation of test-retest reliability across the 8 time-points and used to specify the covariance structure for the power analysis. For the baseline model and full models, sample sizes were set at n=128.

Examining the output of the unconditional model, estimates of mean cohesion change and test-retest reliability estimates were observed to be closely approximated to the population parameters. At n = 128, the average chi-square value for the 10000 replications was $\mu_{\chi 2} = 56.626$ (SD=13.451). At a sample size of 128, power to detect a correlation between intercept and slope values was also strong, equal to .988 (indicating that 99% of the time the model would correctly reject the null hypothesis that cohesion and slope were uncorrelated). Power to detect linear slope value different from zero was near perfect ($1-\beta = 1.000$), as was the power to detect variances and residual variances ($1-\beta = 1.000$). Power to detect quadratic slope was adequately large ($1-\beta = .943$), while the model was unable to converge when estimating a cubic slope. As such, while the present investigation was very likely able to detect differences from the specified growth parameters in the linear and cubic models, it may have had insufficient power to detect a cubic model, should it have existed.

For the full model, the slope and intercept were entered as variables regressed on a hypothesized predictor; leadership behaviors. In the unconditional model an intercept and slope covariance value of -.05 was observed. This was used as the estimated relationship between the

two variables in the full model. With respect to relationships between leader behaviors and group cohesion at a single time point, values have been found to hover around *r* = .70 (Antonuccio, Davis, Lewinsohn, & Breckenridge, 1987). Only one known study has estimated the strength of the association between leadership behaviors and changes in cohesion across time. Chapman and colleagues (2010) found a correlation between leadership behaviors and cohesion at *r* =.40. As such, hypothetical relationships of .10 (weak relationship), .20 (small relationship), .40 (moderate relationship), and .70 (strong relationship) were tested.

Specifying a relationship between cohesion and leader behaviors of .10, power to detect the slope term at n = 128 was 1-β = .23 for intercept and slope terms. At an effect size of .20, power jumped to 1-β = .67. At an effect size of .25, power became adequate at 1-β = .84. This indicated that the present study had sufficient power to detect an effect of .25 or larger for the relationship between leader behaviors and intercept and slope terms under a linear model.

One additional component is necessary in a discussion of power in the present investigation. The data is in a nested structure; with participants nested within groups. Such a data structure results in increases in needed sample size. As a quick shorthand and conservative estimate, the following calculation can be used to determine an adjusted *n* needed to obtain a given power:

$$n' = n(1 + \rho(m - 1))$$

Where ρ is the estimated ICC and *m* is the number of participants sampled per cluster (Doros & Lew, 2010). Average ICC values in the present investigation were .21. Using this calculation and assuming a determined sample size of 128, power to detect an effect at .80 in the present investigation results in an adjusted *n* of between 128(1+.21(6-1). Ideally then, an ideal sample size would be approximately 266 persons, or approximately 40 groups. This was also

consistent with the number of parameters specified under the piecewise linear-quadratic model (df = 40). Given this, future efforts to examine similar study hypotheses should likely aim for a conservative estimate of approximately 266 individuals across 40 psychotherapy groups.

Conclusion

The present study provides a new understanding of the trajectory of cohesion change in short-term psychotherapy groups; identifying a general increasing trend under a piecewise linear-quadratic growth model. The present investigation also finds that, for studies with fewer groups, the trajectory of cohesion can be effectively examined using a more simple linear model. Under such a model, group membership accounts for approximately 20% of the variance in intercept and slope terms with 80% of the variance attributable to individual differences. Individual client characteristics predictive of cohesion change include gender, race, and attachment style, which together explain a small but significant percentage of individual-level variance in intercept and slope terms (11% and 15%, respectively). Group-level predictors of initial cohesion include structuring behaviors, which have a negative relationship with cohesion moderated by levels of emotional climate. Here, structuring behaviors have a non-significant relationship with initial cohesion when emotional climate behaviors are high, but a negative relationship at low-levels of emotional climate. Preliminary support for a positive relationship between verbal behaviors and initial cohesion is also implicated; however results of the present study are inconclusive. The included group-level variables in the present study accounted for approximately 33-44% of the variance in initial cohesion scores.

Given the strong relationship between cohesion and consequential outcomes including symptom reduction and improvements in interpersonal functioning, the present study has important implications for group therapists. Primarily, the present study suggests that while

individual differences have the most robust impact on cohesion development, group characteristics also have a moderate impact on cohesion change and leaders have a small but significant impact on such change. Given this, it will be important for future research to continue to examine the unique roles of client, group, and leader variables in the development of cohesion across the life-span of psychotherapy groups across various populations. Doing so will continue to help empower group therapists to facilitate deeply involving groups, through which they can provide an important antidote to the negative impacts of social isolation and greatly contribute to their clients' psychological health.

CHAPTER 6. REFERENCES

Ahn, H. N., & Wampold, B. E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology*, 48(3), 251.

Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions.* Thousand Oaks, CA: Sage Publications Inc.

Al-Darmaki, F. R. (2005). Counseling self-efficacy and its relationship to anxiety and problem-solving in United Arab emirates. *International Journal for the Advancement of Counselling, 27*(2), 323-335.

Antonuccio, D.O., Davis, C., Lewinson, P.M., & Breckenridge, J.S. (1987). Therapist variables related to cohesiveness in a group treatment for depression. *Small Group Behavior*, 18(4), 557-564.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37, 122-147.

Baumeister, R.F. & Leary, M.R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin,* 117(3), 497-529.

Baumgartner, T. A., Jackson, A. S., Mahar, M. T., & Rowe, D. A. (2003). *Measurement for evaluation in physical education and exercise science* (7th ed.). Boston, MA: McGraw-Hill.

Alonso, J. T. (2012). Cohesion's relationship to outcome in group psychotherapy: A meta-analytic review of empirical research. *Unpublished dissertation, ProQuest Information & Learning*.

Antonuccio, D.O., Davis, C., Lewinson, P.M., & Breckenridge, J.S. (1987). Therapist variables related to cohesiveness in a group treatment for depression. *Small Group Behavior, 18(4),* 557-564.

Baldwin, S. A., Berkeljon, A., Atkins, D. C., Olsen, J. A., & Nielsen, S. L. (2009). Rates of change in naturalistic psychotherapy: Contrasting dose–effect and good-enough level models of change. *Journal of Consulting and Clinical Psychology*, 77(2), 203.

Bakali, J. V., Wilberg, T., Hagtvet, K. A., & Lorentzen, S. (2010). Sources accounting for alliance and cohesion at three stages in group psychotherapy: Variance component analyses. *Group Dynamics: Theory, Research, and Practice, 14*(4), 368.

Barkham, M., Gilbert, N., Connell, J., Marshall, C., & Twigg, E. (2005). Suitability and utility of the CORE-OM and CORE-A for assessing severity of presenting problems in

psychological therapy services based in primary and secondary care settings. *The British Journal of Psychiatry*, *186*(3), 239-246.

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., ... & McGrath, G. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Toward practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, *69*(2), 184.

Baumgartner, T. A., Jackson, A. S., Mahar, M. T., & Rowe, D. A. (2003). *Measurement for evaluation in physical education and exercise science* (7th ed.). Boston, MA: McGraw-Hill.

Bednar, R. L., & Kaul, T. J. (1978). Experiential group research: Current perspectives. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (2nd ed., p. 769–815). New York: Wiley.

Bentler, P. M., & Satorra, A. (2010). Testing Model Nesting and Equivalence. *Psychological Methods*, 15(2), 111–123.

Black, M. (2010). Bridging As a Tool to Avoid Scapegoating. In Fehr, S.S. (Ed.). *101 Interventions in Group Therapy* (p.155-158). New York, NY: Routledge/Taylor & Francis Group.

Bohart, A. C., & Tallman, K. (1999). *How clients make therapy work: The process of active self-healing*. American Psychological Association.

Box GEP, Jenkins GM., and Reinsel, G.C. (2008). *Time series analysis: Forecasting and control*. Hoboken, NK: John Wiley & Sons, Inc.

Braaten, L. J. (1990). The different patterns of group climate critical incidents in high and low cohesion sessions of group psychotherapy. *International Journal of Group Psychotherapy, 40(4),* 477-493.

Brennan, K. A., Clark, C. L., & Shaver, P. R. (1998). Self-report measurement of adult attachment: An integrative overview. In J.A. Simpson & W. S. Rholes (Eds.). *Attachment theory and close relationships* (p. 46-76). New York: Guilford.

Budman, S. H., Soldz, S., Demby, A., Feldstein, M., Springer, T., & Davis, M. S. (1989). Cohesion, alliance and outcome in group psychotherapy. *Psychiatry*,*52*(3), 339-350.

Burlingame, G., Fuhriman, A., & Johnson, J. (2002). *Cohesion in group psychotherapy*. In J. C. Norcross (Ed.). Psychotherapy relationships that work (pp. 71–87). Oxford University Press.

Burlingame, G. M., McClendon, D. T., & Alonso, J. (2011a). Cohesion in group therapy. *Psychotherapy*, *48*(1), 34.

Burlingame, G.M., McClendon, D.T., & Alonso, J. (2011b). Cohesion in group therapy. In Norcross, J.C. (Ed.) *Psychotherapy Relationships That Work: Evidence-Based Responsiveness.* Oxford, UK: Oxford University Press.

Burrow, T. (1928). The Basis of Group-Analysis, or the Analysis of the Reactions of Normal and Neurotic Individuals. In MacKenzie, K.R. (Ed.). (1992). *Classics in Group Psychotherapy,* New York, NY: The Guilford Press.

Campbell, D.T., & FiskeD.W. (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin,* 56, 81-105.

Castonguay, L., Pincus, A., Agras, W., & Hines, C. (1998). The role of emotion in group cognitive-behavioral therapy for binge eating disorder: When things have to feel worse before they get better. *Psychotherapy Research*, *8*(2), 225-238.

Chapman, C. L., Baker, E. L., Porter, G., Thayer, S. D., & Burlingame, G. M. (2010). Rating group therapist interventions: The validation of the Group Psychotherapy Intervention Rating Scale. *Group Dynamics: Theory, Research, and Practice*, *14*(1), 15.

Chen, E. C., & Mallinckrodt, B. (2002). Attachment, group attraction and self-other agreement in interpersonal circumplex problems and perceptions of group members. *Group Dynamics: Theory, Research, and Practice, 6(4),* 311.

Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: A meta-analytic review. *Psychological Bulletin, 116,* 457-475.

Coredor, G.W. & Foreman, D.I. (2014). *Nonparametric statistics: a step-by-step approach.* Hoboken, NJ: John Wiley & Sons, Inc.

Crowe, T. P., & Grenyer, B. F. (2008). Is therapist alliance or whole group cohesion more influential in group psychotherapy outcomes? *Clinical Psychology & Psychotherapy*, *15*(4), 239-246.

Curran, P.J., Obeidat, K., & Losardo, D. Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development: Official Journal of the Cognitive Development Society, 11*(2), 121-136.

Delucchi, K. L., & Bostrom, A. (2004). Methods for analysis of skewed data distributions in psychiatric clinical studies: working with many zero values. American Journal of Psychiatry, 161(7), 1159-1168.

Dinger U, Schauenburg H. (2010). Effects of individual cohesion and patient interpersonal style on outcome in psychodynamically oriented inpatient group psychotherapy. Psychotherapy Research 20(1): 22-9.

Doros and Lew (2010). Design based on intra-class correlation coefficients. *American Journal of Biostatistics 1 (1):* 1-8, 2010.

Dovidio, J. F., Gaertner, S. L., & Kawakami, K. (2003). Intergroup contact: The past, present, and the future. *Group Processes & Intergroup Relations*, 6(1), 5-21.

Duncan, T.E., Duncan S.C. & Strycker, L.A. (2006). *An introduction to latent variable growth curve modeling.* Mahwah, NJ: Lawrence Erlbaum Associates.

Elkin, I., Falconnier, L., Martinovich, Z., & Mahoney, C. (2006). Therapist effects in the NIMH Treatment of Depression Collaborative Research Program. *Psychotherapy Research*, 16, 144–160.

Enders, C.L. (2003). Using the expectation maximization algorithm to estimate coefficient alpha for scales with item-level missing data. *Psychological Methods, 8(3),* 322-337.

Ettin, M. F., Cohen, B. D., & Fidler, J. W. (1997). Group-as-a-whole theory viewed in its 20th-century context. *Group Dynamics: Theory, Research, and Practice*, *1*(4), 329.

Evans, C. R., & Dion, K. L. (1991). Group cohesion and performance a meta-analysis. *Small Group Research*, *22*(2), 175-186.

Fehr, S. S. (Ed.). (2012). *101 interventions in group therapy*. New York, NY: Routledge.

Field, A.P. (2013) *Discovering statistics using IBM SPSS statistics.* London, Sage.

Forsyth, D. R., & Elliott, T. R. (1999). Group dynamics and psychological well-being: The impact of groups on adjustment and dysfunction. In Kowalski, R.M., & Leary, M.R. (Ed*). The social psychology of emotional and behavioral problems: Interfaces of social and clinical psychology*. 339-361. Washington, DC, US: American Psychological Association.

Forsyth, D.R. (2012). Social influence and group behavior. In Weiner, I.B. (Editor-in-Chief), Tennen, H.A., & Suls, M. (Ed.) *Handbook of psychology, Volume 5: Personality and social psychology, 2nd Edition.* Wiley.

Frances, A., Clarkin, J. F., & Perry, S. (1984). *Differential therapeutics in psychiatry: The art and science of treatment selection*. New York: Brunner/Mazel.

Frank, J. (1957). Some determinants, manifestations, and effects of cohesiveness in therapy groups. *International Journal of Group Psychotherapy*, *7*, 53-63.

Fuhriman, A., Drescher, S., Hanson, E., Henrie, R., & Rybicki, W. (1986). Refining the Measurement of Curativeness An Empirical Approach. *Small Group Research*, *17*(2), 186-201.

Garcia, R. L., Meagher, B. R., & Kenny, D. A. (2014). Analyzing the effects of group members' characteristics: A guide to the group actor–partner interdependence model. *Group Processes & Intergroup Relations*, 1-14.

George, D., & Mallery, M. (2010). *SPSS for Windows Step by Step: A Simple Guide and Reference, 17.0 update* (10a ed.). Boston: Pearson.

Gold, P. B., Kivlighan Jr, D. M., & Patton, M. J. (2013). Accounting for session-level dependencies in longitudinal associations of group climate and therapeutic factors in interpersonally focused counselor-training groups. *Group Dynamics: Theory, Research, and Practice,* 17(2), 81.

Guo, S. (2005). Analyzing grouped data with hierarchical linear modeling. *Children and Youth Services Review, 27*, 637-652.

Guo, S., & Hussey, D. (1999). Analyzing longitudinal rating data: A three-level hierarchical linear model. *Social Work Research*, 23(4), 258–269.

Gully, S.M., Devine, D.J., & Whitney, D.J. (2012). A meta-analysis of cohesion and performance: Effects of level of analysis and task interdependence. *Small Group Research, 43*(6), 702-725. doi: 10.1177/104649641248069

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in Quantitative Methods for Psychology, 8(1*), 23–34.

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice, 9*, 329–343.

Heinrich, C. J., & Lynn Jr., L. E. (2001). Means and ends: A comparative study of empirical methods for investigating governance and performance. *Journal of Public Administration Research and Theory, 11(1),* 109– 138.

Henretty, J. R., & Levitt, H. M. (2010). The role of therapist self-disclosure in psychotherapy: A qualitative review. *Clinical Psychology Review, 30,* 63–77.

Henry, K. B., Arrow, H., & Carini, B. (1999). A tripartite model of group identification theory and measurement. *Small Group Research, 30(5*), 558-581.

Hill, W. F. (1971). The Hill interaction matrix. *Personnel and Guidance Journal, 49*, 619–623.

Hill, C. E., & Lambert, M. J. (2004). Methodological issues in studying psychotherapy processes and outcomes. *Bergin and Garfield's handbook of psychotherapy and behavior change, 5*, 84-135.

Honos-Webb, L., Stiles, W.B., & Greenberg, L.S. (2003). A method of rating assimilation in psychotherapy based on markers of change. *Journal of Counseling Psychology*, 50 (2), 189-198.

Hornsey, M. J., Dwyer, L., & Oei, T. P. (2007). Beyond Cohesiveness Reconceptualizing the Link Between Group Processes and Outcomes in Group Psychotherapy. *Small Group Research*, *38*(5), 567-592.

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–response relationship in psychotherapy. *American Psychologist*, *41*, 159–164.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1), 1-55.

Huang, T., Hill, C., & Gelso, C. (2013). Psychotherapy engagers versus non-engagers: Differences in alliance, therapist verbal response modes, and client attachment. *Psychotherapy Research*, *23*(5), 568-577.

Hurley, J.R. & Brooks, L.A. (1988). Primacy of affiliativeness in ratings of group climate. *Psychological Reports, 62,* 123-133.

Johnson, J. E., Burlingame, G. M., Olsen, J. A., Davies, D. R., & Gleave, R. L. (2005). Group Climate, Cohesion, Alliance, and Empathy in Group Psychotherapy: Multilevel Structural Equation Models. *Journal of Counseling Psychology*, *52*(3), 310.

Jones, W. H., & Carver, M. D. (1991). Adjustment and coping implications of loneliness. In Snyder, C. R. & Forsyth, D.R. (Ed). *Handbook of social and clinical psychology: The health perspective.* Pergamon general psychology series, 162, 395-415. Elmsford, NY: Pergamon Press.

Joyce, A. S., Piper, W. E., & Ogrodniczuk, J. S. (2007). Therapeutic alliance and cohesion variables as predictors of outcome in short-term group psychotherapy. *International Journal of Group Psychotherapy*, *57*(3), 269-296.

Kendall, M. G., & Stuart, A. (1958). *The advanced theory of statistics.* New York: Hafner

Kiecolt-Glaser, J. K., Ricker, D., George, J., Messick, G., Speicher, C. E., Garner, W., and Glaser, R. (1984b). Urinary cortisol levels, cellular immunocompetency, and loneliness in psychiatric patients. *Psychosom. Med*. *46*, 15-24.

Kivlighan, D. M. & Angelone, E. O. (1992). Interpersonal problems: Variables influencing participants' perception of group climate. *Journal of Counseling Psychology, 39,* 468-472.

Kivlighan, D.M. & Lilly, R. (1997). Developmental changes in group climate as they relate to therapeutic gain. *Group Dynamics: Theory, Research, and Practice*, 1(3), 208-221.

Kivlighan Jr, D. M., & Goldfine, D. C. (1991). Endorsement of therapeutic factors as a function of stage of group development and participant interpersonal attitudes. *Journal of Counseling Psychology*, *38*(2), 150.

Kivlighan Jr, D. M., Patton, M. J., & Foote, D. (1998). Moderating effects of client attachment on the counselor experience–working alliance relationship. *Journal of Counseling Psychology*, *45*(3), 274.

Kline, R.B. (2011). *Principles and practice of structural equation modeling* (3rd Ed.). New York, NY: The Guilford Press.

Knox, S., & Hill, C. E. (2003). Therapist self-disclosure: Research-based suggestions for practitioners. *Journal of Clinical Psychology*, *59*(5), 529-539.

Kohli, N., Hughes, J., Wang, C., Zopluoglu, C., & Davison, M. L. (2015). Fitting a linear–linear piecewise growth mixture model with unknown knots: A comparison of two common approaches to inference. *Psychological methods*, *20(2)*, 259.

Lambert, M. J., & Barley, D. E. (2001). Research summary on the therapeutic relationship and psychotherapy outcome. *Psychotherapy: Theory, research, practice, training*, 38(4), 357.

Larson, L. M., Clark, M. P., Wesely, L. H., Koraleski, S. F., Daniels, J. A., & Smith, P. L. (1999). Videos Versus Role Plays to Increase Counseling Self-Efficacy in Prepractica Trainees. *Counselor Education and Supervision*, *38(4)*, 237-248.

Larson, L.M., Daniels, J.A. (1998). Review of the counseling self-efficacy literature. *The Counseling Psychologist, 26(2),* 179.

Larson, L. M., Suzuki, L. A., Gillespie, K. N., Potenza, M. T., Bechtel, M. A., & Toulouse, A. L. (1992). Development and validation of the counseling self-estimate inventory. *Journal of Counseling Psychology*, *39*(1), 105.

Levine, J. M., & Moreland, R. L. (1990). Progress in small group research. In J. T. Spence (Vol. Ed.), *Annual review of psychology*, *41*, pp. 585- 634. Palo Alto, CA: Annual Reviews.

Lent, R. W., Hoffman, M. A., Hill, C. E., Treistman, D., Mount, M., & Singley, D. (2006). Client-specific counselor self-efficacy in novice counselors: Relation to perceptions of session quality. *Journal of Counseling Psychology*, *53*(4), 453.

Lerner, M.D., McLeod, B.D., & Amori, M.Y. (2013). Preliminary evaluation of an observational measure of group cohesion for group psychotherapy. *Journal of Clinical Psychology,* 69(3), 191-208.

Levy, K.N., Ellison, W.D., Scott, L.N., & Bernecker, S.L. (2011). Attachment Style. *Journal of Clinical Psychology* 67(2), 193-201.

Lewin, K., Kippitt, R., & White, R.K. (1939). Patterns of aggressive behavior in experimentally created social climates. *Journal of Social Psychology, 10,* 271-272.

Little, R.J.A. (1998). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83(404),* 1198-1202.

Lorentzen, S., Sexton, H. C., & Høglend, P. (2004). Therapeutic alliance, cohesion and outcome in a long-term analytic group. A preliminary study. *Nordic Journal of Psychiatry, 58*(1), 33-40.

Lutz, W., Leon, S., Martinovich, Z., Lyons, J.S., & Stiles, W.B. (2007). Therapist effects in outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling Psychology, 54*(1), 32-39.]

MacKenzie, K. R. (1983). The clinical application of a group climate measure. *.Advances in group psychotherapy: Integrating research and practice,* 159-170.

MacKenzie, K. R., & Tschuschke, V. (1993). Relatedness, group work, and outcome in long-term inpatient psychotherapy groups. *The Journal of psychotherapy practice and research, 2*(2), 147.

MacNair-Semands, R. R., Ogrodniczuk, J. S., & Joyce, A. S. (2010). Structure and initial validation of a short form of the Therapeutic Factors Inventory. *International journal of group psychotherapy, 60*(2), 245-281.

Makuch, L.M. (1997). Development and validation of the leadership characteristics inventory. *Proquest Dissertations and Theses.*

Marcus, D. K., Kashy, D. A., & Baldwin, S. A. (2009). Studying psychotherapy using the one-with-many design: The therapeutic alliance as an exemplar. *Journal of Counseling Psychology*, 56(4), 537.

Margolin, G., Oliver, P. H., Gordis, E. B., O'Hearn, H. G., Medina, A. M., Ghosh, C. M., & Morland, L. (1998). The nuts and bolts of behavioral observation of marital and family interaction. *Clinical Child and Family Psychology Review*, *1*(4), 195-213.

Marmarosh, Cheri, Ari Holtz, and Michele Schottenbauer. (2005) Group Cohesiveness, Group-Derived Collective Self-Esteem, Group-Derived Hope, and the Well-Being of Group Therapy Members. *Group Dynamics: Theory, Research, and Practice* 9.1: 32.

McKnight, D.K. (1994). A study of the relationship between the type of leader verbal intervention and cohesion in therapy groups. *Proquest Dissertations & Theses*.

McNeil, D.C. (2005). Patient, therapist, and observer perspectives on cohesion and alliance and their relationship to outcome in psychodynamic group psychotherapy for persons experiencing complicated grief. *Proquest Dissertations & Theses*.

Menard, S. (1995). *Applied logistic regression analysis*: [Sage University Paper Series on Quantitative Applications in the Social Sciences Series No. 07-106]. Thousand Oaks, CA: Sage.

Moreno, J. L. (1932). *Applications of the group method to classification.* New York: National Committee on Prisons and Prison Labor.

Morran, D. K., Stockton, R., & Harris, M. (1991). Analysis of group leader and member feedback messages. *Journal of Group Psychotherapy, Psychodrama, and Sociometry, 43*, 126-135.

Muthén, L.K. and Muthén, B.O. (1998-2012). *Mplus User's Guide (7th Ed.).* Los Angeles, CA: Muthén & Muthén.

Neimeyer, G. J. & Thomas, V. M. (1982). Group structure and group process: Personal construct theory and group development. *Small Group Behavior*, 13, 150-164.

Nickerson, A.B., & Coleman, M.N. (2006). An exploratory study of member attraction, climate, and behavioral outcomes of anger-coping group therapy for children with emotional disturbances. *Small Group Research, 37,* 115.

Norcross, J. C., & Wampold, B. E. (2011). Evidence-based therapy relationships: research conclusions and clinical practices. *Psychotherapy*, *48*(1), 98.

Nuijens, K.L., Teglasi, H., Simcox, A.G., Kivlighan, D.M., & Rothman, L. (2006). The development and validation of the group leader intervention system. *Group Dynamics: Theory, Research, and Practice, 10(2)*, 116-135.

Ogrodniczuk, J. S., Piper, W. E., & Joyce, A. S. (2004). Differences in men's and women's responses to short-term group psychotherapy. *Psychotherapy Research*, *14*(2).

Ormont, L. (1990). The craft of bridging. *International Journal of Group Psychotherapy*, 40(1), 3.

Ormont, L. R. (1997). Bridging in group analysis. *Modern Psychoanalysis*, *22*(1), 59-77.

Peng, C. Y. J., Harwell, M., Liou, S. M., & Ehman, L. H. (2006). Advances in missing data methods and implications for educational research. *Real data analysis*, 31-78.

Raudenbush, S., Bryk, A., Cheong, Y. F., & Congdon, R. (2001*). HLM 5: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.

Raykov, T., Dimitrov, D. M., von Eye, A., & Marcoulides, G. A. (2013). Interrater Agreement Evaluation A Latent Variable Modeling Approach. *Educational and Psychological Measurement*, *73*(3), 512-531.

Rogers, C., (1961). *On Becoming a Person: A Therapist's View of Psychotherapy.* Boston, MA: Mariner Books.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: University Press.

Russell, D., Cutrona, C. E., Rose, J., & Yurko, K. (1984). Social and emotional loneliness: an examination of Weiss's typology of loneliness. *Journal of Personality and Social Psychology*, 46(6), 1313.

Schauenburg, H., Sammet, I., Rabung, S., & Strack, M. (2001). On the differential importance of group experience in inpatient psychotherapy of depressive patients. *Gruppenpsychotherapie und Gruppendynamik*, *37*(4), 349-364.

Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best practices for missing data management in counseling psychology. *Journal of Counseling psychology*, 57(1), 1.

Shapiro, J.L., Peltz, L.S., & Bernadett-Shapiro, S. (1997). *Brief group treatment: Practical training for therapists and counselors.* Farmington Hills, MI: Brooks-Cole Publishing.

Shea, K., & Sedlacek, W. (1997). *Group cohesiveness and client satisfaction in theme-oriented and general counseling groups.* Unpublished manuscript.

Shechtman, Z., & Dvir, V. (2006). Attachment style as a predictor of behavior in group counseling with preadolescents. *Group dynamics: Theory, research, and practice*, *10*(1), 29.

Shechtman, Z., & Rybko, J. (2004). Attachment Style and Observed Initial Self-Disclosure as Explanatory Variables of Group Functioning. *Group Dynamics: Theory, Research, and Practice*, *8*(3), 207.

Sheng, Y., & Sheng, Z. (2012). Is Coefficient Alpha Robust to Non-Normal Data? *Frontiers in Psychology*, 3(34).

Singer, J. D. & Willett, J. B. (2003). *Applied longitudinal data analysis; Modeling change and event occurrence.* New York: Oxford University Press.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*, Sage.

Spybrook, J., Bloom, H., Congdon, R., Hill, C., Martinez, A., Raudenbush, S., & TO, A. (2011). Optimal design plus empirical evidence: Documentation for the "Optimal Design" software. *William T. Grant Foundation. Retrieved on July 14th, 2015.*

Stein, M. B., Pesale, F. P., Slavin, J. M., & Hilsenroth, M. J. (2010). A training outline for conducting psychotherapy process ratings: An example using therapist technique. *Counselling and Psychotherapy Research*, *10*(1), 50-59.

Stice, E., & Bearman, S. K. (2001). Body-image and eating disturbances prospectively predict increases in depressive symptoms in adolescent girls: a growth curve analysis. *Developmental psychology*, *37*(5), 597.

Stiles, W. B., & Snow, J. S. (1984). Dimensions of psychotherapy session impact across sessions and across clients. *British Journal of Clinical Psychology*, *23*(1), 59-63.

Stiles, W.B., Shapiro, D.A., & Firth-Cozens, J. (1990). Correlations of Session Evaluations with Treatment Outcome. *British Journal of Clinical Psychology* 29(1), 13-21.

Stiles, W. B., Reynolds, S., Hardy, G. E., Rees, A., Barkham, M., & Shapiro, D. A. (1994). Evaluation and description of psychotherapy sessions by clients using the Session Evaluation Questionnaire and the Session Impacts Scale. *Journal of Counseling Psychology*, *41*(2), 175.

Streiner, D.L. (2003). Being Inconsistent About Consistency: When Coefficient Alpha Does and Doesn't Matter, *Journal of Personality Assessment*, 80:3, 217-222.

Suhr, D. (2003, March). SUGI 28: *Reliability, exploratory & confirmatory factor analysis for the Scale of Athletic Priorities.* Paper presented at the SAS Users Group International Proceedings, Seattle, WA.

Tabachnick, B., & Fidell, L. (2001). *Using multivariate statistics.* Boston, MA: Allyn & Bacon.

Tasca, G.A., Illing, V., Joyce, A.S., & Ogrodniczuk, J.S. (2009). Three-level multilevel growth models for nested change data: A Guide for group treatment researchers. *Psychotherapy Research Methods, 19(4-5)*, 453-461.

Taube-Schiff, M., Suvak, M. K., Antony, M. M., Bieling, P. J., & McCabe, R. E. (2007). Group cohesion in cognitive-behavioral group therapy for social phobia. *Behaviour Research and Therapy*, *45*(4), 687-698.

Tschuschke, V., & Dies, R. R. (1994). Intensive analysis of therapeutic factors and outcome in long-term inpatient groups. *International Journal of Group Psychotherapy*, 44(2), 185-208.

Tuckman, B. (1965) Developmental Sequence in Small Groups. *Psychological Bulletin*, 63, 384-399.

Tuckman, B. & Jensen, M. (1977) Stages of Small Group Development. *Group and Organizational Studies, 2*, 419-427.

Van Baarsen, B., Snijders, T. A., Smit, J. H., & Van Duijn, M. A. (2001). Lonely but not alone: Emotional isolation and social isolation as two distinct dimensions of loneliness in older people. *Educational and Psychological Measurement*, 61(1), 119-135.

Vocisano, C., Klein, D. N., Arnow, B., Rivera, C., Blalock, J. A., Rothbaum, B., ... & Thase, M. E. (2004). Therapist Variables That Predict Symptom Change in Psychotherapy With Chronically Depressed Outpatients. *Psychotherapy: Theory, Research, Practice, Training, 41(3), 255.*

Wade, N. G., Cornish, M. A., Post, B. C., Tucker, J. R., Worthington, E. L., Jr., Rye, M., & Sandange, S. (2014). Interventions to promote forgiveness: Treatment type, client characteristics, and group processes. *Unpublished manuscript.* Iowa State University, Ames, IA.

Wampold, B. E., & Brown, G. S. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73*, 914–923.

Wei, M., Russell, D. W., Mallinckrodt, B., & Vogel, D. L. (2007). The Experiences in Close Relationship Scale (ECR)-short form: Reliability, validity, and factor structure. *Journal of Personality Assessment*, 88(2), 187-204.

Weiss, R.S. (Ed.). (1973). *Loneliness: The experience of emotional and social isolation.* Cambridge, Mass: MIT Press.

Wilson, P.A., Hansen, N.B., Nalini, T., Neufeld, S., Kochman, A., & Sikkema, K.J. (2008). Scale development of a measure to assess community-based and clinical intervention group environments. *Journal of Community Psychology, 36(3)*, 271-288.

Wittig, J. (2010). Bridging in group therapy using movement improvisation. In Fehr, S. S. (ed.), *101 interventions in group therapy (p.* 83-86). New York, NY: Routledge/Taylor & Francis Group.

Woody, S. R., & Adessky, R. S. (2003). Therapeutic alliance, group cohesion, and homework compliance during cognitive-behavioral group treatment of social phobia. *Behavior Therapy*, *33*(1), 5-27.

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

Yalom, I.D.(2005). *The theory and practice of group psychotherapy*. New York, NY: Basic Books.

APPENDIX A. IRB APPROVAL

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Institutional Review Board
Office for Responsible Research
Vice President for Research
1138 Pearson Hall
Ames, Iowa 50011-2207
515 294-4566
FAX 515 294-4267

DATE: January 28, 2010

TO: Nathaniel Wade
W112 Lagomarcino

FROM: Office for Responsible Research

TITLE: **ISU Group Counseling Study**

IRB ID: **09-202**

Approval Date: 28 January 2010 **Date for Continuing Review: 18 May 2010**

Submission Type: Modification **Review Type**: Full Committee

The project referenced above has received approval from the Institutional Review Board (IRB) at Iowa State University. Please refer to the IRB ID number shown above in all correspondence regarding this study.

Your study has been approved according to the dates shown above. To ensure compliance with federal regulations (45 CFR 46 & 21 CFR 56), please be sure to:

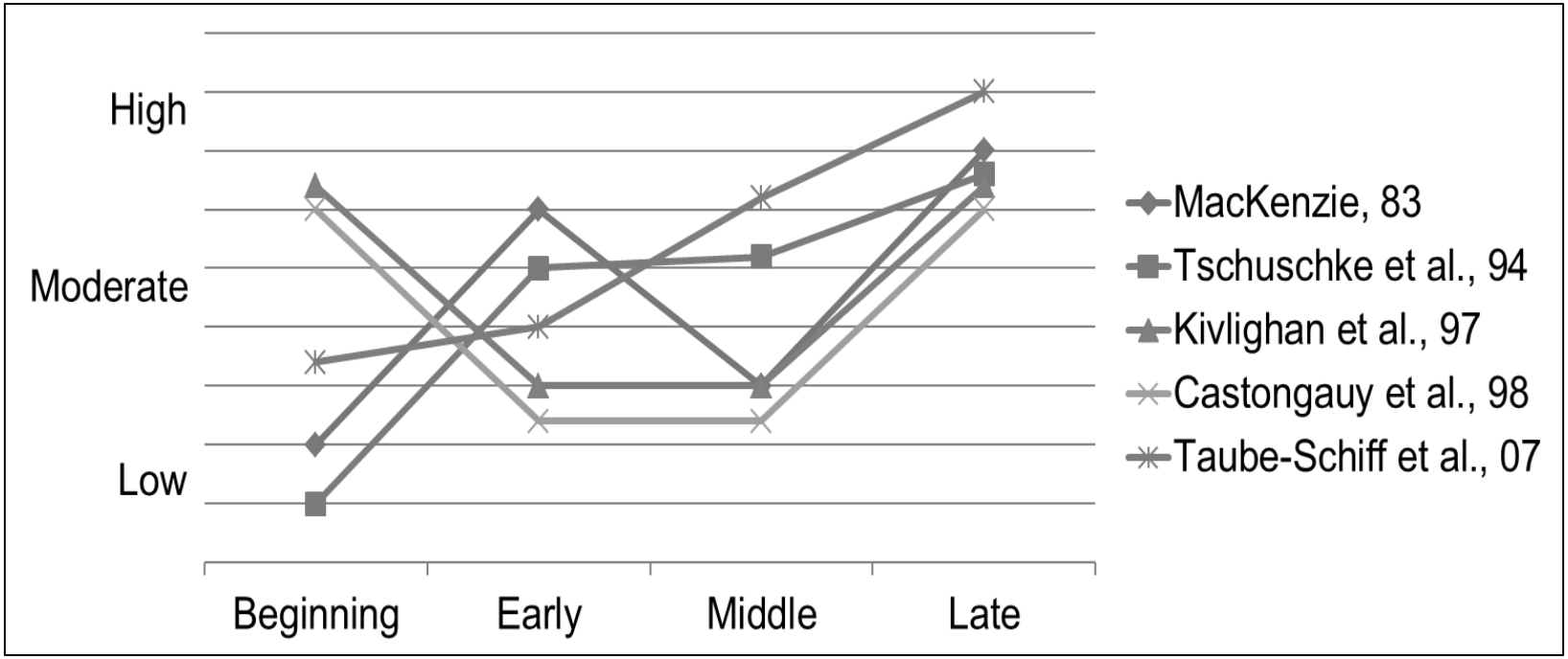
- **Use only the approved study materials** in your research, including the **recruitment materials and informed consent documents that have the IRB approval stamp**.
- **Obtain IRB approval prior to implementing any changes** to the study by submitting the "Continuing Review and/or Modification" form.
- **Immediately inform the IRB of** (1) **all serious and/or unexpected adverse experiences** involving risks to subjects or others; and (2) **any other unanticipated problems involving risks** to subjects or others.
- **Stop all research activity if IRB approval lapses**, unless continuation is necessary to prevent harm to research participants. Research activity can resume once IRB approval is reestablished.
- **Complete a new continuing review form** at least three to four weeks prior to the **date for continuing review** as noted above to provide sufficient time for the IRB to review and approve continuation of the study. We will send a courtesy reminder as this date approaches.

Research investigators are expected to comply with the principles of the Belmont Report, and state and federal regulations regarding the involvement of humans in research. These documents are located on the Office for Responsible Research website [www.compliance.iastate.edu] or available by calling (515) 294-4566.

Upon completion of the project, please submit a Project Closure Form to the Office for Responsible Research, 1138 Pearson Hall, to officially close the project.

ORR 09/09

APPENDIX B. THEORETICAL AND EMPIRICAL FINDINGS ON THE DEVELOPMENT OF COHESION ACROSS TIME



APPENDIX C. CLIENT, LEADER, AND GROUP-LEVEL PREDICTORS EXAMINED IN THE PRESENT INVESTIGATION

	Variable	Proposed link	Representative Citation(s)
Client	Psychological Distress	Low levels of psychological distress predict early termination. High levels of distress in early phases predict increased cohesion in late stages.	Ogrodniczuk, Piper, & Joyce, 2006
	Attachment Style	Attachment anxiety and avoidance lead to unstable group relationships, lower cohesion, early dropout, and poor outcomes.	Schectman & Bybko, 2004; Marmarosh et al., 2006; Huang et al., 2013
	Self-Esteem	Higher self-esteem predicts higher rates of self-disclosure, empathy, and cohesion.	Marmarosh, Holtz, & Schottenbaur, 2005
Leader	Counseling Self-efficacy	Higher counselor self-efficacy leads to greater responsiveness and improved problem-solving; leading to improved cohesion.	Al-Darmaki, 2005; Larson & Daniels, 1998; Lent et al., 2006
	Structuring Behaviors	Establishing a framework for members to understand the group's purpose, boundaries, rules, and roles facilitates a safe working environment for engaging meaningfully in the primary group tasks.	Burlingame, McClendon, & Alonso, 2011a
	Verbal Interactions	Leader interventions that model and facilitate appropriate verbal interactions and style facilitate the creation of cohesive relationships.	Burlingame, McClendon, & Alonso, 2011a
	Facilitating an Emotional Climate	Facilitating clients' emotional expression in a constructive manner allows for openness, authenticity, and a sense of belongingness and acceptance in the group.	Burlingame, McClendon, & Alonso, 2011a
Group	Attendance	Higher attendance indicates higher commitment to the group and its goals.	Burlingame et al, 2011a
	Group Size	Sufficient membership is required for a sense of "we-ness" to emerge	Burlingame et al., 2011a
	Treatment Protocol	Process-oriented groups focus intentionally on fostering interpersonal connections and show more positive development of cohesion across time	Burlingame et al, 2011a

APPENDIX D. LEADER INTERVENTION CATEGORIES

Intervention	Description	Category	Stage(s)
Protecting	Preventing members from taking unnecessary psychological risks such as sharing before safety is established.	Promoting Safety; Emotional Climate	Initial
Drawing Out	Directly inviting comments or involvement from one or more group members	Energizing; Verbal Interaction	Initial
Bridging	Connecting what one group member is saying or doing to the concerns of one or more other members	Energizing; Group Structuring	Initial, Transition
Supporting	Directly reassuring members and encouraging/reinforcing appropriate participation	Promoting Safety; Emotional Climate	Initial, Transition
Self-Disclosure	Revealing own personal feelings, experiences, or here-and-now reactions to group members	Energizing; Verbal Interactions	Initial, transition
Blocking	Cutting-off or intervening when a member is inappropriately probing, gossipy, or invading others' privacy	Promoting Safety; Emotional Climate	Transition
Processing	Facilitating reflection on significant events in group to better understand one's thoughts, feelings, and actions	Energizing; Identification	Working, Ending
Interpreting	Offering possible explanations for members' behaviors or symptoms based on empathic attunement with client.	Energizing; Emotional Climate	Working, Ending
Corrective Feedback	Sharing observations about potentially problematic thoughts, feelings, or behaviors to provide an opportunity for self-reflection and self-appraisal	Energizing; Verbal Interactions	Working, Ending
Positive Feedback	Sharing positive observations of group or individual members to reinforce appropriate behaviors.	Energizing; Verbal Interactions	All
Modeling	Demonstrating skills, attitudes, or other beneficial characteristics such as appropriate self-disclosure, giving/receiving feedback, or openness	Energizing; Verbal Interaction	All

Categories are discussed in Morran, Stockton, & Whittingham, 2004; Chapman et al., 2010; and Ormont, 1990

APPENDIX E. CLIENT AND THERAPIST PRE, MID,AND POST-TREATMENT QUESTIONAIRES

Client.

Your Age: _____ Gender: _____ Race/Ethnicity: _____

Religious affiliation: (Circle the one that best fits your current religious affiliation.)

a. Buddhist
b. Hindu
c. Christian (Catholic)
d. Christian (Protestant)
e. Jewish
f. Muslim
g. None
h. Other: _____

What is your current relationship status? (circle one) Single Married Separated Divorced Widowed

Are you currently receiving counseling or therapy (i.e., meeting regularly with a professional trained to deal with relationship or emotional problems)? YES NO

If so, approximately how many sessions have you had with your current therapist? ____
What are you currently receiving counseling for? (Circle ALL that apply).

a. Anxiety
b. Career/School concerns
c. Concerns with Family
d. Depression
e. Marriage Concerns/Divorce
f. Problems relating to Others
g. Dealing with a Traumatic Event(s)
h. Religious/Spiritual Concerns
i. Other: _____

ECR: The following statements concern how you feel in emotionally intimate relationships. We are interested in how you generally experience relationships, not just in what is happening in your current relationships. Respond to each statement by indicating how much you agree or disagree with it. Mark your answer using the following rating scale:

1=Strongly Disagree 2=Disagree 3=Slightly Disagree 4=Neutral 5=Slightly Agree 6=Agree 7=Strongly Agree

1. It helps to turn to others in times of need.
2. I need a lot of reassurance that I am loved by others.
3. I want to get close to others, but I keep pulling back.
4. I find that other people don't want to get as close as I would like.
5. I turn to others for many things, including comfort and reassurance.

6. My desire to be very close sometimes scares people away.
7. I try to avoid getting too close to others.
8. I do not often worry about being abandoned.
9. I usually discuss my problems and concerns with close others.
10. I get frustrated if other people are not available when I need them.
11. I am nervous when others get too close to me.
12. I worry that others won't care about me as much as I care about them.

GP-CORE: This page has 28 statements about how you have been OVER THE LAST WEEK. Please read each statement and think how often you felt that way last week.

Over the last week …
0=not at all 1=Only Occasionally 2=Sometimes 3=Often 4=Most or all of the time

1. I have felt terribly alone and isolated
2. I have felt tense, anxious, and nervous
3. I have felt that I have someone to turn to for support when needed
4. I have felt O.K. about myself
5. I have felt totally lacking in energy and enthusiasm
6. I have felt able to cope when things go wrong
7. I have been troubled by aches, pains, or other physical problems
8. Talking to people has felt too much for me
9. Tension and anxiety have prevented me from doing important things
10. I have been happy with the things I have done
11. I have been disturbed by unwanted thoughts and feelings
12. I have felt like crying
13. I have felt panic or terror
14. I have felt overwhelmed by my problems
15. I have had difficulty getting to sleep or staying asleep

16. I have felt warmth or affection for someone
17. My problems have been impossible to put to one side
18. I have been able to do most of the things that I needed to
19. I have felt despairing or hopeless
20. I have felt criticized by other people
21. I have thought I have no friends
22. I have felt unhappy
23. Unwanted images or memories have been distressing me
24. I have been irritable when with other people
25. I have thought I am to blame for my problems and difficulties
26. I have felt optimistic about my future
27. I have achieved the things I wanted to
28. I have felt humiliated or shamed by other people

RSE: Please respond to each of the items below by selecting the *one* number that *most closely* describes the extent to which you agree or disagree with the statement.

1=strongly disagree 2=Disagree 3=Agree 4=Strongly Agree

1. On the whole, I am satisfied with myself.
2. At times I think I am no good at all.
3. I feel that I have a number of good qualities.
4. I am able to do things as well as most other people.
5. I feel that I do not have much to be proud of.
6. I certainly feel useless at times.
7. I feel that I am a person of worth, at least on an equal plane with others.
8. I wish I could have more respect for myself.
9. All in all, I am inclined to feel that I am a failure.
10. I take a positive attitude toward myself.

Therapist.

Your Age: _____ Gender: _____ Race/Ethnicity: _____

Do you have a current religious affiliation? YES NO

If so, what is your current affiliation? _____

Are you currently a graduate student? YES NO

Are you currently licensed as a mental health professional? YES NO

What is your professional degree (attained or in progress)?
a. Ph.D.
b. MSW
c. Psy.D.
d. M.A./M.S.
e. Other: _____

How many years have you been conducting individual therapy (including graduate training)? ___

What is your main therapeutic orientation (e.g., humanistic, CBT, etc)?_____

How many years have you conducted therapy groups (including graduate training)? _____

What type of counseling groups are you most experienced with (themed, process, etc.)?____

COSE: Indicate the degree to which you agree or disagree with each statement below by using the following scale:

1=Strongly Disagree 2=Disagree 3=Slightly Disagree
4=Slightly Agree 5=Agree 6=Strongly Agree

1. When using responses like reflection of feeling, active listening, clarification, probing, I am confident I was concise and to the point.
2. I am likely to impose my values on the client during the interview.
3. When I initiate the end of a session I am positive it was in a manner that is not abrupt or brusque and that I will end the session on time.
4. I am confident that I will respond appropriately to the client in view of what the client will express (e.g., my questions was meaningful and not concerned with trivia and minutia).

5. I am certain that my interpretation and confrontation responses was concise and to the point.
6. I am worried that the wording of my responses (e.g., reflection of feeling, clarification, and probing) may be confusing and hard to understand.
7. I feel that I will not be able to respond to the client in a non-judgmental way with respect to the client's values, beliefs, etc.
8. I feel I will respond to the client in an appropriate length of time (neither interrupting the client nor waiting too long to respond).
9. I am worried that the type of response I use at a particular time, i.e., reflection of feeling, interpretation, etc., may not be the appropriate response.
10. I am sure that the content of my responses, i.e., reflection of feeling, clarification, and probing, was consistent with and not discrepant from what the client is saying.
11. I feel confident that I will appear competent and earn the respect of my client.
12. I am confident that my interpretation and confrontation responses were effective in that they were validated by the client's immediate response.
13. I feel confident that I have resolved conflicts in my personal life so that they will not interfere with my counseling abilities.
14. I feel that the content of my interpretation and confrontation responses was consistent with and not discrepant from what the client is saying.
15. I feel that I have enough fundamental knowledge to do effective counseling.
16. I may not be able to maintain the intensity and energy level needed to produce client confidence and active participation.
17. I am confident that the wording of my interpretation and confrontation responses was clear and easy to understand.
18. I am not sure that in a counseling relationship I will express myself in a way that is natural without deliberating over every response or action.
19. I am afraid that I may not understand and properly determine probable meanings of the client's nonverbal behaviors.
20. I am confident that I will know when to use open or closed-ended probes and that these probes will reflect the concerns of the client and not be trivial.
21. My assessments of client problems may not be as accurate as I would like them to be.
22. I am uncertain as to whether I was able to appropriately confront and challenge my

client in therapy.
23. When giving responses, i.e., reflection of feeling, active listening, clarification, probing, I'm afraid that they may not be effective in that they won't be validated by the client's immediate response.
24. I do not feel that I possess a large enough repertoire of techniques to deal with the different problems my clients may present.
25. I feel competent regarding my abilities to deal with crisis situations that may arise during the counseling sessions—e.g., suicide, alcoholism, abuse, etc.
26. I am uncomfortable about dealing with clients who appear unmotivated to work towards mutually determined goals.
27. I may have difficulty dealing with clients who do not verbalize their thoughts during the counseling session.
28. I am unsure as to how to deal with clients who appear noncommittal and indecisive.
29. When working with ethnic minority clients I am confident that I was able to bridge cultural differences in the counseling process.
30. I was an effective counselor with clients of a different social class.
31. I am worried that my interpretation and confrontation responses may not over time assist the client to be more specific in defining and clarifying their problem.
32. I am confident that I was able to conceptualize my client's problems.
33. I am unsure as to how I will lead my client towards the development and selection of concrete goals to work towards.
34. I am confident that I can assess my client's readiness and commitment to change.
35. I feel I may give advice.
36. In working with culturally different clients I may have a difficult time viewing situations from their perspective.
37. I am afraid that I may not be able to effectively relate to someone of lower socioeconomic status than me.

APPENDIX F. CLIENT AND THERAPIST POST-SESSION QUESTIONNAIRES

Client.

GCQ-E: Please circle the number that represents your perceptions of today's group session.

0 = Not at all 3 = Moderately 6 = Extensively

1. The group members liked and cared about each other.	0	1	2	3	4	5	6
2. The members tried to understand why they do the things they do, tried to reason it out.	0	1	2	3	4	5	6
3. The members felt what was happening was important and there was a sense of participation.	0	1	2	3	4	5	6
4. Members challenged & confronted each other in their efforts to sort things out.	0	1	2	3	4	5	6
5. The members revealed sensitive personal information or feelings.	0	1	2	3	4	5	6

SEQ: Please circle the appropriate number to show how you feel about this session.

This session was:								
1.bad	1	2	3	4	5	6	7	good
2.difficult	1	2	3	4	5	6	7	easy
3. valuable	1	2	3	4	5	6	7	worthless
4. shallow	1	2	3	4	5	6	7	deep
5. relaxed	1	2	3	4	5	6	7	tense
6. unpleasant	1	2	3	4	5	6	7	pleasant
7. full	1	2	3	4	5	6	7	empty
8. weak	1	2	3	4	5	6	7	powerful
9. special	1	2	3	4	5	6	7	ordinary

10. rough	1	2	3	4	5	6	7	smooth
11. comfortable	1	2	3	4	5	6	7	uncomfortable

TFI: Please rate the following statements as they apply to your experience in your group by circling the corresponding number, using the following scale:

1= Strongly Disagree 7= Strongly Agree

1. I feel a sense of belonging in this group.
2. It's okay for me to be angry in group.
3. It touches me that people in group are caring toward each other.
4. In group, the members are more alike than different from each other.
5. Even though we have differences, our group feels secure to me.
6. I get to vent my feelings in group.
7. I can "let it all out" in my group.

GES: Please rate the following statements as they apply to your experience in your group by circling the corresponding number, using the following scale:

1= Strongly Disagree 7= Strongly Agree

1. Group members feel a sense of belongingness to the group.
2. Group member feel close to each other.
3. The group is a good place to make friends.
4. Group members show that they care for one another.
5. Group members are committed to the group.
6. Group members can understand what others in the group are going through.
7. Group members are supportive of one another.
8. The atmosphere of the group is a friendly one.
Comments about today's group

Therapist.

GCQ-E: Please circle the number that represents your perceptions of today's group session.

0 = Not at all 3 = Moderately 6 = Extensively

1. The group members liked and cared about each other.	0	1	2	3	4	5	6
2. The members tried to understand why they do the things they do, tried to reason it out.	0	1	2	3	4	5	6
3. The members felt what was happening was important and there was a sense of participation.	0	1	2	3	4	5	6
4. Members challenged & confronted each other in their efforts to sort things out.	0	1	2	3	4	5	6
5. The members revealed sensitive personal information or feelings.	0	1	2	3	4	5	6

SEQ: Please circle the appropriate number to show how you feel about this session.

This session was:								
Bad	1	2	3	4	5	6	7	good
Difficult	1	2	3	4	5	6	7	easy
Valuable	1	2	3	4	5	6	7	worthless
Shallow	1	2	3	4	5	6	7	deep
Relaxed	1	2	3	4	5	6	7	tense
Unpleasant	1	2	3	4	5	6	7	pleasant
Full	1	2	3	4	5	6	7	empty
Weak	1	2	3	4	5	6	7	powerful
Special	1	2	3	4	5	6	7	ordinary
Rough	1	2	3	4	5	6	7	smooth

Comfortable	1	2	3	4	5	6	7	uncomfortable

GES: Please rate the following statements as they apply to your experience in your group by circling the corresponding number, using the following scale:

1= Strongly Disagree to 7= Strongly Agree

1. Group members feel a sense of belongingness to the group.
2. Group member feel close to each other.
3. The group is a good place to make friends.
4. Group members show that they care for one another.
5. Group members are committed to the group.
6. Group members can understand what others in the group are going through.
7. Group members are supportive of one another.
8. The atmosphere of the group is a friendly one.

APPENDIX G. RESEARCH ASSISTANT TRAINING OUTLINE

Week	Meeting focus	Assignment
1	a. Trainees read manual prior to initial meeting. b. Discuss clinical premises for scores/criteria and review questions c. Watch Training Session A together with instructor experienced with the measure. d. Stop tape throughout, have instructor and trainees think/talk out loud about any observed items	Watch and rate training session 1 on own
2	a. Review training session 1 scores in detail b. Watch session together, stop tape and have trainees talk out loud through their scoring decisions.	Watch and rate training session 2 on own
3	Identify and review scoring differences on items that are more than 1 point from one another from sessions 1 & 2.	Begin coding taped sessions (hrs 1 – 30)
5	a. Meet as group to discuss coding progress thus far. Identify inconsistencies and systematic errors b. Focus remediation with trainees on items that seem especially problematic.	Code taped sessions (hrs 30-60)
7	Meet as group to discuss coding progress thus far. Identify inconsistencies and systematic errors	Code taped sessions (hrs 60-90)
9	Meet as group to discuss coding progress thus far. Identify inconsistencies and systematic errors	Code taped sessions (hrs 90-120)
11	Meet as group to discuss coding progress thus far. Identify inconsistencies and systematic errors	Code taped sessions (hrs 120-150)
13	Meet as group to discuss coding progress thus far. Identify inconsistencies and systematic errors	Code taped sessions (hrs 150-180)
15	Meet as group to discuss coding progress thus far. Identify inconsistencies and systematic errors	Code taped sessions (hrs 180-210)
17	Meet as group to discuss coding progress thus far. Identify inconsistencies and systematic errors	Code taped sessions (hrs 210-250)

Note. Table adapted from Stein, M. B., Pesale, F. P., Slavin, J. M., & Hilsenroth, M. J. (2010). A training outline for conducting psychotherapy process ratings: An example using therapist technique. *Counselling and Psychotherapy Research*, *10*(1), 50-59.

APPENDIX H. OBSERVATIONAL CODING MEASURES

Modified GPIRS items

GROUP PSYCHOTHERAPY INTERVENTION RATING SCALE (GPIRS)

Intervention did not occur = 0
Ambiguous - occurred but clarity could be improved = 1
Intervention was performed with clarity = 2

Group Structuring

□ 0 □ 1 □ 2	**1.** Set group agenda (such as discussion topic or group activities); can also be agenda for future sessions (e.g. next week we will…)
□ 0 □ 1 □ 2	**2.** Explicitly describes rationale underlying treatment
□ 0 □ 1 □ 2	**3.** Discussed group rules (such as time, attendance, absences, tardiness, confidentiality, and participation; need to discuss at least one of these. Just mentioning or asserting it is not enough)
□ 0 □ 1 □ 2	**4.** Provide structure that facilitates overall member interaction (not member-member interaction, but member-group)

Verbal Interaction

□ 0 □ 1 □ 2	**5.** Interact in a present – centered (here and now) manner to something occurring explicitly in the group.
□ 0 □ 1 □ 2	**6.** Model informational self-disclosure
□ 0 □ 1 □ 2	**7.** Modeled feeling self-disclosure—has to be a feeling word (sad, angry, joyful, happy, etc).
□ 0 □ 1 □ 2	**8.** Maintained functional control of group when group is "out of control"
□ 0 □ 1 □ 2	**9.** Encouraged self-disclosure relevant to the current group agenda without "forcing it" (must occur "in-the-moment")
□ 0 □ 1 □ 2	**10.** Encouraged present – centered (here and now exchange) vs. story-telling disclosure (past or future)
□ 0 □ 1 □ 2	**11.** Interrupted (or redirected) ill-timed or excessive member disclosure that dominates group
□ 0 □ 1 □ 2	**12.** Leader shared brief personal experience relevant to group agenda (without being judgmental or overly-intellectual)
□ 0 □ 1 □ 2	**19.** Leader discusses fears/concerns about participating in group
□ 0 □ 1 □ 2	**20.** Explicitly discussed roles and responsibilities beyond setting agenda
□ 0 □ 1 □ 2	**21.** Modelled member-member behavior
□ 0 □ 1 □ 2	**22.** Facilitated member-member interaction (not member-leader or member-group)
□ 0 □ 1 □ 2	**23.** Educated members on the value of self-disclosure
□ 0 □ 1 □ 2	**24.** Elicited member – member feeling disclosure
□ 0 □ 1 □ 2	**25.** Reframed injurious member-member feedback (interrupting, if necessary)
□ 0 □ 1 □ 2	**26.** Restated corrective feedback by member
□ 0 □ 1 □ 2	**27.** Used consensus to reinforce m-m or m-l feedback that is present in the

	group
□ 0 □ 1 □ 2	**28.** Attended to balance between positive and corrective member-member feedback
□ 0 □ 1 □ 2	**29.** Encouraged positive member-member feedback
□ 0 □ 1 □ 2	**30.** Therapist helped members apply in-group member-member feedback to out-of-group situations

Emotional Climate

□ 0 □ 1 □ 2	**13.** Leader was not defensive when they made a mistake or an intervention failed or when confronted by a member
□ 0 □ 1 □ 2	**14.** Maintained an active style of engagement throughout the section with the group and its work (active engagement can be entirely nonverbal)
□ 0 □ 1 □ 2	**15.** Used nonjudgmental language/non-verbal behavior with members
□ 0 □ 1 □ 2	**16.** Modeled expressions of open and genuine warmth to all members who share throughout section
□ 0 □ 1 □ 2	**17.** Recognized and responded to the *meaning* of groups members' comments
□ 0 □ 1 □ 2	**18.** Stopped attacking and judgmental situations and expressions between members
□ 0 □ 1 □ 2	**31.** Helped members recognize why they feel a certain way (identifying underlying concerns or motives)
□ 0 □ 1 □ 2	**32.** Encouraged active emotional engagement between group members
□ 0 □ 1 □ 2	**33.** Fostered a climate of supportive challenge—challenge is intended to push a member to behave differently either in/out of the group.
□ 0 □ 1 □ 2	**34.** Assisted members in describing their emotions
□ 0 □ 1 □ 2	**35.** After member sharing, elicited verbal expressions that connect members to one another
□ 0 □ 1 □ 2	**36.** Refrained from conveying personal feelings of hostility and anger in response to disruptive member behavior

APPENDIX I. SCORING MANUAL

Group Leader Interventions and Cohesion Study: Research Assistant Coding Manual

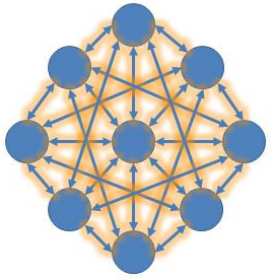


Table of Contents

Introduction

This scoring manual is designed to provide research assistants with a comprehensive system for scoring recorded psychotherapy group sessions as a part of the Group Leader Interventions and Cohesion study. Along with in-person training in behavioral coding, this manual can help in clarifying questions of interpretation for individual items used in this study. It is recommended that this manual serve as a first resource when there is uncertainty regarding the meaning of a particular question or how best to score what one has observed. If at any time this manual is unable to address a given concern, please do not hesitate to contact the principal investigator of this study, Jeritt Tucker, M.S., by e-mail at jrtucker@iastate.edu, or by phone at 515.291.5393.

Below is a brief overview of the variables of interest in the present study. Following this are general scoring guidelines including scoring strategies to help you acquire and maintain coding reliability across measures and coding sessions. Finally, you were presented with detailed information about coding of both of the scales used in the study; the Group Psychotherapy Intervention Rating Scale (GPIRS) and the Group Climate Questionnaire – Engagement subscale (GCQ-E; MacKenzie, 1983).

The Study Variables: Leader Behaviors and Cohesion

There are two primary observational components of this study; *leader interventions* and *group cohesion*. *Leader interventions* refer to any purposeful actions of a leader to ensure safety and/or to initiate, energize, or enhance the therapeutic efficacy of a counseling group (Morran, Stockton, & Whittingham, 2004). Thus, leader interventions typically focus on promoting altruism, reality testing, interpersonal learning, belongingness, and the expression of emotions.

Three dimensions or "types" of leader behaviors have emerged from recent empirical research: G*roup Structuring, Verbal Interactions,* and *Creating an Emotional Climate* (Burlingame, Fuhriman, & Johnson, 2002). *Group Structuring* refers to any verbalized intervention meant to establish rules, norms, and roles; address fears of self-disclosure; and implement structured group exercises. *Verbal Interactions* measure the leader's attempts to model and reinforce appropriate member relations, provide feedback, and facilitate members sharing feedback with one another. This might be seen as the "here-and-now" work of helping members share openly about their present-moment thoughts and feelings. Finally, *Creating an Emotional Climate* refers to attempts to create a safe environment for self-exploration by directly supporting or protecting members. In the present study, the Group Psychotherapy Intervention Rating Scale (GPIRS; Chapman, Baker, Porter, Thayer, & Burlingame, 2010) is used as an observer-rated measure of the presence and quality of these three categories of leader interventions.

The second component of this study, *cohesion*, is a broad term with a rich research history. It can best be defined as the quality of relationships between group members and leaders. Cohesion has both a *structural* dimension as well as a *qualitative* dimension (Burlingame, McClendon, & Alonso, 2011). The structural component of cohesion refers to the direction of the relationship. That is,

is this a relationship between specific members, between a particular member and a leader, or between the member and the whole group? The second component of cohesion, and the one we are primarily interested in this study, is the quality of these relationships. Studies on the qualitative aspects of cohesion have revealed three primary components (called 'factors'): *Bonding*, *Working*, and *Negative relationship factors* (Johnson, Burlingame, Olsen, Davies, & Gleave, 2005). *Bonding* refers to how much fondness, closeness, and belonging exists in the relationships in the group. *Working* refers to how unified members are in accomplishing therapeutic goals, such as understanding the origins of their concerns and increasing their psychological well-being. *Negative relationship factors* refer to the presence of conflict and a lack of empathy between both leaders and group members. For this study, the GCQ-E (MacKenzie, 1983) is used as an observational measure of group cohesion.

In summary, the two variables of interest in this study can be modeled as follows:

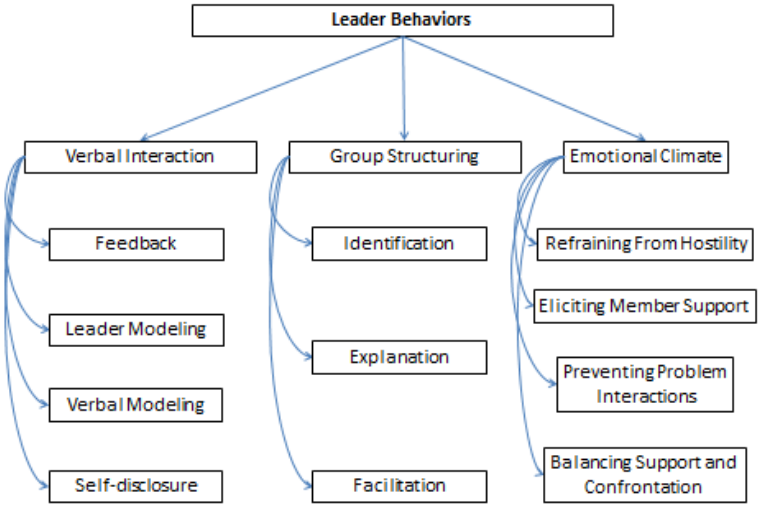


Figure 1. Structural Model of Leadership Behaviors (Chapman, Baker, Porter, Thayer, & Burlingame, 2010)

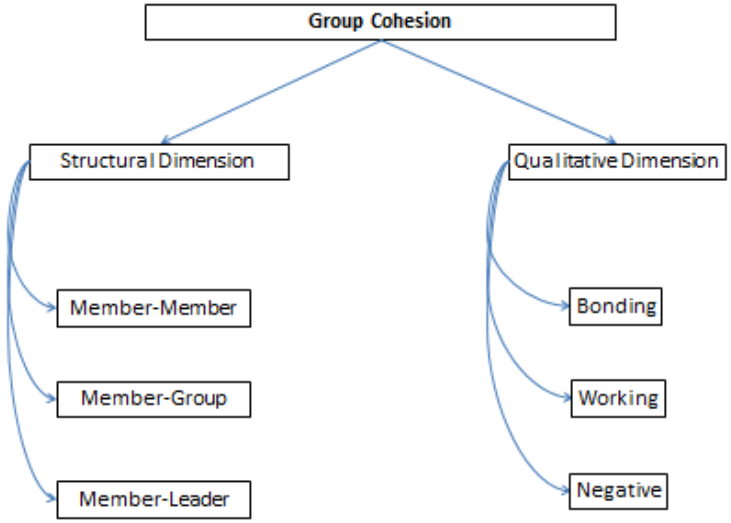


Figure 2. Structural Model of Group Cohesion (Johnson, Burlingame, Olsen, Davies, & Gleave, 2005)

When coding client and group leader behaviors, you were operating within the two above structures to identify specific interventions or facets of cohesion. This will help us in understanding which specific interventions relate to cohesion and ultimately to positive client outcomes.

General Scoring Guidelines

This section provides an overview of scoring guidelines intended to help research assistants score therapy sessions in an efficient, standardized, and reliable manner. Please read this section carefully.

Scoring Strategies

Scoring "Is" not "Oughts": All scoring focuses upon leader or member behavior. Thus, coders should only score what a leader or a client actually does in session, not what might have been done or should have been done. Here is a brief summary of important guidelines for rating "is" and not "ought":

(a) Code only client/leader behaviors.
(b) Rate only what a client or leader does, not what you believe they should have done, and not what you believe they intended to do.
(c) Never assume or guess what a client or leader might be thinking. If there is no behavioral evidence, in the form of something the individual says or does, then do not give the corresponding item a positive score.

Being Thorough: Carefully read each scale item every time an item is scored. Although you may have become familiar with the scales after using it for a time, reading the item over again will make sure the full content is considered in formulating a final decision. When coding, always have this manual present and refer to it whenever there is any confusion about scoring an item. This will also help to prevent "halo" effects (described below).

Because coding tapes is a demanding and work-intensive process, please do not do other tasks when scoring. If you need a break, which you most certainly will, feel free to take a moment away from the tapes, but please do not try and multi-task while coding.

Coder Caveats

Avoiding Halo Effects: Coders should be careful to avoid instances of "halo" effects. Halo effects refer to situations where the scoring for one item is biased or influenced by the scoring awarded to another item, or by a global judgment about the whole session. Halo effects come in many forms; here are some relevant examples:

(a) A coder decides s/he really likes the leader of this particular group. As a result, the coder tends to give high scores on every item.
(b) A coder is particularly impressed with a specific therapeutic segment. As a result, the coder gives high scores too many items, even those they did not directly observe.
(c) When coding cohesion, a coder observes early on that, if the session were stopped, the overall cohesion of the group would be low. Having formed a negative opinion, the coder does not give sufficient weight to behaviors that appear later in the session. The coder therefore gives low scores for most items.
(d) A coder decides s/he really dislikes a particular leader. As a result, the coder tends to give low scores on every item.
(e) A coder intentionally decides or unintentionally acts as though two different items naturally go together.

To avoid halo effects coders have to follow the consistent criteria provided by this manual. Coders must score each item as a separate, independent entity that is not influenced by other items.

Essentially, coders should treat each GPIRS and GCQ-E item as if it is completely uncorrelated with every other item, even if that item appears to have similar characteristics.

Call'em Like you See'em: Please remember that not every aspect of a leader's intervention or group cohesion can be scored. The GPIRS and GCQ-E are not exhaustive lists of all dimensions of these variables. Coders should therefore not stretch the assessment of client or leader behavior just so it will fit into one of the items (even if it seems like a particularly potent therapeutic moment). When client/leader behavior is forced to fit certain items (or vice-versa), coder reliability is severely compromised.

Coding using the GPIRS

The GPIRS interventions are organized into three categories of interventions. The first category—*Structuring*—relates to how a group leader sets up a group so that all the members know what the group session is about (agenda) and how it relates to their recovery (treatment rationale). This category also deals with setting the boundaries of the group (rules) so that the group is a safe and predictable container for treatment. Finally, it captures whether a group leader attempts to structure activities to draw in members of the group (interaction). All of the groups in the present study should have a clear purpose, general goals and session's objectives to achieve these goals. The degree to which the group leader conveys the purpose, goals and objectives to group members is partially captured by the *Structuring* dimension of the GPIRS.

The second category of GPIRS interventions is the *Verbal Interaction* category. This category captures both positive and negative influences on the group. Five items on this scale capture leader attempts to model behaviors that have been linked to positive outcomes. This includes present-centeredness instead of "story-telling" as well as well-timed information giving and the disclosure of

feelings. There are two verbal interaction items that address the leader proper and have been linked to negative processes and outcomes in group. For instance, one item captures how well the leader is managing the group and keeping it on track, especially when there are disruptive moments or members. Another item captures leader interventions directed toward members who dominate the group or are engaged in ill-timed self-disclosure.

The last category addresses how the leader is facilitating the *Emotional Climate* in the group. *Emotional Climate* is highly correlated with successful groups and good member outcomes. Like the verbal interaction category, it captures both positive and negative influences. Three items relate to how the leader engages group members, expresses an open and warm relationship, and reflects the meaning of member comments. Three items also reflect potential negative influences on the group, including how the leader responds to challenges, relates to members in a nonjudgmental manner, and handles conflict or aggression in the group.

In the present investigation, you are being asked to rate each session for both presence and quality of specific types of leader behaviors. These are broken down into the above three categories and captured by 36 possible intervention types. In rating, the hope is that you was familiar enough with each intervention type to mark all that apply during the course of a session. The rating scale is from 0 to 2; where 0 represents the intervention type not occurring, 1 represents a poorly or ambiguously executed intervention, and 2 represents a well-executed or clear intervention. Descriptions and example behaviors for each item and each rating level are provided.

A leader "speaking turn" is defined as the time between when a leader starts speaking until another member speaks. In coding, you will mark the time on the tape when the leader begins speaking. This was in the form of 'hh:mm:ss'. For example, '01:22:13' equates to 1 hour, 22 minutes, and 13 seconds into the tape. You will also note when the leader stops speaking so that we can compute data regarding the length of each speaking turn and the overall amount of time a leader is speaking in a given group.

What follows is a brief description of what sort of leader behaviors would fall under each intervention category.

Structuring

1. Set group agenda (such as discussion topic or group activity); can also be agenda for future sessions (e.g. next week we will…)

Level 0 Proposed Anchors

No group agenda set.

Level 1 Proposed Anchors

Mentions initial step only: "Let's start with a check in" or "Welcome to the group" with no further discussion of agenda.

"Let's talk about our concerns." "We are going to work on forgiveness" And then no further guidance.

Level 2 Proposed Anchors

Outlines agenda for group (e.g. "Let's check in first, then we will discuss what is bringing us to group").

"Today we are going to..." (then lists focus of group and talk about it) "We are going to work on forgiveness and will be doing role plays"

2. Explicitly describes rationale underlying treatment

Level 0 Proposed Anchors

Rationale (i.e. reason the agenda is this way) for treatment not mentioned or discussed at all.

Level 1 Proposed Anchors

Generic/surface level reason given for why group agenda is this way:

"Forgiveness is good for your mental health." "Talking about this will help get you ready to end our group and say goodbye"

Level 2 Proposed Anchors

Leader gives deeper "reason why", outlines the *purpose* of the agenda "Saying exactly how you were hurt helps to identify what values or needs were undermined so you can…" or "Working on understanding your own tendencies in group helps to improve our self-awareness and our relationships…"

3. Discussed group rules (such as time, attendance, absences, tardiness, confidentiality, and participation; need to discuss at least one of these. Just mentioning or asserting it is not enough)

Level 0 Proposed Anchors

No mention of group rules or no group rules established.

Level 1 Proposed Anchors

"You know the rules; we've already talked about that!"

Level 2 Proposed Anchors

"Remember our rule about confidentiality, so what you say about your feelings will stay here in the group to make members feel safe."

"Attendance is important because it communicates members' dedication to group and helps build safety"

4. Provide structure that facilitates overall member interaction (not member-member interaction, but member-group).

Level 0 Proposed anchor

Leader does not provide structure for group interaction; shares without asking for member sharing.

Level 1 Proposed Anchors

Leader loosely structures member interactions "I wonder who would like to begin first today?"

Level 2 Proposed Anchors
Leader uses a structured exercise/activity that encourages member interaction. "Ok fill out your paper and then let's discuss your answers in pairs. Who would be willing to share first?" Leader facilitates interactive discussion.

19. Leader discusses fears/concerns about participating in group

Level 0 Proposed Anchors
Not addressed or a member expressed concern about participating and it was not acknowledged by the leader.
Level 1 Proposed Anchors
There is an acknowledgement of fear/concern by the leader without deeper discussion. "Thanks for expressing your concern (then moves on)."
Level 2 Proposed Anchors
A response to a specific concern raised by a group member or ongoing acknowledgement of concerns or emotions that arise with support.
"Thank you for expressing your concern" and then discusses it... "Sometimes it can be scary to talk in group but this is okay because it may be important…

20. Explicitly discussed roles and responsibilities beyond setting agenda

Roles and responsibilities refer to expectations of leaders and members. These can be broad (e.g. leader role to create safety and trust) or specific (e.g. one member's role in a given group session to role-play another member's father). Roles must be explicitly discussed (as in "you are expected to" or "it will be your role to.."), and cannot be implied only.
Level 0 Proposed Anchors
No explicit description of roles or responsibilities. Questions about roles/responsibilities ignored. "We have Susan with us today and she will be a silent observer." (here does not discuss role of silent observer).
Level 1 Proposed Anchors
Explicit description of either a member or leader role.
"Now, Larry, remember your role in here." "That's not your responsibility."
"It's my role in here to make sure members feel safe and yet challenged"
Level 2 Proposed Anchors
Explicit description of both member and leader roles during this section.
Leader starts by saying "I will… or I'm going to… and then you will have a chance to…"
"In group today I will bring in this information – and I'd like the group to give feedback – how do you feel about that?"

Verbal Interaction

5. Interact in a present – centered (here and now) manner to something occurring explicitly in the group.

Level 0 Proposed Anchors
Focus on the past: "Tell me about that time in your life? "I want to tell you a story"
Level 1 Proposed Anchors
Leader reacts to something occurring in the group but does not facilitate an interaction or provides little detail. Leader says "Mark when you were talking about the side-effects of your medication, I noticed the group really listened."

Level 2 Proposed Anchors
Relates to a leader interacting with a member in the "here and now", by helping members express what is going on for them moment-to-moment.
Leader says "John, I noticed that you began to tear up when you were talking about the side-effects you were having with your medication. Can you tell the group what is going on for you right now?"

6. Modeled informational self-disclosure
Informational self-disclosure refers to hard facts about the leader. For example "I have a Ph.D. in counseling psychology and work primarily with
college students."
Level 0 Proposed Anchors
Informational self-disclosure did not occur or too much, inappropriate, or only irrelevant information was given. E.G. Leader says "I played golf the other day as well, where did you go?"
Level 1 Proposed Anchors
Some sharing without any detail "I've even experienced that." (no information provided)
Level 2 Proposed Anchors
Personal example related to topic and included members– "We've been discussing spirituality, and I wanted to let you know my own personal background in the area…"

7. Modeled feeling self-disclosure—has to be a feeling word (sad, angry, joyful, happy, etc).
Feeling self-disclosure refers to the leader sharing their past or present-moment feelings.
Level 0 Proposed Anchors
Feeling disclosure did not occur. Uses non-specific feeling word (e.g. rushed)
Level 1 Proposed Anchors
Can be general or indirect; uses a feeling word but does not model or follow up. "Sometimes I feel angry."
Level 2 Proposed Anchors
Acknowledges feeling with an explicit feeling word and includes a modeling component
""When I have been in situations similar to yours, I felt scared because I didn't know how things would turn out for me."
"When you said that, I felt…"

8. Maintained functional control of group when group is "out of control"
"Out of control" could include times when the group is off topic, dealing with a member outburst that is not related to the group theme, member-member or member-leader conflict. In short, it is when the content and/or emotional group processes stray from the structure.
Level 0 Proposed Anchors
There is a need for functional control interventions but leader fails to do so or attempts to redirect inappropriately "Knock it off!" " I need everyone to control themselves." Alternatively, the group is functioning fine.
Level 1 Proposed Anchors
Leader addresses "out of control" group but does not steer group clearly back or does not acknowledge group is out of control "Let's get back to the manual".

Level 2 Proposed Anchors
there is a need for functional control and leader redirects effectively to regain control of group
When this condition is met the group leader first acknowledges that the group is out of control and then makes efforts to bring it back on track. "I noticed that the group has started to become a little cognitive here, I wonder if we can push ourselves to get back in touch with our emotions".

9. Encouraged self-disclosure relevant to the current group agenda without "forcing it" (must occur "in-the-moment")
This must be an explicit encouragement for a member or members to share information about themselves (can be information or feeling). "Ken, can you tell us more about what you are feeling right now?" This is not just support after a member has shared or an open invitation to share.
Level 0 Proposed Anchors
Encouragement of self-disclosure does not occur.
Level 1 Proposed Anchors
A style of encouraging participation that appears disjointed or does not flow. For example, turn-taking without the ability to pass: "Kevin, it's your turn to participate"
Level 2 Proposed Anchors
An open inviting style to participate—"What are some other feelings about…that members are noticing?" "I noticed Ed and Linda were talking about goals. would other members like to share their goals with the group?"

10. Encouraged present – centered (here and now exchange) vs. story-telling disclosure (past or future)
Level 0 Proposed Anchors
Does not interrupt members who are telling long stories or there is no need to interrupt because story-telling is not present
Level 1 Proposed Anchors
Interruption without redirection "Mark, stop telling that story."
Level 2 Proposed Anchors
Early interruption of story with redirection to present. "Lonnie, that's an interesting story, this is how I see it relating to what we're talking about…"

11. Interrupted (or redirected) ill-timed or excessive member disclosure that dominates group
Requires judgment that a member is sharing excessively. This can be indicated when a member interrupts others frequently or continues to share when other members interject or begin to share.
Level 0 Proposed Anchor
Does not interrupt member, takes passive role, yawns, says disrespectful things ("Too much information").
Level 1 Proposed Anchor
Abrupt redirection or interruption: "Let's not talk about that right now"
Level 2 Proposed Anchor
Directly intervenes in a respectful way. "Let me stop you from going on, I think that what you're talking about is beyond the scope of our group and I'd be happy to visit with you after group about…"

12. Leader shared brief personal experience relevant to group agenda (without being judgmental or overly-intellectual)
Level 0 Proposed Anchors

No relevant personal experience related: "This weekend I got stuck in traffic too..." Or, personal experience takes up too much time so that members don't have an opportunity to interact and share

Level 1 Proposed Anchors

Relates personal information without relating it to the group agenda "This weekend I was in traffic and noticed myself getting very angry and resentful."

Level 2 Proposed Anchors

Tells brief personal story and then relates to topic "this weekend I was stuck in traffic and noticed the feeling of getting very resentful toward the persons ahead of me, my face got really warm and I noticed feeling tightness in my body. What do others notice in their bodies when they feel resentful or angry?"

21. Leader models member-member behavior

Member-member behaviors include direct interactions between members (asking questions, responding to one another, etc.). Modeling this behavior means providing education on how one member might share directly with another.

Level 0 Proposed Anchors

Modelling did not occur. "When Taylor asked you that question, what would have been a more helpful response?" (this does not model what a more effective response is).

Level 1 Proposed Anchors

Models member-member behavior in a way that interrupts the flow of group or does not help member to experiment "When Taylor asked you that question, a more effective response would have been…" with no follow-up

Level 2 Proposed Anchors

Models member-member behavior in a way that is flexible, educative, and supportive. "When Taylor asked you that question, I wonder if a more effective response might have been to... Let's try that now…"

22. Facilitate member-member interaction (not member-leader or member-group)

Level 0 Proposed Anchors

Leaders encourage members to only talk to leader or share with the group as a whole. "What do you all think?"

Level 1 Proposed Anchors

Provide a turn-taking task that requires member to member interaction. "Ted, what do you think of Sharon's picture?"

Level 2 Proposed Anchors

Helps direct one member to share directly with or ask questions of another member. "Rosalee, can you share with Tom what you just told me?"

23. Educated members on the value of self-disclosure

Level 0 Proposed Anchors

Does not occur; sarcasm directed toward self-disclosure.

Level 1 Proposed Anchors— provides a rationale for why they are encouraging disclosure

Encourages self-disclosure generically but does not explain or provide education. "It helps to talk about these kinds of things"

Level 2 Proposed Anchors—models with a member how disclosure can lead to resolution

Explicitly educates on the value of self-disclosure. Takes steps to implement it in the group. "Although we sometimes feel like we are all alone in our concerns, often by sharing our story we actually help others who are quietly suffering the same things. Who would be willing to make

the first step?"

24. Elicited member–member feeling disclosure

Level 0 Proposed Anchors

Does nothing to elicit member feelings; or invalidates feeling disclosure.

Level 1 Proposed Anchors

Does not elicit but describes: "does that make you feel sad?" "Any other feelings?"

Level 2 Proposed Anchors

Elicits and facilitates interaction: "So Melissa, Can you tell Amanda directly what emotion comes up in you when you hear her say that?" "How do you feel when you hear that?" "Does anyone feel similar to Doug with respect to…."

25. Reframed injurious member-member feedback (interrupting, if necessary)

Level 0 Proposed Anchors

No injurious feedback occurs or ignores injurious feedback—"well, moving along…"

Level 1 Proposed Anchors

Reframes injurious feedback in a way that does not flow or is judgmental "Hold on a second Scott, that language is not acceptable in group."

Level 2 Proposed Anchors

"Hold on a second Alan, when I heard you say Becky is incompetent, I wonder if your underlying message could have been… is that accurate?"

26. Restated corrective member-member feedback

Level 0 Proposed Anchors

Stays silent or corrective member-member feedback does not occur.

Level 1 Proposed Anchors

Restates corrective feedback in a way that does not check for accuracy: "It sounds like Tom doesn't want you to say that word to her."

Level 2 Proposed Anchors

Restates and then checks in for accuracy. "It sounds like what Natalie is saying is that she doesn't want you use that word because it reminds her of something her father used to say…. Is that accurate Natalie?"

27. Used consensus to reinforce member-member or leader-member feedback that is present in the group

Level 0 Proposed Anchors

There is no member-member or leader-member feedback that is given or stays silent when feedback is given.

Level 1 Proposed Anchors

Leader asks directly for consensus without restating or opening discussion up "Does everyone agree?" in response to m/m or m/l feedback.

Level 2 Proposed Anchors

Clarifies feedback, uses consensus, and opens up to group. "What I hear several of you saying is that it is important to talk with your doctor before stopping your medications. Are there any other opinions?"

28. Attended to balance between positive and corrective member-member feedback

Level 0 Proposed Anchors
No feedback given. Or no opportunity for corrective feedback.
Level 1 Proposed Anchors
Just corrective or positive feedback when there is clearly opportunity for both.
Level 2 Proposed Anchors
Balances positive and negative feedback during the group. "I really like how you….what would have happened if you would have…Good job!" (although may not be spoken aloud in one complete sentence as seen above)

29. Encouraged positive member-member feedback

Level 0 Proposed Anchors
Giving feedback without encouraging others.
Level 1 Proposed Anchors
Encouraged feedback but did not facilitate processing. "I like the way that Donald did…don't you guys?" or can include non-verbal feedback such as smiling
Level 2 Proposed Anchors
Provides encouragement and then engages the group. "I noticed that the interaction between Donald and Melissa was very authentic. What did you all notice about how Donald shared?"

30. Therapist helped members apply in-group member-member feedback to out-of-group situation

Level 0 Proposed Anchors
No member-member feedback or application; "that's something we only talk about in group."
Level 1 Proposed Anchors
Focuses on applying to out of group but instructs members only. "You should do what Kathy said on your next visit home."
Level 2 Proposed Anchors
Encourages out-of-group behavior and facilitates processing. "Mark, can you think of any examples outside of our group that relate to the feedback that Rick just shared with you" "Doug, I notice that you're angry with Susan. Have you had similar feelings towards others outside our group?" "So how can what you just heard from Kathy help you at home this week when you interact with your family?"

Emotional Climate

13. Leader was not defensive when they made a mistake or an intervention failed or when confronted by a member

Level 0 Proposed Anchors
leader defensive or blames others. "I wouldn't have said that if you guys hadn't acted that way." "I'm not the one in need of therapy here." "Oh boy, I'm rambling again". Alternatively, no mistake was clearly made or there was no opportunity for this intervention (e.g. leader made mistake but said nothing afterward)
Level 1 Proposed Anchors
Neutral acknowledgement of failed intervention or member confrontation: "You're right."
Level 2 Proposed Anchors

Non-defensive acknowledgement as they address the issue: "You're right; there may be a more effective way for me to handle this."

14. Maintained an active style of engagement throughout the section with the group and its work (active engagement can be entirely nonverbal)

Level 0 Proposed Anchors

Playing with phone the whole time, yawning, reading manual directly or session notes.

Level 1 Proposed Anchors

Established eye contact and talked with select group members who shared.

Level 2 Proposed Anchors

Established eye contact and talked to each (most) group member who shared, watching, listening, and responding throughout section

15. Used nonjudgmental language or non-verbals with members

Level 0 Proposed Anchors

Does not speak during section, or, if does, uses judgmental language or negative evaluative statements: "You're wrong." "You're always late." Behaviors may include rolling eyes, sighing at group member comment/behavior.

Level 1 Proposed Anchors

Describes behavior with a label or implicit judgment. Comments that have a "hint" of condescension "t"

Level 2 Proposed Anchors

Describes behavior without a label, stating things in a factual, non-emotional manner, "How did that affect you?" "What did you do as a result of that?"

16. Modeled expressions of open and genuine warmth to all members who share throughout section

Level 0 Proposed Anchors

Little acknowledgment of members, focuses primarily on content of group or session manual.

Level 1 Proposed Anchors

This is a group-level or global expression of warmth without targeting specific members. "I appreciate you all staying with me even though we're going a little long"

Level 2 Proposed Anchors

Expresses warmth to all members who share. Open, friendly comments such as "John, I really appreciate how you continue to make contributions to each exercise that we do in the group"

17. Recognized and responded to the *meaning* of group member's comments

Level 0 Proposed Anchors

No acknowledgement or acknowledgement of statement without identifying meaning "Thanks for sharing"

Level 1 Proposed Anchors

Weak acknowledgement of meaning of member comment: "How interesting." "How exciting!"

Level 2 Proposed Anchors

Verified or restated deeper meaning of member response, "What I hear you saying is… "Let me make sure I'm on the same page with you." "It sounds like you were really miserable"

18. Stopped attacking or judgmental situations and expressions between members

Level 0 Proposed Anchors

Allowed attacking to continue. (If no opportunity for this intervention occurred, mark 0)

Level 1 Proposed Anchors

group leader tells member to stop. "Cut it out you guys." "Stop that right now"

Level 2 Proposed Anchors

group leader stops attack/judgment and redirects

Stops action and invites members to try something new "I'd like to pause the action right now. What do we observe is going on in group?"

31. Helped members recognize why they feel a certain way (identifying underlying concerns or motives)

Level 0 Proposed Anchors

Ignoring feelings, does not reflect feelings. Or if no feelings present or if nonspecific feeling words are used (nonspecific = busy, hard, tired)

Level 1 Proposed Anchors

Assist member with directly acknowledging feelings (used specific feeling words such as sad, angry, scary, etc.)

"It seems this group is making you angry."

Level 2 Proposed Anchors

not only directly acknowledges feeling but also deepens affect and helps member understand "why"

"Thanks for sharing that concern, any thoughts about why you feel that way?"

"When you said that you felt angry at Lee, was that related to his earlier comment?"

32. Encouraged active emotional engagement between group members

Level 0 Proposed Anchors

Keeping exchanges of information focused, or lecture style, or avoiding emotional engagement between members.

Level 1 Proposed Anchors

Minimal encouragement or passive engagement between members "Gary, how do you feel about what Doug just said?"

Level 2 Proposed Anchors

Specific, "I really appreciate how you helped Doug share his thoughts, Dave; I wonder if you can push yourself to share similarly with the group"

33. Fostered a climate of supportive challenge—challenge is intended to push a member to behave differently either in/out of the group.

Level 0 Proposed Anchors

Passive acceptance of a client's behavior, or no clear opportunity for challenging feedback

Level 1 Proposed Anchors—challenge without support or support without challenge

Attempts to balance support/challenge but errs on one-side "Russell, you seem to be having a hard time talking today" (support only)

Level 2 Proposed Anchors

Effectively balances support and then challenge: "Russell, as we've talked about, you seem to be having a difficult time speaking up today. How can you take this opportunity to reach out to group rather than withdraw inward?"

34. Assisted members in describing their emotions

Level 0 Proposed Anchors

Does not facilitate member's describing their emotions.

Level 1 Proposed Anchors

Leader interprets/describes member feelings without checking in: "You're afraid." "Overwhelmed, huh? That must be a big burden to carry, like a heavy weight"

Level 2 Proposed Anchors—

Leader identifies feelings and then helps member elaborate further. "I wonder if I hear anger in what you're saying? Can you talk a little about it?" "Share with us what it's like to be so frustrated."

35. After member sharing, elicited verbal expressions that connect members to one another

Level 0 Proposed Anchors

Not addressed, leader-centric, lecture type, focused on content only.

Level 1 Proposed Anchors

Checks in with group after member shares "Anyone else feel that way?" "Anyone else have that experience that we just heard John describe?"

Level 2 Proposed Anchors

Checks in with specific members after someone shares: "David, have you ever had a similar experience to the one that John just shared?" "David, I watched you have a reaction to what John just described, have you had a similar experience?"

36. Refrained from conveying personal feelings of hostility and anger in response to disruptive member behavior

Level 0 Proposed Anchors

Did not refrain, "I'm upset by what you said, I can't believe you said that." "You are rude.

Level 1 Proposed Anchors

leader response exhibits irritability but not to the point of hostility or anger

Gritting teeth, using some body language that conveys anger, incongruent with a statement of support.

Level 2 Proposed Anchors

Redirect the client or handle the disruption without expressing anger. Requires acknowledgment of disruption: "I wonder if that comment might not be helpful for us in feeling safe to share in here. Could you help us understand what is upsetting you?"