

## Process-Outcome Relationships in Short- and Long-Term Psychodynamic Group Psychotherapy: Results From a Randomized Clinical Trial

Steinar Lorentzen

University of Oslo and Oslo University Hospital,  
Oslo, Norway

Bernhard Strauss and Uwe Altmann

University Hospital, Jena, Germany

This study focuses on the associations between cohesion, baseline predictors (e.g., symptoms, interpersonal problems, psychosocial functioning, presence of personality disorders [PDs]), and outcomes in short-term (6 months) and long-term (24 months) group-analytic psychotherapy (called STG and LTG, respectively). This study is part of a project investigating the impact of treatment duration on outcome in a randomized trial and relates to a subsample of 144 outpatients with complex diagnoses. Cohesion, commonly investigated in nonpsychodynamic short-term groups, was measured using the cohesiveness subscale of the Therapeutic Factors Inventory. We used hierarchical linear models for statistical analyses. Cohesion in STGs appeared to be sensitive to the patients' initial impairment (i.e., a higher level of interpersonal problems was associated with lower cohesion). The initial level of symptom distress and presence of PDs did not significantly affect cohesion in any of the therapy formats. Higher cohesion was associated with improvement in symptoms and interpersonal problems at the end of treatment in the STG, but not at follow-up. In the LTG cohesion was not significantly associated with outcome at any time-point. Thus, cohesion is an important predictor of improvement in STG. There were no significant differences between cohesion-predictor/outcome associations in STG and LTG. More research on larger samples using more frequent process measures and updated group relationship measures are needed to gain more knowledge of the possible change mechanisms within process-outcome research and the impact of duration of therapies on the cohesion-outcome associations.

*Keywords:* group psychotherapy, psychodynamic, cohesion, randomized clinical trial

The efficacy and effectiveness of group psychotherapy has been firmly established for a variety of mental disorders (Burlingame, Strauss, &

Joyce, 2013; Burlingame et al., 2016). According to an integrative model, components that may explain the outcome of group treatments are the

---

Steinar Lorentzen, Institute of Clinical Medicine, University of Oslo, and Department for Research and Development, Clinic for Mental Health and Addiction, Oslo University Hospital, Oslo, Norway; Bernhard Strauss and Uwe Altmann, Institute of Psychosocial Medicine and Psychotherapy, University Hospital Jena, Jena, Germany.

We want to express gratitude to the following institutions and persons for important contributions to this study: outpatient clinics at Community Mental Health Centers in Ålesund, Alna (Oslo) and in Sandnes and Ryfylke (Stavanger) and Laila Hjulstad, Knut Skjøstad, Vibeke Lohne, Allan Larsen, Kirsten Høbye,

Helge Knudsen, Marthe Horneland, Dagny Sande Børnes, Ole Inge Gjøyen, Wibeke Kløvning, Synnøve A. Kristiansen, Martin Mydske Nilssen, Ørjan Berg, Anette Fjeldstad, Jan V. Bakali, Torleif Ruud, and Per A. Høglend. Ole Klungsøyr and Svein Friis have given important comments in the final stages of the publication process.

Correspondence concerning this article should be addressed to Steinar Lorentzen, University of Oslo, Institute of Clinical Medicine, Clinic for Mental Health and Addiction, P.O. Box 1039, Blindern, 0315 Oslo, Norway. E-mail: [steinar.lorentzen@medisin.uio.no](mailto:steinar.lorentzen@medisin.uio.no)

formal change theory, aspects of the small group processes, and group leader and patient characteristics, as well as different structural aspects, such as treatment duration, session frequency, group size, and pregroup preparation (Burlingame, MacKenzie, & Strauss, 2004). Even though several characteristics of the patients and therapists predict or moderate the outcome of group therapy, it is unclear why specific group therapies work or which underlying mechanisms are most important for change (Burlingame et al., 2013). Small group processes have been the focus of research during the last decade, and there has been increasing interest in process factors. Group cohesion has long been the most central theoretical concept developed to capture the essence of small group processes (Yalom & Leszcz, 2005), and it is considered the group equivalent of therapeutic alliance in individual psychotherapy (Burlingame, Fuhrman, & Johnson, 2002; MacKenzie, 1998). Unfortunately, cohesion has lacked a uniform definition, which has led to the development of an array of measures reflecting different, but overlapping, aspects of the group process (Burlingame et al., 2004). Consequently, results related to cohesion reported in the clinical and empirical literature can hardly be compared. This has slowed the development of a broader understanding of the complex relationships between process and outcome and their roles as potential mechanisms of change.

In a meta-analysis, Burlingame, McClendon, and Alonso (2011) found an average correlation between cohesion and outcome of .25, which is within a similar range as the correlation between helping alliance and outcome in individual psychotherapy (Horvath, Del Re, Flückiger, & Symonds, 2011). Burlingame et al.'s meta-analysis included 40 studies based on counseling, psychotherapy, and personal growth groups and, in contrast to generally held beliefs, demonstrated that only 17 (43%) of the 40 studies reported significant associations between process and outcome. The selection criteria in the meta-analysis were groups of at least three members, at least one quantitative measure of cohesion and outcome, and the availability of information that allowed the calculation of effect sizes as weighted correlations. Only 10 studies (25%) focused on psychodynamic/existential groups, only 4 (10%) studied group as primary treatment, and most

groups had a short-term duration ( $M = 23.5$  sessions).

As our main objective was to study the associations between cohesion and outcome in short- and long-term psychodynamic group psychotherapy, we restricted our review to studies based on psychodynamic theory. Five of the psychodynamic papers covered in the meta-analysis presented results from studies of short-term (12 weekly sessions) supportive/interpretive group therapies for patients with complicated grief. Four of these studies used a cohesion measure developed by Piper, Marrache, Lacroix, Richardsen, and Jones (1983). One study found that one out of five cohesion measures is significantly associated with improvement in one out of three outcome measures (Joyce, Piper, & Ogrodniczuk, 2007). Kipnes, Piper, and Joyce (2002) used a participant-rated member-targeted Group/Member/Leader/Cohesion Scale (Piper et al., 1983) and an observer-rated global-targeted Group Cohesive Scale (Budman et al., 1989) but did not find any significant process-outcome associations. In a study of alexithymic patients, Ogrodniczuk, Piper, and Joyce (2005) reported a significant association between a therapist-rated cohesion measure (the patient's positive qualities, compatibility with the group, and significance as a group member) and outcome. In a study of patients with personality disorders (PDs), Ogrodniczuk, Piper, and Joyce (2006) found that cohesion in the group mediated the effect of interpersonal distress on attendance in supportive groups. The last of these studies demonstrated that a higher early rating on the Engaged subscale of the Group Climate Questionnaire (MacKenzie, 1983) is associated with less grief and general symptoms after therapy (Ogrodniczuk & Piper, 2003).

Two of the remaining five psychodynamic studies in the meta-analysis were observational and related to the same small sample of only 16 inpatients with mixed disorders. Higher cohesiveness correlated with better outcome at 12 and 18 months following treatment (Tschuschke & Dies, 1994), "relatedness" (similar to group cohesion) was positively associated with improvement at 18 months, and "positive group work" was unrelated to outcome (MacKenzie & Tschuschke, 1993). Crowe and Grenyer (2008) studied the relationships between alliance, group climate, cohesion, and outcome in a sample of 30 depressed patients receiving 16 sessions of expressive/supportive group therapy, finding that higher cohesion in the

groups was associated with better outcomes. In a study of a long-term analytic group with 12 patients, early cohesion ratings using repeated measures of alliance and cohesion did not predict changes in symptoms or interpersonal problems (Lorentzen, Sexton, & Høglend, 2004). Finally, Beutel et al. (2006) studied 144 inpatients with vocational strain and conflicts in a multimodal inpatient treatment program at a rehab hospital with a mean stay of 1.5 months. They found that group climate was a predictor of favorable outcome (i.e., more positive attitudes about returning to work).

Following the meta-analysis by Burlingame et al. (2011), only a few studies have been published that relate cohesion to outcome in psychodynamic or interpersonal psychotherapy. Dinger and Schauenburg (2010) studied cohesion, interpersonal style, and outcome in a psychodynamically oriented inpatient group treatment and found that high cohesion and an increase in cohesion predicted symptom improvement. Two studies from a group investigating process and outcome in group psychotherapy in women with binge eating disorder found that the convergence of individually rated cohesion and group perception of cohesion was related to improved self-esteem (Gallagher, Tasca, Ritchie, Balfour, Maxwell, et al., 2014). In addition, attachment anxiety moderated the relationship between growth in group cohesion and improvement in binge eating, as only patients with higher early values for attachment anxiety improved (Gallagher, Tasca, Ritchie, Balfour, & Bissada, 2014).

From a methodological point of view, all of these studies were heterogeneous in relation to group theory, subjects, focus of therapy, and study design. Furthermore, there was great variation between the studies regarding the rating unit (individual or group), measurement (mean of items or global), rater source (patient, therapist, observer), and even which elements of the cohesion measure (Piper et al., 1983) were used. Most studies in Burlingame et al.'s (2011) meta-analysis were based on bivariate correlations, which are not robust against bias caused by confounders (Sun, Shook, & Kay, 1996). The reviewed studies also have inconsistent results, as some demonstrate positive associations between cohesion and outcome and other studies do not.

Therefore, the empirical literature on cohesion and outcome in group therapy lacks studies on long-term group therapies and therapies with a psychodynamic background. Many studies are also methodologically weak (Blackmore, Tantam, Parry, & Chambers, 2011), do not take long-term effects into account, do not attempt to address problems with the nested data that are invariably present in group studies (Baldwin, Murray, & Shadish, 2005), and have small sample sizes. Because of these limitations, we were interested in studying the role of treatment length in psychodynamic group psychotherapy and designed a randomized clinical trial (RCT) of outpatients with mixed diagnoses who were allocated to either short-term group (STG) or long-term group (LTG) analytic therapies of 20 or 80 weekly 90-min sessions, respectively. We previously published outcomes in this RCT using repeated self-rated measures of symptoms (Symptom Check List-90-R [SCL-90-R]; Derogatis, 1977) and interpersonal problems (Inventory of Interpersonal Problems; IIP-C; Alden, Wiggins, & Pincus, 1990), and an observer-rated measure of psychosocial functioning (Global Assessment of Functioning [GAF]; Endicott, Spitzer, Fleiss, & Cohen, 1976). On a group level, using linear mixed model analyses, we found similar changes in patients with both treatment formats over 3 years (Lorentzen, Ruud, Fjeldstad, & Høglend, 2013). When we tested the presence of a PD at baseline as a potential moderator of treatment effects, patients with PD in the STG and LTG changed equally on the Global Severity Index (GSI) of the SCL-90-R and IIP during the first 6 months. However, at the 3- and 7-year follow-up assessments, LTG patients had significantly larger effects on the GSI and IIP than STG patients. Patients with PD in the LTG also changed significantly more when the GAF was used as an outcome measure (Fjeldstad, Høglend, & Lorentzen, 2016; Lorentzen, Ruud, Fjeldstad, & Høglend, 2015).

Using data from this RCT, we wanted to study potential associations between group process variables and treatment outcome. After several attempts to include different process variables, we finally chose cohesion, which is the process variable most commonly studied in

groups.<sup>1</sup> Specifically, we raised the following research questions:

1. What are the associations between the process variable cohesion and outcome, and do associations with various outcome variables differ between the STG and LTG?
2. To what degree do the baseline variables symptom severity, interpersonal problems, psychosocial functioning, and presence of PD predict cohesion in the STG and LTG? Are there differences between the two treatment formats?

Our specific hypothesis, based upon reviews of the role of cohesion in group psychotherapy, was that cohesion is a predictor of outcome (i.e., less symptomatic distress, lower degree of interpersonal problems, and better psychosocial functioning) in both treatment formats. We further hypothesized that greater initial symptoms, including symptoms of a PD, would be associated with lower cohesion in both therapy formats.

## Method

### Sites, Patient Referral, Inclusion, and Randomization

This multisite study was performed in three urban areas of Norway and has been described extensively elsewhere (Lorentzen et al., 2013, 2015). The study was implemented within the regular public mental health system, and patients were locally evaluated after referral from community mental health centers, general practitioners, and practicing psychiatrists/psychologists. Inclusion criteria were one or more axis I or II diagnoses, interest in working with problems and relationships in groups, and willingness to accept randomization. Exclusion criteria were psychosis, main diagnosis of alcohol/drug addiction, and organic brain disease. The project was approved by the Data Inspectorate and the Regional Committee on Ethics in Health Research and registered at ClinicalTrials.gov under NCT00521.

A total of 175 patients were initially referred to and interviewed by trained evaluators (psychologists and psychiatrists) who were not in-

involved in the treatment (Schulz & Grimes, 2002). The interviewers had participated in several seminars with the principal investigator to ensure standard procedure for evaluation/randomization and had been trained to secure satisfactory interrater reliability in diagnostic interviews and the use of observer-rated measures. Eight patients were not eligible for participation and 167 patients were randomized to the two treatment conditions. Randomization was performed by the local evaluators with a stratified randomization for gender, drawing at least two women and two men for each group to secure mixed gender groups. Time from referral to assessment/randomization varied from 1 to 7 weeks ( $M = 2.5$  weeks). The time from randomization to the start of therapy varied from 2 to 12 weeks ( $M = 7$  weeks).

<sup>1</sup> Initially, we tried to extend multilevel structural equation models consisting of data from three key relationship constructs in group therapy (Bakali et al., 2009), group climate, cohesion, and alliance, by adding predictors of process and outcome variables. We failed to obtain an adequate model fit when outcome variables were included and abandoned this strategy. We then decided to study if and how alliance measured with the Working Alliance Inventory and cohesion are related to outcome.

When we combined all three waves in which the process variable had been assessed and added predictors (GSI, IIP, and GAF at baseline) and outcome (GSI, IIP, and GAF) variables to the SEM, the number of parameters was higher than the number of subjects in our sample. As expected, the predictors cohesion and alliance highly correlated (SGT:  $r = .777$ ; LGT:  $r = .838$ ), which usually causes problems of multicollinearity in regression models. Therefore, we tried various ways of combining the process variables (i.e., a model with the average of cohesion and alliance, a model with an average of the two predictors and a difference variable). In all cases, the model fit was insufficient (RMSEA > .08, CFI < .93), and we finally decided to restrict the analysis of process-outcome associations to one manifest process variable: (cohesion), using hierarchical linear models. Though the process variable and multiple outcome variables are estimated in one model in SEM associations between multiple variables at the beginning of therapy, we applied multiple HLMs to reduce the number of model parameters, e.g., one HLM for the prediction of cohesion and another HLM for the prediction of GSI at T3.

In our initial HLM analysis, each regression coefficient was modelled as random, but no random effect had significant variance, including the association between cohesion and outcome. Therefore, we used only one random effect, a random intercept, which is a compromise between the theoretical assumption of a multilevel data structure and the recommendation to exclude all random effects without significant variance.

Of the 167 patients, 144 were included in the process-outcome analyses presented in this paper. Excluded patients were early drop-outs ( $n = 18$ ), one extreme outlier who was deleted from all analyses of longitudinal data, and four patients who did not contribute any data on cohesion. A flowchart of the sample selection was published elsewhere (Lorentzen et al., 2013).

### Evaluation Procedure

The sample time-points are depicted in Figure 1. Prerandomization evaluation (2 hr to 4 hr) included full clinical and diagnostic interviews using Mini International Neuropsychiatric Interview (MINI PLUS; Sheehan et al., 2002) and the Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders Axis II (SCID-II; First, Gibbons, Spitzer, Williams, & Benjamin, 1997). The SCID-II was audiotaped. All patients completed the SCL-90-R (Derogatis, 1977) and the IIP-C (Alden et al., 1990). Both measures were completed every 6 months during the LTG and then 3 years after baseline, a total of six times (see Figure 1). The GAF (Axis V; DSM-IV-R; Endicott et al., 1976) was used to evaluate psychosocial functioning at baseline and 3 years later. This interview was also audiotaped. The cohesion measure was administered to the patients by their therapists at the end of specific sessions defined as three waves of data collection: Sessions 3 and 4 (Wave 1); 10 and 11 (Wave 2); and 17 and 18 (Wave 3; see Figure 1). Patients usually completed the measure right after the session. When participants did not attend either of the two sessions in a wave, ratings from the session

closest in time were used. Only one response per instrument per wave was included in the analysis, and the main selection principle was to include the ratings from the session with the highest attendance. The numbers of patient ratings were 139 (Wave 1), 130 (Wave 2), and 126 (Wave 3). All included patients fulfilled a minimum of one wave.

### Groups, Therapies, and Therapists

Eighteen psychotherapy groups (nine STG and nine LTG) were established. Each group consisted of seven to eight patients and one therapist, and each therapist conducted both a STG and a LTG. The groups were closed, but to secure the 'survival of the group' in case of serious early attrition, patients who left the LTG during the first 6 months ( $n = 8$  patients) were replaced by the next patient of the same gender. The STG and LTG therapies consisted of 20 or 80 weekly sessions lasting 90 min each (approximately 6 or 24 months, respectively). The length of therapy chosen was a compromise based on clinical judgment, local routines, and knowledge of the research literature. The mean number of sessions attended was 18.9 ( $SD = 3.8$ ) for STG and 57.7 ( $SD = 26.9$ ) for LTG. Thus, the attendance rates were 94.5% and 72.1%, respectively.

The nine therapists (7 women) comprised two psychiatrists, three psychologists, three psychiatric nurses, and one social worker with a mean age of 52.7 years ( $SD = 3.7$ ), mean time in practice of 19.7 years ( $SD = 4.4$ ), and mean formal postgraduate psychotherapy training of 12.5 years ( $SD = 3.7$ ). All of the therapists had completed a formal group education of at least 5 years, and they were

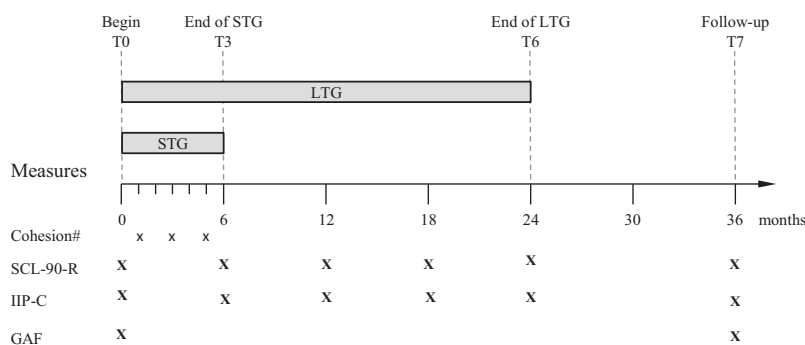


Figure 1. Assessment points of the study.

trained in both formats and met regularly for supervision by the principal investigator while they were running their groups.

Both therapy formats were manualized group psychotherapies (Lorentzen, 2014) built on a psychodynamic/group-analytic understanding that emphasizes a developmental perspective of personality, the existence of internal representations of interpersonal relationships, and the influence of unconscious individual and group processes on behavior (psychological 'causation'). The STG format was built partly on MacKenzie's (1997) generic time-limited phase-oriented group psychotherapy, whereas the LTG format was mainly a modification of Foulkes' group analytic therapy (Foulkes & Anthony, 1984; Lorentzen, 2014). The therapist used facilitating or supportive techniques to promote interaction in combination with traditional psychodynamic techniques of confrontation and interpretation. Group members were asked to interact and focus on conscious, as well as potential derivatives of unconscious, processes among themselves, others, and the group. When interacting with each other, the group members' individual patterns (both adaptive and dysfunctional) were commonly activated and appeared as multiple transferences and resistances (or functional coping behavior) in the group. The aims of the therapy were to become aware of intrapsychic conflicts and dysfunctional interpersonal patterns, and to increase the understanding of self, others, and interpersonal relationships. Thus, the treatments were not primarily targeted to specific symptom disorders, but therapists in the STG were instructed to have a more circumscribed focus, to be more active, to work more in the present, and to be more attentive to the impending termination phase.

### Treatment Fidelity

Thirty-nine audio recordings from the first 6 months of therapy (Sessions 3, 10, and 17) for both formats were randomly selected from the pool of 54 recordings. Two evaluators blindly and independently rated sessions for the therapist's activity level, degree of focus, group work in the present, and competence using Likert scales ranging from 0 (not at all) to 4 (very much). The intraclass correlation coefficients (ICCs; Shrout & Fleiss, 1979) on the therapy process scales ranged from 0.70 to 0.94. There was significantly more work on a circumscribed

problem focus (STG:  $M = 3.7$ ,  $SD = 1.8$ ; LTG:  $M = 2.3$ ,  $SD = 1.4$ ; Cohens  $d = 0.9$ ) and work in the present (STG:  $M = 1.8$ ,  $SD = 0.8$ ; LTG:  $M = 1.4$ ,  $SD = 0.5$ ; Cohens  $d = 0.6$ ) in the STG as suggested in the treatment manuals. The therapist's activity level was equal in the two groups (STG:  $M = 2.1$ ,  $SD = 0.5$ ; LTG:  $M = 2.2$ ,  $SD = 0.8$ ; Cohens  $d = 0.2$ ). The therapist's mean competence was similar in both formats, ranging from moderate to high (STG:  $M = 2.8$ ,  $SD = 0.5$ , range 1.7–3.4; LTG:  $M = 2.6$ ,  $SD = 0.5$ , range 1.9–3.1; Cohens  $d = 0.4$ ).

### Measures

**Outcome measures.** The SCL-90-R (Derogatis, 1977) consists of 90 items rated on a 5-point Likert scale. The symptom load is reflected by the GSI provided by the mean of all 90 symptom ratings. The internal consistency using Cronbach's alpha was 0.91.

The IIP-C (Alden et al., 1990) is a well-validated and widely used self-report measure for assessing problems in interpersonal relationships. Sixty-four items are rated on a 5-point Likert scale. The first 39 items begin with the phrase, 'It is hard for me to . . .' and the next 25 items describe 'Things that I do too much.' The IIP has demonstrated high test-retest stability over 10 weeks but is still sensitive to change (Horowitz, Rosenberg, Baer, Ureño, & Villaseñor, 1988). We used the mean total score of the IIP, which reflects the overall degree of distress associated with interpersonal problems. Cronbach's alpha was 0.94.

The GAF interview occurred at baseline and 3 years later (i.e., 2.5 years after termination of the STG and 1 year after termination of the LTG). The ICCs among five evaluators rating 20 cases from tapes was 0.76 for the GAF-S (symptoms) and 0.80 for the GAF-F (functioning). We used the average of the GAF-S and GAF-F scores as a combined GAF score. This is an observer-rated measure, and higher values on the GAF scale (0–100) indicate better psychosocial functioning. Seventy-nine of 90 patients (88%) completed the 3-year follow-up interview after LTG analytical therapy versus 71 of 77 patients (92%) in STG therapy.

**PD.** The presence of PDs was used as a predictor of outcome (Lorentzen et al., 2015), and the PD diagnoses were established through SCID-II interviews. The ICC was 0.85 between

evaluators using SCID-II and a blind evaluator rating 30 cases. For the patients with PD ( $n = 75$ ), 91% and 93% participated in the STG and LTG follow-up evaluations, respectively.

**Process measure.** Group cohesion was assessed with the Cohesiveness subscale (COH) of the Therapeutic Factors Inventory (Lese & MacNair-Semands, 2000), which is included in the revised Core Battery (Strauss, Burlingame, & Bormann, 2008). The COH is a nine-item self-report questionnaire using a 7-point Likert scale ranging from ‘strongly disagree’ to ‘strongly agree.’ Based on a sample from a counseling center population, Lese and MacNair-Semands (2000) found a test–retest reliability of 0.93 over 1 week, and a Cronbach’s alpha of 0.90 for this measure. In the present study, the Cronbach’s alphas for COH scores at Waves 1, 2, and 3 were 0.83, 0.87, and 0.90, respectively. For the analyses, we used the average of Waves 1, 2, and 3.

## Data Analysis

**Missing data.** In our data set, variables assessed at the beginning of therapy (age, sex, PD, IIP, GSI, GAF) were complete. As in other longitudinal studies, not all patients completed the questionnaires at the end of therapy. For 12 patients (8.3%), we had no IIP, GSI, and GAF measures at T3 and T7 (see Figure 1). Cohesion was assessed at Wave 1 by 132 patients, at Wave 2 by 129 patients, and at Wave 3 by 125 patients. However, for each of the 144 patients, we had a minimum of one cohesion score. Taking age, sex, treatment condition, PD, GAF, IIP, GSI, and cohesion measures into account, Little’s MCAR (missing completely at random) test indicated that missing data could be assumed as completely at random ( $\chi^2 = 112.028$ ,  $df = 136$ ,  $p = .934$ ). Accordingly, missing questionnaire data were imputed using the R package missForest (Stekhoven, & Bühlmann, 2012), a nonparametric approach that provides better imputation results than other methods (Waljee et al., 2013). The parameters were as follows: maximum iterations = 20, random trees = 1000, seed = 47. Fit indices of the imputation can be classified as good (normalized root-mean-square error [NRMSE] = 0.112; proportion of falsely classified [PFC] = 0.0).

**Descriptive statistics and randomization check.** Descriptive statistics were used to characterize the sample. We then tested for se-

lection effects because only 144 of the 167 randomized patients were included in this study. We also checked whether the randomization to the two treatment formats was successful. In both cases, we used univariate tests ( $t$  test for metric variables and chi-square tests for categorical variables) for variables assessed at the beginning of therapy (sociodemographic and clinical data).

**Dependency in nested data.** Data obtained from several small groups revealed a multilevel structure (Baldwin et al., 2005), normally requiring multilevel models for data analysis. We considered the multilevel structure and tested whether membership in one of the therapy groups explained the variance in the outcome variables (GSI, IIP, GAF) at the different time points using random intercept models. In two random intercept models, the variance of the random effect was zero (GSI at T7 and GAF at T7). In the other models, the variances could be estimated but were not significantly different from zero. The ICCs were between 0.03 (GSI at T3) and 0.11 (IIP at T3). For the cohesion measure, we found a small random effect (ICC = 0.23). Given the sample size ( $N = 144$  patients), 18 treatment groups, and 7 to 10 patients per group, there was at least some support for a multilevel structure of the variables. Accordingly, we decided to model the data using a hierarchical linear model (HLM) including the nested variable ‘patient within therapy group.’

**Testing predictor-process-outcome relationships.** To analyze the associations between initial symptoms, cohesion, and treatment outcome, we applied correlations and HLMs. Process-outcome associations are commonly examined using correlations (Burlingame et al., 2011). We computed bivariate correlations between COH and GSI, IIP, and GAF assessed at the different sampling points and for each of the treatment formats.

In contrast to correlations, HLMs consider the presence of PD, initial symptoms, and initial interpersonal problems as confounders and allows the consideration of nested data. We applied a two group model (i.e., respecting the two treatment conditions) using the following model equation:

$$Y_{ij} = I_{X_{ij}=STG}(\beta_{00} + \beta_{01}PD_{ij} + \beta_{02}GSI_{T0ij} + \beta_{03}IIP_{T0ij} + \beta_{04}COH_{ij})$$

$$+ I_{X_{ij}=LTG}(\beta_{10} + \beta_{11}PD_{ij} + \beta_{12}GSI_{T0ij} + \beta_{13}IIP_{T0ij} + \beta_{14}COH_{ij}) + U_{0j} + e_{ij}$$

The indices *i* and *j* indicate the patient and therapy group, respectively. The dependent variable is *Y<sub>ij</sub>* (which can be GSI at T3, GSI at T7, IIP at T3, and so on). The dummy variables *I<sub>X<sub>ij</sub>=STG</sub>* and *I<sub>X<sub>ij</sub>=LTG</sub>* indicate (with 1 and 0) whether patient *i* in therapy group *j* was treated with STG or LTG. *PD<sub>ij</sub>*, *GSI<sub>T0ij</sub>*, and *IIP<sub>T0ij</sub>* are confounders assessed at the beginning of therapy, and *COH<sub>ij</sub>* is the process variable. *U<sub>0j</sub>* is a level-2-residual variable for the nesting ‘patient within therapy group,’ and *e<sub>ij</sub>* is the level-1-residual variable. The regression coefficients  $\beta_{00}, \dots, \beta_{04}$  refer to the associations between outcome and independent variables in the STG condition, and  $\beta_{10}, \dots, \beta_{14}$  to the corresponding associations in the LTG condition. In a previous analysis, we modeled all regression coefficients as random and found that none of the random effects had a significant variance. Accordingly, we minimized the number of random effects to address both the theoretical suggestion of the nested data structure and the limitations of the data (sample size). For these models, we also computed the ICCs:  $ICC = \frac{Var(U_{0j})}{(Var(U_{0j}) + Var(e_{ij}))}$ . Note that we used GAF at T0 instead of GSI at T0 as a confounder for the GAF at T7 as a dependent variable.

Next, we compared the STG and LTG with respect to the association between cohesion and

outcome using the corresponding regression coefficients. In the STG,  $\beta_{04}$  refers to the cohesion-outcome association, and the corresponding coefficient in the LTG is  $\beta_{14}$ . A *t* test examined whether the strength of the cohesion-outcome association was equal in both treatment conditions ( $H_0: \beta_{04} - \beta_{14} = 0$ ).

To obtain standardized coefficients as a measure of effect size, we applied an HLM with z-transformed variables. Standardized coefficients can be interpreted as (partial) correlations indicating small ( $\beta \geq .1$ ), medium ( $\beta \geq .3$ ), or large effect sizes ( $\beta \geq .5$ ).

## Results

### Descriptive Statistics

Table 1 summarizes the patients’ baseline characteristics for the total sample (*N* = 167), the examined sample (*N* = 144), and the examined STG (*n* = 69) and LTG (*n* = 75; for more details, see Lorentzen et al., 2013). Comparing the subsamples, we found slightly more patients with a PD in the LTG subsample,  $\chi^2(1, 143) = 6.421, p = .011$ , Cramer’s *V* = .211. This suggested that PD should be used as a covariate in the regression model to control for confounding effects. All other differences, including the mean number of SCID-II criteria for PD, were not significant. Because this is a categorical (yes/no) diagnosis, it means that some patients in the LTG were just above and some in the

Table 1  
Sample Characteristics, Including Scores on Outcome Measures at Therapy Start

Characteristic	Randomized sample ( <i>N</i> = 167)		Examined sample ( <i>N</i> = 144)		STG ( <i>n</i> = 69)		LTG ( <i>n</i> = 75)	
	<i>M</i>	( <i>SD</i> )	<i>M</i>	( <i>SD</i> )	<i>M</i>	( <i>SD</i> )	<i>M</i>	( <i>SD</i> )
Gender (% females)	62.9%		61.8%		62.3%		61.3%	
Age (average in years)	38.4	(9.4)	38.5	(9.3)	38.6	(9.1)	38.4	(9.5)
Number of comorbid diagnoses	3.3	(2.0)	3.3	(2.0)	3.1	(1.9)	3.5	(2.1)
Personality disorder (PD)	44.9%		47.2%		36.2%		57.3%	
Number of PD criteria	8.0	(6.8)	8.0	(6.6)	7.2	(6.6)	8.8	(6.4)
GAF at T0	58.9	(7.8)	59.0	(7.8)	59.6	(8.2)	58.5	(7.1)
GSI at T0			1.0	(.6)	.9	(.5)	1.1	(.6)
IIP at T0			1.4	(.5)	1.3	(.6)	1.4	(.4)
Early cohesion			5.3	(.9)	5.3	(.9)	5.3	(.9)

Note. PD = personality disorder; GAF = Global Assessment of Functioning; GSI = Global Severity Index; IIP = Inventory of Interpersonal Problems.



STG just below the cut-off point (i.e., numbers of positive SCID-II criteria), and that the load of personality pathology is similar in the two subsamples. We did not find any significant difference between the two therapy conditions for the process variable cohesion.

### Bivariate Correlations

All correlations between our process variable, cohesion, and the predictors/outcome measures at T0, T3, and T7 are reported in Table 2.

In both the STG and LTG, higher cohesion was significantly associated with fewer symptoms at T0 and T3, and fewer interpersonal problems at T3 and T7. A higher cohesion was significantly associated with less symptoms at T7 only in the LTG. For the GAF, we did not find any significant correlations. A higher initial (T0) level of interpersonal problems and presence of a PD, were both significantly associated with lower cohesion in the STG.

### Hierarchical Linear Models

We examined the associations between PDs, initial symptoms (or initial GAF), initial interpersonal problems, cohesion, and outcome at different time-points using the HLM. Detailed statistics for two groups are reported in Table 3.

Table 2  
*Correlations Between Cohesion and Predictors/Outcomes Depending on Treatment Condition*

Predictors/Outcomes	STG ( <i>N</i> = 69)	LTG ( <i>N</i> = 75)
GSI at T0	-.260*	-.321**
GSI at T3	-.456**	-.321**
GSI at T7	-.199	-.301**
IIP at T0	-.306*	-.224†
IIP at T3	-.417**	-.301**
IIP at T7	-.355**	-.329**
GAF at T0	.179	.097
GAF at T7	-.058	.162
PD (1 = yes, 0 = no) <sup>a</sup>	-.249*	-.142

*Note.* Cohesion reflects mean of three time-points. STG = short-term group; LTG = long-term group; GAF = Global Assessment of Functioning; GSI = Global Severity Index; IIP = Inventory of Interpersonal Problems; T0 = baseline; T3 = end of STG; T7 = 1 year after end of LTG.

<sup>a</sup> Spearman correlation coefficient, otherwise Pearson correlation.

†  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ .

Higher initial levels of interpersonal problems significantly predicted lower cohesion in the STG, but not the LTG. On the other hand, the degree of symptom distress or presence of PD at baseline did not significantly affect cohesion in any of the therapy formats. Furthermore, higher cohesion significantly predicted both fewer symptoms and fewer interpersonal problems in the STG at T3 (6 months). At T7 (36 months after baseline), the same associations were no longer significant. In the LTG, there were no significant associations between baseline levels of predictors and cohesion, or between cohesion and improvement in symptoms or level of interpersonal problems at T3 or T7. GAF at T7 was not significantly associated with cohesion in any of the group formats. A summary of all significant associations is provided in Figure 2. Comparing STG and LTG regarding the strength of the cohesion-outcome associations, we did not find any significant differences between the two treatment formats (see Table 3).

### Discussion

Group cohesion is the most frequently examined process variable in group psychotherapy. However, limited knowledge is available regarding the role of cohesion for the outcome in psychodynamic group therapy, especially in LTG. Thus far, measures have been shown to reflect different, though overlapping, aspects of the group process (Burlingame et al., 2004).

### Cohesion Outcome Associations

Our first research question concerned potential associations between cohesion and outcome in several areas, and whether these associations would differ between STG and LTG. We hypothesized that cohesion would be a significant predictor of improvement in all outcome variables in both treatment formats.

Using a HLM, we found significant associations between higher cohesion and improvement of both symptom distress and the level of interpersonal problems after therapy in the STG at T3 (i.e., 6 months; the effect sizes were small). These associations were not sustained at T7 (36 months after baseline, which is 30 months after the end of the STG). In the LTG, there were no significant associations between

Table 3  
*Statistics of Hierarchical Linear Model for the Prediction of Cohesion, GSI, IIP, and GAF Assessed at Different Time Points*

Predictors	STG			LTG			Difference between STG and LTG		
	<i>b</i>	<i>SE</i>	<i>p</i>	<i>b</i>	<i>SE</i>	<i>p</i>	<i>b</i> <sub>STG</sub> - <i>b</i> <sub>LTG</sub>	<i>SE</i>	<i>p</i>
Cohesion as dependent variable									
Intercept	5.89	(.26)	.000	5.76	(.31)	.000	.13	(.40)	.749
PD	-.23	(.20)	.242	-.10	(.19)	.602	-.13	(.27)	.620
GSI T0	.02	(.21)	.908	-.22	(.16)	.173	.24	(.26)	.360
IIP T0	-.41	(.18)	.026	-.15	(.24)	.538	-.26	(.30)	.380
Var( <i>e</i> ) = .46, <i>SE</i> = .06, <i>p</i> < .001 Var( <i>U</i> <sub>0</sub> ) = .15, <i>SE</i> = .07, <i>p</i> = .045, ICC = 24.4%									
GSI at T3 as dependent variable									
Intercept	1.31	(.42)	.002	.66	(.40)	.103	.64	(.58)	.270
PD	.21	(.12)	.067	.01	(.11)	.946	.21	(.16)	.189
GSI T0	.55	(.12)	.000	.62	(.09)	.000	-.07	(.15)	.640
IIP T0	-.01	(.11)	.910	.01	(.14)	.926	-.02	(.17)	.886
COH	-.20	(.07)	.003	-.07	(.06)	.282	-.14	(.09)	.147
Var( <i>e</i> ) = .16, <i>SE</i> = .02, <i>p</i> < .001 Var( <i>U</i> <sub>0</sub> ) = .01, <i>SE</i> = .01, <i>p</i> = .507, ICC = 4.2%									
GSI at T7 as dependent variable									
Intercept	.33	(.43)	.438	.92	(.40)	.023	-.59	(.59)	.318
PD	.12	(.12)	.317	.04	(.11)	.721	.08	(.16)	.613
GSI T0	.54	(.13)	.000	.37	(.10)	.000	.18	(.16)	.272
IIP T0	-.02	(.11)	.851	-.08	(.14)	.581	.06	(.18)	.748
COH	-.02	(.07)	.781	-.10	(.06)	.112	.08	(.09)	.386
Var( <i>e</i> ) = .18, <i>SE</i> = .02, <i>p</i> < .001 Var( <i>U</i> <sub>0</sub> ) = .00, <i>SE</i> = .00, <i>p</i> = .999, ICC = .0%									
IIP T3 as dependent variable									
Intercept	1.03	(.36)	.005	1.37	(.34)	.000	-.35	(.50)	.488
PD	.16	(.10)	.107	.10	(.09)	.257	.06	(.13)	.665
GSI T0	-.05	(.10)	.664	.11	(.08)	.186	-.15	(.13)	.248
IIP T0	.62	(.09)	.000	.26	(.12)	.027	.36	(.15)	.015
COH	-.13	(.06)	.023	-.10	(.05)	.064	-.03	(.08)	.695
Var( <i>e</i> ) = .12, <i>SE</i> = .02, <i>p</i> < .001 Var( <i>U</i> <sub>0</sub> ) = .00, <i>SE</i> = .01, <i>p</i> = .543, ICC = 4.0%									
IIP T7 as dependent variable									
Intercept	.88	(.47)	.062	1.15	(.44)	.011	-.27	(.64)	.682
PD	.32	(.13)	.016	.05	(.12)	.677	.27	(.18)	.128
GSI T0	.21	(.14)	.127	.13	(.11)	.230	.08	(.17)	.639
IIP T0	.30	(.12)	.012	.29	(.15)	.063	.02	(.19)	.938
COH	-.12	(.08)	.130	-.13	(.07)	.057	.02	(.10)	.858
Var( <i>e</i> ) = .21, <i>SE</i> = .03, <i>p</i> < .001 Var( <i>U</i> <sub>0</sub> ) = .00, <i>SE</i> = .01, <i>p</i> = .964, ICC = .2%									
GAF at T7 as dependent variable									
Intercept	46.05	(13.45)	.001	16.08	(14.42)	.267	29.97	(19.72)	.131
PD	-1.59	(2.91)	.586	5.35	(2.68)	.048	-6.94	(3.96)	.082
GAF T0	.55	(.16)	.001	.72	(.17)	.000	-.17	(.23)	.464
IIP T0	.23	(2.45)	.927	-2.28	(3.13)	.467	2.51	(3.97)	.529
COH	-2.01	(1.73)	.246	2.21	(1.51)	.146	-4.23	(2.30)	.068
Var( <i>e</i> ) = 108.62, <i>SE</i> = 13.27, <i>p</i> < .001 Var( <i>U</i> <sub>0</sub> ) = .00, <i>SE</i> = .00, <i>p</i> = .999, ICC = .0%									

*Note.* GSI = Global Severity Index; IIP = Inventory of Interpersonal Problems; GAF = Global Assessment of Functioning; PD = personality disorder; STG = short-term group; LTG = long-term group; Var(*e*) = variance of Level 1 residual variable; Var(*U*<sub>0</sub>) = variance of Level 2 residual variable; ICC = intra-class correlation; *N*<sub>STG</sub> = 69; *N*<sub>LTG</sub> = 75.

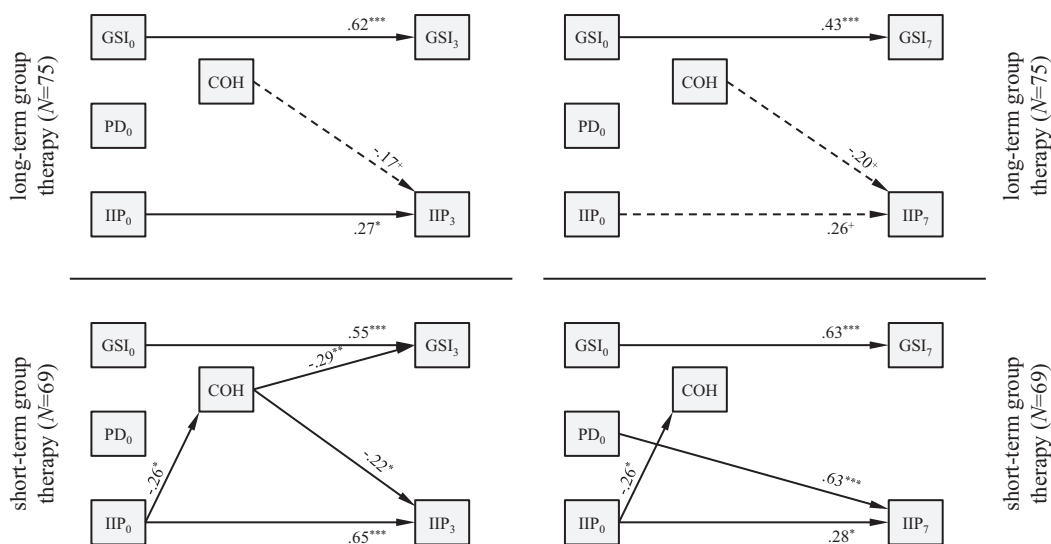


Figure 2. Significant associations between baseline variables, cohesion, and outcome with standardized regression coefficients using HLM. <sup>+</sup>  $p < .10$ ; \*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$ .

cohesion and outcome (level of symptoms and interpersonal problems) at T3 or T7. There were no significant associations between cohesion and observer-rated GAF scores at T7 in either group format.

Our results, especially the association between cohesion and fewer symptoms after STG, are consistent with findings from several time-limited studies (e.g., Crowe & Grenyer, 2008; Dinger & Schauenburg, 2010; Ogradniczuk, & Piper, 2003). Tschuschke and Dies (1994) also found fewer symptoms (and improved psychosocial functioning) after 6 months of inpatient group treatment ( $M = 93$  sessions), but it is difficult to compare those results with ours because of differences in session frequency and potential influence of other factors in a complex hospital treatment. The lack of cohesion-outcome associations after LTG in this study is similar to what was found in a previous study of cohesion-outcome associations in another sample of outpatients (Lorentzen et al., 2004). We find the association between higher cohesion and improvement in the IIP in STG interesting, because there is scarce evidence for such results in previous cohesion-outcome studies in clinical populations (Burlingame et al., 2011). We found no significant differences in the strength of the cohesion-outcome associations (represented by

the regression coefficients) between the STG and LTG.

From a clinical point of view, a possible explanation why higher cohesion is significantly associated with improvement in both symptoms and interpersonal problems during STG may be due both to the reduced time and/or a specific therapist attitude in the short-term format. The limited duration of the STG may create a sense of urgency and trigger a need in patients to work on their problems. This may then speed up the therapeutic work and promote earlier improvement, especially of symptoms. Another factor that may speed up changes during the STG is that therapists are advised to focus on specific problems, be active, and to work mostly in the present. In addition, the group process in the STG is systematically conceptualized in group stages (engagement, differentiation, interpersonal work, termination) as a back drop for both understanding and intervening in the process. These elements give structure to the STG and 'push' the group and the individual members forward.

In our more limited review of psychodynamic group therapies, compared to Burlingame et al. (2011), positive cohesion-outcome relationships were found in the majority, but not all, of the studies (e.g., Kipnes et al., 2002; Lorentzen et al., 2004). On the basis of the general

assumptions in group therapy, we could have expected a stronger relationship between cohesion and all measured outcomes. However, although some researchers have stated that cohesion has shown a positive association with patient improvement in nearly every published report (e.g., Tschuschke & Dies, 1994), only 17 (43%) of the 40 studies in Burlingame et al.'s (2011) meta-analysis reported a significant association between cohesion and patient improvement. Our results indicate that cohesion averaged across a short-term therapy, is an important (significant) therapeutic factor in STG.

Regarding the cohesion-outcome associations, we found more significant associations in correlation analyses than in the HLM analyses. This could be expected because, in contrast to what happens in correlations, the HLM considers potential confounders and the nested structure of the data. There has been an increase in the use of more complex models in group research because a HLM (and sometimes SEM) is more appropriate for the specific characteristics of the group data (e.g., Bakali, Baldwin, & Lorentzen, 2009; Lo Coco, Gullo, & Kivlighan, 2012; Tasca, Balfour, Ritchie, & Bissada, 2007). On the other hand, an HLM usually requires large study samples, which usually encompass large investments of money and human resources, especially if long-term therapies are examined.

### Predictors of Cohesion

Our second research question was whether the level of the baseline variables symptom severity, interpersonal problems, psychosocial functioning, and presence of PD can predict cohesion in the STG and LTG. We also wanted to know if there are differences between the two treatment formats. Our hypothesis was that more symptoms at baseline, including symptoms of PD, would be associated with lower cohesion in both therapies. This hypothesis was not confirmed for any of the formats. However, a higher initial level of interpersonal problems predicted a lower cohesion in STG, but not the LTG. Again, there were no significant differences between the strength of the predictor-cohesion associations when we compared the STG and LTG. A possible clinical explanation for the significant predictor-cohesion association in STG might be that a sense of urgency

and a high therapist demand for interactional work in the here-and-now, which signifies STG, strongly activates problematic relationship patterns.

### Strengths and Limitations

This study has several strengths. First, carefully selected patients were randomized to two manualized psychodynamic treatment conditions and treatment fidelity was monitored and found satisfactory. We used multilevel analyses to control for confounder variables and dependency in the data, and we compared the cohesion-outcome associations of two treatment formats within the same analyses. Patient outcomes were followed up to 2.5 years and 1 year after therapy in the STG and LTG, respectively. Second, we applied both correlations and multilevel analyses. Process-outcome associations are often studied by correlating an early or average score of cohesion with outcome (Burlingame et al., 2011), an approach that, according to Sun et al. (1996), is often not robust against bias caused by confounders. Investigators have commonly used samples of convenience and observational designs in such studies and failed to control for initial levels of distress or other potential confounders, such as the presence of PDs. In the case of correlated predictors, correlations can lead to spurious correlations. For example, in our correlation analysis we found significant associations between cohesion and initial GSI, initial IIP, and PD in the STG. Using the more conservative HLM, only the relationship between initial IIP and cohesion was significant. This suggests that the correlations between PD and cohesion, and between initial GSI and cohesion may be spurious.

A limitation of this study is that we did not assess GAF at T3 and T6. Consequently, to compare outcome results, we had to use the assessments at T7, 1 year after termination of the LTG. Although we had a larger sample size than most of the psychodynamic studies summarized by Burlingame et al. (2011), we note small sample size as a further limitation. In our study, the 144 patients were distributed to 18 therapy groups with 7 to 10 patients per group. For 12 (8.3%) of the patients we had no outcome measures and imputed missing data and did not offer optimal conditions for the HLM (Hox, 2010).

## Conclusion and Research Implications

We were only able to demonstrate that cohesion is important for outcome in the time-limited group-analytic therapy. Cohesion in the STG is associated with better outcomes in both symptoms and interpersonal problems after therapy, but not at follow-up. In the LTG, there are no significant associations at any time-point. Cohesion has not been demonstrated to be a mediator of change in groups, and the concept may be too composite to have such a function. Our results indicate that cohesion is a therapeutic factor in STG. Based on our clinical and research experience, we think that a ‘good enough’ cohesion in a STG can be developed by establishing a good relationship with patients during the evaluation/preparation phase, and by negotiating a meaningful focus for the therapeutic work. Further, the therapist should be relatively transparent and model a warm, concerned, and empathetic group leader stance, which is later carried on into the group. Given the limited knowledge of mediators in group therapy, it seems important to perform additional studies with larger samples and more frequent process recordings of specified elements of the group process. The use of updated group relationship measures (i.e., The Group Questionnaire; Krogel et al., 2013), specific measures related to psychodynamic theory (e.g., defense mechanisms and insight concerning intrapsychic conflicts and dysfunctional relational patterns), and use of newer and more advanced statistical methods (Lo Coco et al., 2012; Kivlighan & Kivlighan, 2013) may also provide more answers regarding mechanisms of change. An important extension of carefully designed RCTs is to combine them with qualitative studies based on detailed process observations.

## References

- Alden, L. E., Wiggins, J. S., & Pincus, A. L. (1990). Construction of circumplex scales for the Inventory of Interpersonal Problems. *Journal of Personality Assessment*, 55(3–4), 521–536. <http://dx.doi.org/10.1080/00223891.1990.9674088>
- Bakali, J. V., Baldwin, S. A., & Lorentzen, S. (2009). Modeling group process constructs at three stages in group psychotherapy. *Psychotherapy Research*, 19, 332–343. <http://dx.doi.org/10.1080/10503300902894430>
- Baldwin, S. A., Murray, D. M., & Shadish, W. R. (2005). Empirically supported treatments or type I errors? Problems with the analysis of data from group-administered treatments. *Journal of Consulting and Clinical Psychology*, 73, 924–935. <http://dx.doi.org/10.1037/0022-006X.73.5.924>
- Beutel, M. E., Knickenberg, R. J., Krug, B., Mund, S., Schattenburg, L., & Zwerenz, R. (2006). Psychodynamic focal group treatment for psychosomatic inpatients—With an emphasis on work-related conflicts. *International Journal of Group Psychotherapy*, 56, 285–305. <http://dx.doi.org/10.1521/ijgp.2006.56.3.285>
- Blackmore, C., Tantam, D., Parry, G., & Chambers, E. (2011). Report on a systematic review of the efficacy and clinical effectiveness of group analysis and analytic/dynamic group psychotherapy. *Group Analysis*, 45, 46–69. <http://dx.doi.org/10.1177/0533316411424356>
- Budman, S. H., Soldz, S., Demby, A., Feldstein, M., Springer, T., & Davis, M. S. (1989). Cohesion, alliance and outcome in group psychotherapy. *Psychiatry*, 52, 339–350. <http://dx.doi.org/10.1080/00332747.1989.11024456>
- Burlingame, G. M., Fuhrman, A., & Johnson, J. E. (2002). Cohesion in group psychotherapy. In J. C. Norcross (Ed.), *Psychotherapy relationships that work* (pp. 71–87). New York, NY: Oxford University Press.
- Burlingame, G. M., McClendon, D. T., & Alonso, J. (2011). Cohesion in group therapy. *Psychotherapy: Theory, Research, & Practice*, 48, 34–42. <http://dx.doi.org/10.1037/a0022063>
- Burlingame, G. M., MacKenzie, K. R., & Strauss, B. (2004). Small-group treatment: Evidence for effectiveness and mechanisms of change. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (pp. 647–696). New York, NY: Wiley.
- Burlingame, G. M., Seebeck, J. D., Janis, R. A., Whitcomb, K. E., Barkowski, S., Rosendahl, J., & Strauss, B. (2016). Outcome differences between individual and group formats when identical and non-identical treatments, patients, and doses are compared: A 25-year meta-analytic perspective. *Psychotherapy: Theory, Research, & Practice*, 53, 446–461. <http://dx.doi.org/10.1037/pst0000090>
- Burlingame, G. M., Strauss, B., & Joyce, A. S. (2013). Clinical mechanisms and effectiveness of small group treatment. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (pp. 640–689). New York, NY: Wiley.
- Crowe, T. P., & Grenyer, B. F. S. (2008). Is therapist alliance or whole group cohesion more influential in group psychotherapy outcomes? *Clinical Psychology & Psychotherapy*, 15, 239–246. <http://dx.doi.org/10.1002/cpp.583>

- Derogatis, L. R. (1977). *The SCL-90-R: Administration, scoring, and procedures. Manual I*. Baltimore, MD: Clinical Psychometric Research Unit, Johns Hopkins University School of Medicine.
- Dinger, U., & Schauenburg, H. (2010). Effects of individual cohesion and patient interpersonal style on outcome in psychodynamically oriented inpatient group psychotherapy. *Psychotherapy Research, 20*, 22–29. <http://dx.doi.org/10.1080/10503300902855514>
- Endicott, J., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry, 33*, 766–771. <http://dx.doi.org/10.1001/archpsyc.1976.01770060086012>
- First, M. B., Gibbons, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). *Structural Clinical Interview for DSM-IV Axis I Personality Disorders (SCID-II)*. Washington, DC: American Psychiatric Press, Inc.
- Fjeldstad, A., Høglend, P., & Lorentzen, S. (2016). Presence of personality disorder moderates the long-term effects of short-term and long-term psychodynamic group therapy: A 7-year follow-up of a randomized clinical trial. *Group Dynamics, 20*, 294–309. <http://dx.doi.org/10.1037/gdn0000055>
- Foulkes, S. H., & Anthony, E. J. (1984). *Group psychotherapy: The psychoanalytical approach*. London, UK: Maresfield reprints.
- Gallagher, M. E., Tasca, G. A., Ritchie, K., Balfour, L., & Bissada, H. (2014). Group attachment anxiety moderates the relationship between growth in group cohesion and treatment outcomes in group psychodynamic interpersonal psychotherapy for women with binge eating disorder. *Group Dynamics, 18*, 38–52. <http://dx.doi.org/10.1037/a0034760>
- Gallagher, M. E., Tasca, G. A., Ritchie, K., Balfour, L., Maxwell, H., & Bissada, H. (2014). Interpersonal learning is associated with improved self-esteem in group psychotherapy for women with binge eating disorder. *Psychotherapy: Theory, Research, & Practice, 51*, 66–77. <http://dx.doi.org/10.1037/a0031098>
- Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureño, G., & Villaseñor, V. S. (1988). Inventory of interpersonal problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology, 56*, 885–892. <http://dx.doi.org/10.1037/0022-006X.56.6.885>
- Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy: Theory, Research, & Practice, 48*, 9–16. <http://dx.doi.org/10.1037/a0022186>
- Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge.
- 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*, 269–296. <http://dx.doi.org/10.1521/ijgp.2007.57.3.269>
- Kipnes, D. R., Piper, W. E., & Joyce, A. S. (2002). Cohesion and outcome in short-term psychodynamic groups for complicated grief. *International Journal of Group Psychotherapy, 52*, 483–509. <http://dx.doi.org/10.1521/ijgp.52.4.483.45525>
- Kivlighan, D. M., Jr., & Kivlighan, D. M., III. (2013). Group climate research: Where do we go from here? *International Journal of Group Psychotherapy, 63*, 419–431. <http://dx.doi.org/10.1521/ijgp.2013.63.3.419>
- Krogel, J., Burlingame, G., Chapman, C., Renshaw, T., Gleave, R., Beecher, M., & Macnair-Semands, R. (2013). The Group Questionnaire: A clinical and empirically derived measure of group relationship. *Psychotherapy Research, 23*, 344–354. <http://dx.doi.org/10.1080/10503307.2012.729868>
- Lese, K. P., & MacNair-Semands, R. R. (2000). The Therapeutic Factors Inventory: Development of a scale. *Group, 24*, 303–317. <http://dx.doi.org/10.1023/A:1026616626780>
- Lo Coco, G., Gullo, S., & Kivlighan, D. M., Jr. (2012). Examining patients' and other group members' agreement about their alliance to the group as a whole and changes in patient symptoms using response surface analysis. *Journal of Counseling Psychology, 59*, 197–207. <http://dx.doi.org/10.1037/a0027560>
- Lorentzen, S. (2014). *Group analytic psychotherapy: Working with affective, anxiety, and personality disorders*. New York, NY: Routledge.
- Lorentzen, S., Ruud, T., Fjeldstad, A., & Høglend, P. (2013). Comparison of short- and long-term dynamic group psychotherapy: Randomised clinical trial. *The British Journal of Psychiatry, 203*, 280–287. <http://dx.doi.org/10.1192/bjp.bp.112.113688>
- Lorentzen, S., Ruud, T., Fjeldstad, A., & Høglend, P. A. (2015). Personality disorder moderates outcome in short- and long-term group analytic psychotherapy: A randomized clinical trial. *British Journal of Clinical Psychology, 54*, 129–146. <http://dx.doi.org/10.1111/bjc.12065>
- 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*, 33–40. <http://dx.doi.org/10.1080/08039480310000770>
- MacKenzie, K. R. (1983). The clinical application of a group climate measure. In R. R. Dies & K. R. MacKenzie (Eds.), *Advances in group psychotherapy: Integrating research and practice* (pp. 159–170). Madison, WI: International Universities Press.
- 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, 147–156.
- MacKenzie, R. K. (1997). *Time-managed group psychotherapy: Effective clinical applications*. Washington, DC: American Psychiatric Press.
- MacKenzie, R. K. (1998). The alliance in time-limited group psychotherapy. In J. D. Safran & J. C. Muran (Eds.), *The therapeutic alliance in brief psychotherapy* (pp. 193–215). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10306-008>
- Ogrodniczuk, J., & Piper, W. E. (2003). The effect of group climate on outcome in two forms of short-term group therapy. *Group Dynamics*, 7, 64–76. <http://dx.doi.org/10.1037/1089-2699.7.1.64>
- Ogrodniczuk, J. S., Piper, W. E., & Joyce, A. S. (2005). The negative effect of alexithymia on the outcome of group therapy for complicated grief: What role might the therapist play? *Comprehensive Psychiatry*, 46, 206–213. <http://dx.doi.org/10.1016/j.comppsy.2004.08.005>
- Ogrodniczuk, J. S., Piper, W. E., & Joyce, A. S. (2006). Treatment compliance among patients with personality disorders receiving group psychotherapy: What are the roles of interpersonal distress and cohesion? *Psychiatry*, 69, 249–261. <http://dx.doi.org/10.1521/psyc.2006.69.3.249>
- Piper, W. E., Marrache, M., Lacroix, R., Richardsen, A. M., & Jones, B. E. (1983). Cohesion as a basic bond in groups. *Human Relations*, 36, 93–108. <http://dx.doi.org/10.1177/001872678303600201>
- Schulz, K. F., & Grimes, D. A. (2002). Blinding in randomised trials: Hiding who got what. *Lancet*, 359, 696–700. [http://dx.doi.org/10.1016/S0140-6736\(02\)07816-9](http://dx.doi.org/10.1016/S0140-6736(02)07816-9)
- Sheehan, D., Janavs, J., Baker, R., Harnett-Sheehan, K., Knapp, E., & Sheehan, M. (2002). *Mini international neuropsychiatric interview*. Tampa, FL: Tampa University of South Florida Press.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420–428. <http://dx.doi.org/10.1037/0033-2909.86.2.420>
- Stekhoven, D. J., & Bühlmann, P. (2012). MissForest—Non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 28, 112–118. <http://dx.doi.org/10.1093/bioinformatics/btr597>
- Strauss, B., Burlingame, G. M., & Bormann, B. (2008). Using the CORE-R battery in group psychotherapy. *Journal of Clinical Psychology*, 64, 1225–1237. <http://dx.doi.org/10.1002/jclp.20535>
- Sun, G. W., Shook, T. L., & Kay, G. L. (1996). Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *Journal of Clinical Epidemiology*, 49, 907–916. [http://dx.doi.org/10.1016/0895-4356\(96\)00025-X](http://dx.doi.org/10.1016/0895-4356(96)00025-X)
- Tasca, G. A., Balfour, L., Ritchie, K., & Bissada, H. (2007). The relationship between attachment scales and group therapy alliance growth differs by treatment type for women with binge-eating disorder. *Group Dynamics*, 11, 1–14. <http://dx.doi.org/10.1037/1089-2699.11.1.1>
- Tschuschke, V., & Dies, R. R. (1994). Intensive analysis of therapeutic factors and outcome in long-term inpatient groups. *International Journal of Group Psychotherapy*, 44, 185–208. <http://dx.doi.org/10.1080/00207284.1994.11490742>
- Waljee, A. K., Mukherjee, A., Singal, A. G., Zhang, Y., Warren, J., Balis, U., . . . Higgins, P. D. (2013). Comparison of imputation methods for missing laboratory data in medicine. *British Medical Journal Open*, 3(8), e002847. <http://dx.doi.org/10.1136/bmjopen-2013-002847>
- Yalom, I. D., & Leszcz, M. (2005). *The theory and practice of group psychotherapy*. New York, NY: Basic Books.

Received February 19, 2017

Revision received April 24, 2018

Accepted April 24, 2018 ■