

ABSTRACT

Title of thesis: THE DEVELOPMENT AND VALIDATION OF THE GROUP
LEADER INTERVENTION SCALES

Karen L. Nuijens, Master of Arts, 2004

Thesis advisor: Dr. Hedwig Teglassi-Golubcow, Professor,
Department of Counseling and Personnel Services

The purpose of the study was to develop and validate an instrument for examining therapist interventions in group counseling settings. The Group Leader Intervention Scales (GLIS) is a verbal response mode system designed to code group leader verbalizations on six group process variables (structure, group cohesion, modeling, information, exploration, and feedback), and on eleven subscales that examine various aspects of the main categories. The GLIS was developed through a content analysis of session transcripts from a group intervention for children that used stories and peer group processes to increase social problem-solving. High levels of interrater reliability were established between three raters for the six group process variables, and for nine of the eleven subscales. Initial validity of the new instrument was demonstrated by its ability to distinguish between groups based on treatment response (high or low cognitive treatment response), and based on stage of treatment (early, middle, or late).

THE DEVELOPMENT AND VALIDATION OF THE GROUP LEADER
INTERVENTION SCALES

by

Karen L. Nuijens

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Arts
2004

Advisory Committee:

Dr. Hedwig Teglasi-Golubcow, Chair
Dr. Ruth Fassinger
Dr. Dennis M. Kivlighan, Jr.

© Copyright by
Karen L. Nuijens
2004

TABLE OF CONTENTS

Introduction	1
Background of the Study	
Purpose of the Study	
Significance of the Study	
Overview of the Data Set	
Predictions	
Literature Review	9
Chapter Overview	
History of Group Process Research	
Group Leader Interventions	
STORIES Program	
Rationale for the Development of the Group Leader Intervention Scales	
Methodology	30
Instrument	
Data Set	
Dividing Sessions into Thought Units	
Establishing the Interrater Reliability of the GLIS	
Establishing the Validity of the GLIS	
Results	45
Interrater Reliability	
Overview of the Chi Square Analyses	
Treatment Response Analyses	
Stage of Group Analyses	
Number of Interrupted Interventions	
Discussion	71
Predicted Findings	
Additional Findings	
Limitations of the Study	
Future Directions	
Appendix A: Group Leader Intervention Scales Coding Manual	86
Appendix B: Frequencies and Within-Category Percentages for High and Low Treatment Response Groups across Sessions.....	95
References	99

LIST OF TABLES

1.	Definition of Variables Used in the Predictions	5
2.	Systems for Studying Interventionist Skills/Intentions.....	23
3.	Overview of the Group Process Scale	32
4.	Group Descriptions According to Treatment Response	38
5.	Counseling Sessions Chosen According to Group	39
6.	Sample Coding Sheet for the Group Process Scale	40
7.	Reliability across Raters	46
8.	Number of Interventions across the Main Categories of the GLIS.....	48
9.	Number of Interventions across Treatment Response and Stage of Group....	49
10.	Main Categories of the GLIS across Level of Treatment Response	50
11.	Type of Structure across Level of Treatment Response	50
12.	Type of Structure across High Treatment Response Groups	51
13.	Level of Treatment Response across Type of Group Cohesion	51
14.	Level of Treatment Response across Type of Modeling	52
15.	Initiation of Information across Level of Treatment Response	52
16.	Topic of Exploration across Level of Treatment Response	53
17.	Topic of Exploration across High and Low Treatment Response.....	53
18.	Initiation of Exploration across Level of Treatment Response	55
19.	Valence of Feedback across Level of Treatment Response	56
20.	Type of Feedback across Level of Treatment Response	57
21.	Type of Feedback across High Treatment Response Groups	57
22.	Direction of Interventions across Level of Treatment Response.....	58
23.	Main Categories of the GLIS across Group Stage	59
24.	Type of Structure across Group Stage	60
25.	Type of Group Cohesion across Group Stage	60
26.	Type of Modeling across Group Stage	61
27.	Initiation of Information across Group Stage.....	61
28.	Topic of Exploration across Group Stage	62
29.	Initiation of Exploration across Group Stage	63
30.	Valence of Feedback across Group Stage	63
31.	Type of Feedback across Group Stage	64
32.	Direction of Intervention across Group Stage	64
33.	Number of Interrupted Interventions per Session	65
34.	Summary of Predicted Findings.....	66
35.	Summary of Additional Findings.....	69

INTRODUCTION

Background of the Study

Groups provide a natural setting for therapeutic work with children. Much of children's time is spent in groups (e.g., instruction, play, lunch and snack times) and group work is attractive to children since it closely mirrors their natural environment (Dwivedi, 1993). Group counseling also offers opportunities for peer influences that would be missing in an individual counseling situation. According to Dwivedi, effective group work with children has the potential to: enhance social skills, increase self-esteem and reality testing, teach how to delay gratification and manage feelings more appropriately, explain abstractions and values, and simultaneously increase interdependence with peers while further developing the skills necessary for increased autonomy.

Using stories as tools in a therapeutic setting is an especially effective strategy for addressing children's problems within school settings because it allows individuals to experience their issues through identification with characters' feelings and experiences in a non-threatening manner (Shechtman, 1999). This identification, in turn, allows individuals to reconnect with their own feelings and promotes catharsis.

The STORIES program (Teglasi & Rothman, 2001), a 15-week program designed to increase social problem-solving skills in elementary school children, utilizes stories to structure discussions and interactions among peers. Outcome results from the initial implementation of the program for two classrooms showed that children who completed the STORIES program had lower teacher-reported externalizing behavior and lower levels of beliefs that support aggression than did children who did not receive the

program. The program also has shown favorable changes in cognitive level functioning when compared to more skills oriented social competency programs (Rahill & Teglassi, 2003). These positive outcomes in both studies were predicted both by initial pre-test scores and by children's cognitive responses during the sessions. The group leader interventions have not yet been examined.

The initial focus of this thesis was to examine group leader behavior in STORIES groups in depth to more fully understand the therapeutic mechanisms underlying the success of the program. However, this researcher was unable to locate an established categorical system that would allow for an examination at the desired level of specificity. Bednar and Kaul (1994) speculated that outcome research continues to be prevalent over process research even after the efficacy of groups has been established partially because the conceptual and measurement tools required to ask more specific questions are not available. To obtain higher quality measurement instruments, they recommend observing, defining, and classifying central treatment elements using descriptive techniques.

Purpose of the Study

The purpose of the study was to develop a comprehensive instrument to examine group leader interventions within STORIES groups, and then to establish the interrater reliability of the new instrument and its validity in terms of distinguishing groups on the basis of treatment response (high or low cognitive treatment response), and in terms of distinguishing groups on the basis of group stage (early, middle or late).

The instrument that was developed, named the Group Leader Intervention Scales (GLIS), contains two main scales. The first scale, named the Group Process Scale, examines group process variables such as structure, group cohesion, modeling,

information, exploration and feedback. The second scale, named the Framework Building Scale, examines interventions directed at building a framework for more adaptive social problem-solving.

Significance of the Study

This study is significant in several ways:

1. It provides a measurement instrument for rating group leader interventions that was derived from examining groups with children. Existing measures have been developed almost exclusively from adult groups, or developed for coding individual counseling sessions and then adapted to measure groups.
2. It adds knowledge to the scarce literature currently available on group leader interventions specific to children's groups. While process research is fairly sparse across age groups, there is truly a paucity of process studies involving children's groups (Dagley, Gazda, Eppinger & Stewart, 1994; Hoag & Burlingame, 1997).
3. It supplements the outcome data previously gathered on the STORIES program. The increased accountability from funding sources necessitates that interventions demonstrate positive outcomes. However, in process research it is equally important to understand how and why improvement takes place, to clarify issues of generalizability for a specific treatment to other groups, to help prevent causal misattribution, and to examine aspects of therapy that may be impeding group member progress (Green, 2000).
4. It provides direction for training group therapists, particularly those who will run groups using the STORIES program. Manual guided therapies have become common largely because manuals allow consistent implementation of group programs

independent of therapist philosophy (Beutler, Machado, & Allstetter-Neufeldt, 1994).

The information gathered from this study will be valuable in the expansion of the STORIES program manual and can assist in specific feedback during training and periodic self-assessments of group leaders regarding therapeutic interventions.

Overview of the Data Set

Written transcripts from the original implementation of the STORIES program (Teglasi & Rothman, 2001) were used to establish the interrater reliability of the new instrument and the validity of the new instrument in terms of its ability to predict group differences and stage of session differences. Twelve groups participated in the original implementation of the program. Of the 12 groups, 4 groups were selected on the basis of their average cognitive response during the sessions (referred to as the treatment response).

Treatment response codes were given to individual group members by the group leader at the conclusion of 15 sessions. Treatment response codes also were given to individual group members by an independent rater, who coded group member verbalizations using session transcriptions then averaged these ratings across responses and sessions for individual group members. The percent agreement on treatment response coded between the group leader and the independent rater was 86.4%.

Treatment responses codes ranged from one to three. A score of one indicated a poor treatment response (many verbalizations that missed the mark and little or no improvement over sessions). A score of two indicated a moderate treatment response (few, but appropriate verbal responses and/or little improvement). A score of three

indicated good treatment response (high quality of responding throughout and/or improved quality of responses).

For the purposes of this project, the two groups with the highest average treatment response across members and the two groups with the lowest average treatment response across members were chosen for comparison of group leader interventions using the new instrument. Three sessions were coded for each of the four groups: one session that occurred early in the course of the program, one session that occurred in the middle of the program, and one session that occurred late in the program. Having the sessions evenly distributed across the course of the program allowed for predictions based on stage of group.

Several predictions were made to assess the validity of the GLIS in terms of its ability to distinguish groups on the basis of treatment response and on the basis of group stage (see Table 1 for brief definitions of the variables used in the predictions; see Appendix A for examples and for complete definitions of all of the variables contained in the GLIS):

Table 1
Definitions of Variables Used in the Predictions

VARIABLE	DEFINITION
Structure	Used to us manage the flow of sessions. Three types of structure are coded with the GLIS: long-term structure, routine management, or behavioral management (used to enforce a rule or redirect behavior).
Exploration	Inviting or engaging the group members to think about an idea, feeling or event in order to further clarify or extend the lesson beyond the readings or known facts. Interventions that are coded as exploration are also coded for whether the content of the intervention is focused on a reading, group-event, or group member-experience outside the group.

Feedback	Comments or reactions to a group member's idea, feeling or behavior that stemmed from the readings or an experience within or outside the group.
Valence of feedback	<ol style="list-style-type: none"> 1. Positive - explicit comment or reaction that indicates approval or acceptance of a group member's response. 2. Negative - comment or reaction that indicates disagreement or disapproval of a group member's response. 3. Neutral - a group member's response is reflected, repeated or acknowledged without an indication of acceptance or disagreement.
Types of feedback	<ol style="list-style-type: none"> 1. Simple acknowledgment or disagreement 2. Paraphrase or restatement - a group member's response is repeated or rephrased without changing the meaning of the statement or adding any additional information. 3. Reframing - a group member's response is altered to a more accurate or appropriate answer and/or false information is corrected. 4. Elaboration - a group member's response is extended or connected to an additional interpretation/explanation, but is not contradicted or altered.
Direction of interventions	<ol style="list-style-type: none"> 1. Directed toward the entire group - introduction of a new concept or providing an advanced organizer to the group that is not in response to an individual group member's statement or question or directed to a particular group member. 2. Directed toward an individual - direct reply to an individual group member's statement or question; calling on a particular member to answer a question, or providing an opinion to an individual member 3. Directed toward both - an intervention that was prompted by, or directed toward, an individual but generally applies to the whole group.
Interrupted interventions	When the group leader is interrupted before a thought was finished and he or she picks up on that same thought at the beginning of the next speaking turn, or following a behavioral management intervention within the same speaking turn.

*Predictions:**Structure:*

1. High treatment response groups will have proportionately fewer behavioral management interventions than the low treatment response groups.
2. In terms of group stage, sessions occurring in the early stage of the program will have proportionately more interventions focused on structure compared to middle and late stages.

Exploration:

3. In terms of group stage, the topic of exploration will differ. Specifically, fewer reading related interventions and more group-event related interventions are expected in later sessions.

Feedback:

4. In the valence of feedback category, high treatment response groups will have fewer negative feedback interventions than low treatment response groups.
5. In the type of feedback category, high treatment response groups will have more elaborative feedback and less reframing interventions compared to low treatment response groups.

Direction of Interventions:

6. High treatment response groups will have more interventions directed toward the whole group and more interventions directed at both (i.e., interventions that are prompted by, or directed toward, an individual but generally apply to the whole group), and fewer interventions directed toward individuals compared to low treatment response groups.

Interrupted Interventions:

7. Interrupted group leader interventions are more likely to occur in low treatment response groups than in high treatment response groups.

LITERATURE REVIEW

Chapter Overview

One of the challenges in planning a research project designed to examine group process, especially group process in children's groups, is the scarcity of literature specific to the topic. Hoag and Burlingame (1997) reported only five literature reviews prior to their review that specifically examined group psychotherapy with children and adolescents. They also point out that early reviews did little more than provide a catalog of studies and offered few, if any, overarching conclusions.

Additionally, studies that address the processes that occur within children's groups are much rarer than studies that examine outcomes of children's groups. In a review of the literature on group therapy with children, preadolescents and adolescents spanning 12 years, only 6 process studies were identified (Dagley, Gazda, Eppinger & Stewart, 1994). This researcher located an additional 8 process studies published between 1993 and 2003 focused on therapeutic groups with children. Of the 8 studies, 5 were conducted by Shechtman and colleagues.

Drawing generalizations across child, preadolescent, adolescent, and adult group process research is complicated by differences according to age, developmental level and type of treatment. For instance, maturational and developmental differences, particularly cognitive and emotional developmental stages, will influence an individual's ability to conceptualize problems, cultivate insight and develop empathy (Dwivedi, 1993). Modes of communication and capacity for concentration also have important implications regarding the type of treatment and group leader interventions that are likely to be effective with children and adolescents and are dependent on developmental stage. In

addition, comparing the process of children's groups with adult groups erroneously assumes that children's groups operate like adult groups when even type of treatment is likely to differ since group treatment with children often focuses on developmental or skill-based interventions rather than more traditional group therapy objectives.

Another complication within the group process literature is that most of the measurement instruments that have been used to examine processes in groups were originally developed for individual counseling settings. A potential danger in this practice involves the issue of whether counselor interventions are similar across individual and group settings. Although it is helpful to consider the process research in individual counseling, it is important to keep in mind that counselor interventions in groups must simultaneously meet the needs of multiple participants. Fuhriman and Burlingame (1994b), arguing that there are significant differences between the two modalities, point out that counselors in groups must be more flexible in intervention application strategies by simultaneously maintaining multiple relationships, being aware of the more complex balance of power and influence, being able to aim the focus of the session on content or process, and, at the same time, selecting an individual, interpersonal or group focus. Beck and Lewis (2000) similarly assert that understanding the group processes and influences that determine success is a difficult one because of the interpersonal relationships involved.

While the issue of whether counselor interventions are similar across individual and group settings is still being debated, Hill (1990b) pointed out more similarities than differences when comparing process in individual and group therapy sessions. For example, Hill stated that universality, role flexibility, vicarious learning and reenactment,

all of which are typically thought of as group counseling factors, operate within the individual settings as well but just differ in relative emphasis. This is supported by research comparing group and individual therapy of aggressive boys which displayed different patterns for process variables across group and individual treatments. For instance, Shechtman and Ben-David (1999) found that group therapists used more directives and less self-disclosure than individual therapists, and “experiencing” (defined as exploring feelings, reactions or behaviors) was more frequent within group treatment.

Further complicating this literature review is evidence that adult and children’s groups do not operate similarly when comparing group and individual treatments. For example, Shechtman (2003) used the Group Counseling Helpful Impacts Scales (Kivlighan, Multon & Brossart, 1996) to compare therapeutic factors within group and individual therapy settings of aggressive boys. Results of this study showed that three of the four therapeutic factors examined (relationship-climate, other versus self-focus, and problem definition-change) were similar across individual and group therapeutic settings. The only therapeutic factor examined that was different between group and individual treatment was emotional awareness-insight, which was more prevalent in group treatment. This is contrary to findings by Holmes and Kivlighan (2000), which used the same scale to compare the therapeutic factors between adult group and individual treatments. Results of the Holmes and Kivlighan study showed that relationship-climate and other versus self-focus are more prominent in group treatment, whereas emotional awareness-insight and problem definition-change are more central to the process of individual treatment.

In many regards, the previously discussed debates are irrelevant for this literature review. The literature specific to children's groups is so sparse that drawing tentative parallels between child, adolescent and adult group therapy literature is necessary. Likewise, the dearth of measurement instruments developed specifically to examine group leader interventions necessitates the examination of measurement instruments originally developed from individual counseling settings that have been subsequently adapted for group process research.

Taking the above factors into consideration, this chapter begins with an historic overview regarding the study of therapeutic groups. The chapter then turns to group leader intervention measurement and provides a rationale for the development of the Group Leader Intervention Scales (GLIS).

History of Group Process Research

Joseph Pratt, a Boston physician, is generally credited with initiating group psychotherapy in 1905 when he began a group for hospitalized tuberculosis patients (Horne & Rosenthal, 1997). The purpose of the group was not therapeutic in nature, but rather was designed to disseminate information about treatment in an efficient manner. In fact, Pratt referred to the format as the "class method." Although the groups were not formally studied, supportive effects of the method were recognized.

In 1907, Jesse B. Davis, a high-school principal in Grand Rapids, Michigan, formed the first group implemented formally in an educational setting (Horne & Rosenthal, 1997). Davis' groups were designed to teach educational and vocational guidance on a weekly basis. The same year, Frank Parsons, considered by many as the founder of modern counseling, developed career and vocational counseling in Boston as a

cost effective way of presenting information. Unfortunately, neither approach was studied empirically.

Also early in the nineteenth century, a psychiatrist named Jacob Moreno created a puppetry and drama group in a child guidance clinic in Vienna he called “psychodrama” (McGrath, 1993). By 1934, Moreno had moved to the United States and had begun working with 8-12 year olds in groups aimed at providing a supportive environment, improving ego strength and self-worth, as well as offering an opportunity to develop interests in “leisure activities” and to rebuild distorted personalities (Dwivedi, 1993). Moreno, who was the first to use the term “group therapy,” wrote the first book on group psychotherapy in 1932, founded the American Society for Group Psychotherapy and Psychodrama in 1942, and created the sociogram, which is still a popular means of mapping out significant likes, dislikes, and indifferences among group members.

Samuel Slavson, an engineer turned school teacher who was also interested in psychoanalysis, introduced “Activity Group Therapy” in 1934 (McGrath, 1993). Slavson, who was entirely self-trained in the areas of education and psychology, had a standardized approach for the groups that included the use of games, activities and snacks as incentives for participation. The groups consisted of 5-8 children and the group leader lead the group in an unstructured manner that allowed children to act spontaneously. Conflicts that arose were acted out and reflected upon. Slavson examined the interactions among group members descriptively and concluded that group work facilitated as much growth and change for children as did individual therapy (Horne & Rosenthal, 1997).

During the period of 1930-1945, school guidance became a greater focus and the power of groups was becoming increasingly recognized in education, psychiatry, social

work, psychology and religion (Horne & Rosenthal, 1997). Horne and Rosenthal, after reviewing the literature of the time period, concluded that research continued to be largely descriptive in nature and tried to recreate individual counseling in groups rather than focusing on the specific curative factors of groups. However, group leadership and membership factors were beginning to be addressed toward the end of this period. For instance, measurement of verbal and nonverbal behaviors occurring in response to specific actions of other members was beginning to appear.

During the 1950s, there was an explosion of research on a wide variety of group variables such as group structure, group climate, leadership variables, member characteristics, norms, goal setting and problem-solving (Horne & Rosenthal, 1997). Many credit the proliferation of the formal study of process in therapeutic settings to Carl Rogers, who in the 1940's, had persuasively argued about the benefits of recording therapy sessions for use in training and research (Hill, 1992). Experimental approaches during this time period tended to employ no treatment controls, although comparison with other treatments was becoming more common signaling a shift away from descriptive research toward more outcome-based research (Horne & Rosenthal, 1997).

During the 1960s, there was an overall decline in group research. Horne and Rosenthal (1997) ascribe this to social unrest in response to the Vietnam War, the women's movement, race relations, and the increased focus on individuals over the group. For example, "group think" became a popular term to express the phenomenon of thinking and behaving in ways that are influenced by group pressure and group expectation and a wide spread hostility grew regarding group pressure to conform. Wheelan (1994) also noted the decline in research during the 1960's and cited the

increasing popularity of career oriented specialties that did not fit the interdisciplinary nature of group research as another reason for the decline.

However, group research increased again during the 1970s and 1980s. An indication of this is that 20% of all counseling related articles during the 1970s involved groups (Horne & Rosenthal, 1997). Outcome research during this time period, compared to process research, continued to increase. Literature reviews consistently showed the superior results of group therapy compared to control groups, and comparable results of group therapy and individual therapy (Fuhriman & Burlingame, 1994a). In addition, during the 1980s, there was an increased focus on treatment models for populations with specific diagnostic criteria (e.g., depression, eating disorders, bereavement). Interestingly, there was a push to identify curative factors in greater depth during this time period. For instance, researchers began to acknowledge that it may be necessary to compromise some of the methodological rigor that had become common with the prevalence of outcome studies in exchange for studying factors that are more relevant to practice (Horne & Rosenthal, 1997).

Group Leader Interventions

Process evaluation of groups can encompass three components: the leader, the members and/or the interaction between them (Trotzer, 1999). Specific goals for evaluation include determining leader intentions and methods, determining member attitudes and characteristics, and identifying the focus and themes of group interaction in relation to the goals of the group. Variables used to study process range from discrete variables to abstract, complex variables (Marmar, 1990). Discrete variables have the advantage of being easier to quantify and tend to have high interrater reliabilities. More

abstract or complex variables, such as rating scales, have the advantage of being more sophisticated yet often are more difficult to operationalize and have lower interrater reliabilities.

Green (2000) asserts that the dimension of process research that has captured the greatest empirical attention in recent years is leader behavior. For instance, what leaders say, think about and values they express have become popular topics to research. The focus on therapist behavior, according to Green, reflects a “back to basics quality” (p.32). Many of the more current studies are qualitative in nature and apply descriptive level statistics of behavioral categories across the course of therapy. This is in line with Bednar and Kaul’s (1994) recommendation discussed previously that astute observation and careful description must occur prior to more extensive experimentation.

Despite the frequency of studies that have addressed therapist effects, Fuhriman and Burlingame (1994b) state that conclusions regarding leader interventions are often based on secondary or post hoc findings. Similarly, Riva and Smith (1997) state that group leader effects are rarely studied for particular characteristics, training or interventions, which makes it difficult to know how much of the treatment success or failure is the result of a particular leader.

In the broadest sense, the role of the therapist in a group setting is to provide a meaningful framework for therapeutic change and to provide an environment conducive for this change to occur. Providing a meaningful framework entails the development of group rules, norms and outcome expectations, as well as providing the appropriate structure for group development (Dies, 1994). Providing a positive climate for therapeutic change largely entails the establishment of a high quality therapist-client

relationship. For instance, a positive view of therapist has been consistently associated with client improvement. In addition, individuals who view their group as successful often refer to client-to-client interactions, whereas dissatisfied members are more likely to place blame with the therapist who lead the group.

Similarly, Yalom (1995) proposed that a therapist has three main functions when facilitating a group: (a) creation and maintenance of the group, (b) culture building, and (c) activation and illumination of the here and now. Yalom also asserted that no counseling technique can take precedence over a therapist coming across as concerned, empathetic, genuine, and accepting of group members. The therapist also has a duty to provide summaries and promote work on themes that carry over from session to session. Playing the “group historian” in this fashion increases continuity between sessions and creates a strong group according to Yalom.

Categories of process measures that can be studied include: (a) direct, or within the session, versus indirect, which generally tap global experience through questionnaire administration prior to or after sessions; (b) an examination of perspective, which can occur from the vantage point of the therapist, client or non-participant judge; (c) an exploration of the focus of the session; (d) examination of style, which could include whether interventions were implemented empathically or judgmentally; (e) exploring quality of intervention, or how well it is said or done; and (f) an examination of an aspect of process (Lambert & Hill, 1994). Aspects of process could include a content component (what is said or meant) or an action component (what is done in sessions such as question, self-disclosure). According to Lambert and Hill, action has been studied most often and content has been studied least often.

Hill (1991) describes several types of behavior in counseling settings that can be studied in ascending complexity from observable and discrete behaviors to more abstract behaviors that occur over a longer period of time: (a) ancillary behaviors, such as speech dysfluency and nonverbal behavior; (b) verbal behaviors, such as response modes and client experiencing; (c) covert behaviors, such as therapist intentions and client reactions; (d) content; (e) strategies, such as empty chair technique and analysis of transference; (f) interpersonal manner, such as therapist empathy and client involvement; and (g) therapeutic relationship, such as working alliance, transference/countertransference issues and interpersonal transactions. According to Hill, existing research tends to focus on verbal behavior, such as therapist interventions within sessions.

The most common method to evaluate therapist interventions involves nominal category response systems (Hill, 1990a). Response modes, as they are commonly called, refer to the grammatical structure of therapists' verbal response independent of content or topic. Over 30 nominal category systems have been developed to measure response modes. The large number of nominal category systems that are present is linked to process researchers' tendency to create new scales, or to modify existing scales, according to the goals of a particular study. However, little construct validity information is available for the nominal category systems. In an attempt to unify findings from the existing measures, six developers of therapist rating scales originally designed for individual therapy settings collaborated to rate a common set of individual therapy sessions conducted with adults (Elliott, et al., 1987). Evidence for convergent and discriminant validity was obtained for six common response modes that occurred across

the measures. The six common response modes were question, advisement, information, reflection, interpretation, and self-disclosure.

One of the measures examined in Elliott et al. (1987) was Hill's Counselor Verbal Response Mode System (1978, 1986). The Hill Counselor Verbal Response Mode System is one of four systems Hill developed to examine within group processes (Hill, 1992). The other three systems measure therapist intentions, client reactions, and client behaviors. When used in conjunction with one another, the four systems describe the interaction of both overt and covert behaviors of therapists and clients. Hill describes the process model examined by these measures as a series of continually evolving interactions. First, the therapist develops an intention for the impact he or she would like to have based on theory and clinical observation. Next, a response mode is chosen to implement the intention. The client's reaction to the intervention determines his or her response to the therapist, which in turn, influences selection of the next intention and response mode of the therapist.

Although the Hill Verbal Response Modes System was originally developed through examination of individual therapy with adults, Leichtentritt and Shechtman (1998) used an adapted version of the system to examine therapist, trainee, and participant verbal responses and their development over time in a group therapy setting with children. The authors were particularly interested in how therapists' verbal response modes related to group members' self-disclosure. Therapist response modes that were examined in this study were encouragers, feedback (replaced interpretives from the original system), directives, questions, paraphrase and self-disclosure. Results showed that therapists played the most active role in the group and that they employed a wide

array of responses, especially questions, self-disclosure, feedback and encouragers. In addition, structured activities and questions led to the most self-disclosure by children. The only examined category children used more than group leaders was self-disclosure.

Shechtman and Ben-David (1999) used a similarly adjusted version of Hill's counselor and client verbal response mode systems to compare outcomes and processes between group and individual therapy for aggressive children. The therapist response modes examined in the study were encouragers, directives, questions, paraphrases, interpretives and self-disclosure. Client responses that were examined included asking for advice, experiencing, insight, future goals, simple responses, and therapist and client alliance. Responses that could not be categorized within these categories were not included in analyses. Results found both individual and group counseling formats effectively reduced aggression in children as measured by an abbreviated teacher-reported version of the Child Behavior Checklist. Although the outcome was similar for the two treatment formats, they differed in terms of response modes. For instance, group therapists used more directives and less self-disclosure than individual therapists, and "experiencing" (defined as exploring feelings, reactions or behaviors) was more frequent within group treatment.

Shechtman and Yanov (2001) studied group leaders' verbal responses in children's groups in relation to the productivity of group members' responses. Three group leader verbal response modes from Hill's system utilized in this study were: (a) feedback, defined as direct and honest personal reaction to another person based on observational behavior; (b) confrontations, defined as interventions that focus on incongruity in client verbal or nonverbal behavior; and (c) interpretives, defined as

explanations of one's thoughts, feelings, or behaviors. Children's responses immediately following one of three verbal responses were analyzed through the Client Behavior System (Hill & O'Brien, 1999) and were coded as either unproductive responses (resistance, agreement, request, recounting) or productive responses (cognitive exploration, emotional exploration, insight and change). Of the three leader responses, feedback produced the highest rate of productive responses among group members.

The Hill Counselor Verbal Response Category System has recently been updated and renamed it the Helping Skills System (Hill & O'Brien, 1999). The skills examined in the revised system are conceptually the same as previous versions of the system, but include updated definitions and illustrative examples. The eleven skills included in the Helping Skills System are approval and reassurance, closed question, open question, restatement, reflection of feelings, challenge, interpretation, self-disclosure, immediacy, information and direct guidance. It is important to keep in mind that, although the previous studies cited in this literature review show that the system can be adapted for children's therapeutic groups, the system was originally developed by examining adults in individual treatment.

McLeod (1994) argues that it is more informative to measure leader intentions rather than the specific skills that are used. An example of this is the Helper Intention List by Hill and O'Grady (1985; reprinted in Hill & O'Brien, 1999). The Helper Intention List consists of nineteen intentions: set limits, get information, give information, support, focus, clarify, instill hope, encourage catharsis, identify maladaptive cognitions, identify maladaptive behaviors, encourage self-control, identify and intensify feelings, promote insight, promote change, reinforce change, deal with resistance, challenge, deal with the

therapeutic relationship and relieve helper's needs. This scale was also originally developed by examining adults in individual treatment.

Several other leader behavior scales specific to group work are available, but they are largely evaluative in nature and are mostly used to evaluate professional development. For instance, Trotzer's (1999) account of Group Leadership Skills divides skills into three categories: (a) reaction skills, which include active listening, restatement, reflection, clarifying and summarizing; (b) interaction skills, which include moderating, interpreting, linking, blocking, supporting, limiting, protecting, consensus taking; and (c) action skills, which include questioning, probing, tone setting, confronting, personal sharing and modeling. These skills serve as the building blocks, Trotzer argues, for group leadership.

Similar to Trotzer's system, Corey and Corey (2002) created the Self-Assessment of Group Leadership Skills, which is designed to help leaders assess the following skills: active listening, reflecting, summarizing, facilitating, empathizing, interpreting, questioning, linking, supporting, confronting, blocking, diagnosing, modeling, suggesting, initiating, evaluating and termination (see Table 2 for a comparison of the systems).

Table 2
Systems for Studying Interventionist Skills/Intentions

Helping Skills System (Hill & O'Brien, 1999)	Helper Intention List (Hill & O'Brien, 1999)	Group Leadership Skills (Trotzer, 1999)	Self-Assessment of Group Leadership Skills (Corey & Corey, 2002)
1 Approval and reassurance	1 Set limits	<i>Reaction Skills:</i>	1 Active listening
2 Closed question	2 Get information	1 Active listening	2 Reflecting
3 Open question	3 Give information	2 Restatement	3 Clarifying
4 Restatement	4 Support	3 Reflection	4 Summarizing
5 Reflection of feelings	5 Focus	4 Clarifying	5 Facilitating
6 Challenge	6 Clarity	5 Summarizing	6 Empathizing
7 Interpretation	7 Instill hope	6 Tracking	7 Interpreting
8 Self-disclosure	8 Catharsis	7 Scanning	8 Questioning
9 Immediacy	9 Maladaptive cognitions	<i>Interaction Skills:</i>	9 Linking
10 Information	10 Maladaptive behaviors	8 Moderating	10 Confronting
a. About process of helping	11 Encourage self-control	9 Interpreting	11 Supporting
b. Facts, data or opinions	12 Identify and intensify feelings	10 Linking	12 Blocking
c. Feedback about the client	13 Promote insight	11 Blocking	13 Diagnosing
11 Direct guidance	14 Promote change	12 Supporting	14 Modeling
a. Process advisement	15 Reinforce change	13 Limiting	15 Suggesting
b. Directives	16 Deal with resistance	14 Protecting	16 Initiating
12 Other	17 Challenge	15 Consensus taking	17 Evaluating
	18 Therapeutic relationship	<i>Action Skills:</i>	18 Termination
	19 Helper's needs	16 Questioning	
		17 Probing	
		18 Tone setting	
		19 Confronting	
		20 Personal sharing	
		21 Modeling	

While the increased attention being shown to the study of process in children's groups is promising, the measures typically used were originally developed based on more traditional therapy goals, for adult populations, and for use in individual therapy settings. According to Bednar and Kaul (1994) using systems borrowed from individual counseling settings can be problematic because they do not clarify the uniqueness of group treatments, they limit the understanding of group dynamics, and they use units of analysis and observations not designed or derived from individuals in a group format. Perhaps most important, current measures of group process are inadequate for the original goal of the current study, which was to determine group processes most conducive to positive outcomes of the STORIES program, a school-based group intervention designed to increase children's social problem-solving.

The STORIES Program

The Structure/Themes/Open Communication/Reflection/ Individuality/ Experiential Learning/Social Problem-Solving (STORIES) program is designed to utilize the story form to increase social problem-solving skills in elementary school children (Teglasi & Rothman, 2001). Throughout the 15-week program, stories involving bullies, victims and bystanders are used as a platform to link cognitive, emotional, and behavioral components of social competence.

According to Crick and Dodge (1994), social information processing, also referred to as social problem-solving, is a series of steps that are reciprocally related to one another and includes the following six steps: (a) encoding of social cues, (b) interpretation of social cues, (c) clarification of goals, (d) determining possible responses,

(e) choosing a response, and (f) a behavioral enactment and subsequent evaluation of the response.

One of the challenges for sequentially taught social-skills programs is that several of the social problem-solving steps occur outside conscious awareness, and therefore, are not likely to be altered through didactic means alone (Teglasi & Rothman, 2001). The STORIES program, on the other hand, encourages reorganization of schemas through a more experiential-based approach, such as examining lessons learned from story books, sharing similar personal stories, and through activities. Using a bibliotherapy approach is an effective method when addressing social problem-solving steps because the steps are inherent to the story form (Teglasi, 2001). Most importantly, the story form connects these components in ways that correspond to experience.

The goal of schema-based instruction, or framework building, is to create and expand schemas in the domain where instruction has occurred (Marshall, 1995). This approach differs from more traditional teaching approaches because it deemphasizes the quantity of factual information that is learned and concentrates instead on the quality of the integration of those facts. Framework building emphasizes doing something with the information, or becoming an active problem-solver, as a way of appraising successful learning. In other words, the criterion for success is not just recall of learned information, but rather the ability to integrate and apply learned material to novel situations. Under schema theory, the student is an active learner and the teacher provides new information that is pertinent, explicitly pointing out its links to known information and providing understandable examples. Marshall refers to the teacher acting as a “tour guide” who is

not bound to a fixed route, but rather explores various areas for extended periods of time when the “tourists” want to examine one concept in greater depth.

The focus of framework building in the STORIES program involves the development of more sophisticated, adaptive schemas for social problem-solving. Themes and/or lessons from stories containing social problems children commonly encounter (e.g., bullying) are used as the vehicle to promote meaningful discussion. During the sessions, the social problem-solving steps are introduced and displayed. The steps are used as a way to organize and connect the external events of the story, inner world of the characters, and the connection between actions and consequences.

As stated in the introductory chapter, outcome results from the initial implementation of the program for two classrooms (one fourth-grade classroom and one fifth-grade classroom) showed that children who completed the STORIES program had lower teacher-reported externalizing behavior (Teglasi & Rothman, 2001). Furthermore, higher individual cognitive treatment responses, which were reliably determined by the group leader and an independent rater reviewing written transcripts, predicted lower teacher-rated externalizing behavior beyond the contribution of pretest scores. The program also promoted greater cognitive change for emotionally disabled children when compared to the Skills Streaming Program (Rahill & Teglasi, 2003).

The above links between the STORIES program and favorable outcome data provide encouraging information about the efficacy of the program. However, the reasons for the program’s effectiveness are not yet fully understood. Teglasi and Rothman (2001) argue that treatment integrity of the program includes how the “functional unit” is working (i.e., actions of the interventionist and responses of the targeted individuals). In

order to examine the functional unit, group leader interventions and group member responses must first be reliably identified. The focus of this study is on identifying the role the interventionist plays in this process.

Rationale for the Development of the Group Leader Intervention Scales

Bednar and Kaul (1994) argue that group researchers focus on experimentation before accurate description and precise measurement are in place. They state, “The reason we know so little about group process variables is that we devote so little time to clarifying their essential nature and meaning with astute observation and careful description” (p. 640).

Descriptive research is often seen as a primary source for generating insight into clinical hunches and change agents. Bergin and Garfield (1994) reported a growing endorsement of descriptive approaches and methodological pluralism for process research, and Meehl (1979; as cited in Bednar & Kaul, 1994, p.659) noted that, “observation and description seem to be more valued and better understood in the more well-developed physical sciences than the social sciences.” Similarly, Hill (1990a, p.289) stated that, “it is important to realize that most of our theories are the personal impressions and biases of gifted therapists rather than highly developed frameworks that integrate and explain a large body of empirical findings.”

Hill (1991) argues that we need to build on each other’s work in order to accumulate knowledge in the process area and that in order to do that we should use existing measures rather than create new ones. Hill acknowledged that existing measures are not going to fit data exactly, partially because individuals all view the topic slightly differently and focus on various components. Still, she recommends revising existing

measures to fit a particular study if necessary rather than create a new system of measurement.

There were several problems with following Hill's (1991) recommendation for the current study. First, there is a lack of existing systems that have been created specifically for examining children's groups. Lack of specific information about children's groups is problematic as such groups are likely to require different group leader interventions than adult groups (e.g., children's groups likely have more behavioral issues to attend to). Second, no single response mode system listed in Table 1 includes all of the process-based leader interventions that were identified through preliminary open coding of STORIES program transcripts. Third, and perhaps most problematic for understanding the change agent in the STORIES program, no current system includes a method for evaluating framework building of social problem-solving. Although social problem-solving components are well defined theoretically (e.g., Crick & Dodge, 1994; Lemerise & Arsenio, 2000), this researcher was unable to locate any systematic attempts to determine what aspects of social problem-solving are typically addressed during intervention.

Given these issues, it became clear that existing systems did not provide a method for exploring how group leader interventions targeted the central goal of the STORIES program and would not allow for examination of interventions at the level of specificity that was desired. In the end, this researcher and her advisor decided that developing a new set of scales through a content analysis of transcribed STORIES sessions would provide an instrument that would allow for a more accurate representation of the group leader interventions that are present within the program. The focus of this project,

therefore, shifted from the original goal of an in-depth examination of the processes within the STORIES program to developing a new measurement instrument designed to study group leader interventions, and then establishing the new instrument's interrater reliability and validity in terms its ability to distinguish groups on the basis of treatment response (high or low cognitive treatment response) and on the basis of group stage (early, middle, or late).

METHODOLOGY

The purpose of the current study was to develop a new instrument to code group leader interventions in a children's group counseling setting geared at increasing social problem-solving (i.e., the STORIES program), and to establish the validity of the new instrument. The study utilized written transcripts from four STORIES groups to examine: (a) the interrater reliability of the scales across three raters, (b) treatment response differences across high and low responding groups, and (c) stage of group differences.

Instrument

The scale development process began with a content analysis of group leader interventions within a STORIES-based counseling session. Open coding, or the process by which concepts are identified in data, and conceptual ordering (Strauss & Corbin, 1998), or the organizing of the concepts along a dimension according to a set of properties, were used until a series of nominal scales emerged. The scales are designed to code thought units, which are defined as verbalizations by the interventionist that together make a cohesive idea. The scales continued to be refined through repeated coding of practice sessions by this researcher and her advisor until the categories contained within the new instrument appeared exhaustive and well-defined.

From the content analysis process, two broad dimensions for examining group leader behavior emerged: (a) the Group Process Scale, which tracks interventions by the group leader such as the provision of structure, group cohesion, modeling, information, exploration and feedback; and (b) the Framework Building Scale, which tracks interventions by the group leader that are designed to increase children's social problem-

solving. Collectively these two broad dimensions, and their various subcategories, were named the Group Leader Intervention Scales (GLIS).

Each thought unit is assigned to one of six categories within the Group Process Scale, as well as to one of the respective subcategories within the larger categories (see Table 3 for an overview of the categories and subcategories contained in the Group Process Scale; see also Appendix A to view the GLIS Coding Manual, which contains a more detailed account of each category/subcategory than is provided here).

The Group Process Scale

1. Structure, defined as interventions that are used to manage the flow of sessions. If this category is selected, raters then chose between three types of structure: (a) long-term structure, or the provision of an advanced organizer for what to expect later in the session or for future sessions; (b) routine management, which includes redirecting comments or topics, and facilitating turn taking by responding to verbal or nonverbal initiatives; and (c) behavioral management, which includes interventions that enforce a rule or redirect behavior.
2. Group Cohesion, defined as efforts to engage members in the group and to foster a sense of group identity or belonging. If this category is selected, raters then chose between two types of group cohesion: (a) team building, which refers to creating an atmosphere or building traditions that lead individual group members to identify more closely as a team; and (b) emotional engagement, which refers to fostering an investment in relationships among team members and conveying the importance of each individual member to the group.

Table 3
Overview of the Group Process Scale

MAIN CATEGORY	SUBCATEGORIES
1. Structure	Types of structure: 1. Long-term 2. Routine management 3. Behavioral management
2. Group cohesion	Types of group cohesion: 1. Team building 2. Emotional engagement
3. Modeling	Types of modeling: 1. Self-disclosure 2. Interaction with others
4. Information	Types of information: 1. New, factual 2. Reason or explanation 3. Review or summary Initiation of information: 1. Spontaneously offered 2. Given in response to a group member's statement/question
5. Exploration	Topic of exploration: 1. Readings 2. Group event 3. Member experience Initiation of exploration: 1. Spontaneously generated 2. Given in response to a group member's statement/question Level of exploration: 1. Abstract 2. Concrete
6. Feedback	Valence of feedback: 1. Positive 2. Negative 3. Neutral Types of feedback: 1. Simple acknowledgement 2. Paraphrase or restatement 3. Reframing 4. Elaboration

3. Modeling, defined as the demonstration of how to perform an action or express an idea or emotion. If this category is chosen, raters then chose between two types of modeling: (a) self-disclosure, which refers to the sharing of a personal thought, feeling or experience with the group; and (b) interaction with others, which refers to the demonstration of prosocial interactions with group members.
4. Information, defined as providing new factual information to group members, providing a reason or explanation to group members, or reviewing previously covered information through progress checks or summaries. If this category is selected, raters would choose between whether the information was: (a) new factual information, (b) a reason or explanation, or (c) a review or summary of previous information. Raters would indicate whether the intervention was: (a) spontaneously offered by the group leader, or (b) given in response to a group member's statement or question.
5. Exploration, defined as inviting or engaging the group members to think about an idea, feeling or event in order to further clarify or extend the lesson beyond the readings or known facts. If this category is selected, raters would choose whether the exploration was: (a) spontaneously generated by the group leader, or (b) initiated by a group member's statement or question. In addition, raters would also specify whether the topic of exploration was focused on: (a) the readings, (b) a group event, or (c) a group member's experience outside the group. Finally, the rater would choose whether the intervention contained an: (a) abstract question, or (b) a concrete question.
6. Feedback, defined as comments or reactions to a group member's idea, feeling or behavior that stemmed from the readings or an experience within or outside the

group. If this category is selected, raters first chose whether the valence of the intervention was: (a) positive in that it indicates approval or acceptance; (b) negative in that it indicates disagreement or disapproval; or (c) neutral in that it provides neither acceptance nor disagreement. Next, raters chose from four types of feedback: (a) simple acknowledgement or disagreement; (b) paraphrase or restatement, where a group member's response is repeated or rephrased without changing the meaning of the statement or adding any additional information; (c) reframing, where a group member's response is altered to a more accurate or appropriate answer and/or false information is corrected; or (d) elaboration, where a group members response is extended to an additional interpretation/explanation.

The Direction of Interventions

An additional scale within the Group Process Scale examines the direction of the intervention. Following the coding of the Group Process Scale, all thought units are also coded according to whether they are directed toward the group, toward an individual, or toward both in that the intervention was prompted by an individual but generally applies to the whole group. Determining who the intervention is directed toward is a unique aspect of the GPS. According to Posthuma (1996), how interventions are directed has important implications for group functioning. For instance, asking questions to one individual may invite one-to-one interaction, while relating the intervention to the whole group may promote greater overall investment among group members.

The Framework Building Scale

The Framework Building Scale examines how group leader interventions foster the development of improved social-problem solving. The Framework Building Scale

follows the general layout of the social problem-solving components proposed by Crick and Dodge (1994), taking into account the role affect plays in the process (Lemerise & Arsenio, 2000). The focus of this scale is on identifying which of the social problem-solving steps is being addressed by the interventionist through exploration of story content or through the application of the steps to group member experiences.

While all interventions are coded on the Group Process Scale and its appropriate subcategories, only interventions geared at providing information, exploration and feedback are coded on the Framework Building Scale since they are the more cognitively-focused interventions.

The Framework Building Scale is comprised of five main categories:

1. Encoding, defined as helping group members take stock of, or clarify social cues. This includes drawing attention to facts, physical appearances and other relevant information from the environmental surroundings.
2. Interpretation, defined as helping group members understand or explore the reasons or causes of emotions, a behavior or sequence of behaviors or events, or the relationship among them.
3. The exploration and/or clarification of goals or intentions prior to taking action in a particular situation.
4. The formulation and evaluation of potential responses.
5. Discussion of the moral of a story or the lesson(s) learned.

As with the Group Process Scale, the Framework Building Scale is a mutually exclusive nominal scale, so for each thought unit eligible to be coded, the rater would first determine which of the above categories is being targeted through the intervention.

After the main category has been chosen, there are also two subscales to choose from. First, the rater determines whether the intervention was focused on: (a) internal factors, such as thoughts, feelings, goals, motives, intentions and physiological processes; (b) external factors, or observable behaviors, events or physical characteristics; or (c) an integrative focus, or a coordination of internal and external factors. Second, the rater would determine whether the content of the interventions was: (a) story-based, (b) experience-based, or (c) a combination of story-based content and experience-based.

Data Set

As previously stated, the study utilized verbatim transcriptions from the original implementation of the STORIES program (Teglasi & Rothman, 2001). Participants in the original study included 59 children (31 males, 28 females) from two intact classrooms (one fourth-grade, one fifth-grade) from a single county in Central Maryland. All but two of the participants were African-American. Parents of all students enrolled in the two classrooms signed permission forms for their children to participate in the program. Only first names are included in the session transcriptions.

School personnel, including past teachers, present teachers and administrators, identified children in the two classrooms according to previous disruptive and/or aggressive behaviors. Of the 59 children, 28% were identified as aggressive (11 males, 6 females). In addition, scores on the Behavioral Assessment System for Children (BASC; Reynolds & Kamphaus, 1992) showed that the identified children were in or near the clinical range for externalizing behavior according to current teacher ratings. Participants were divided into 12 groups consisting of 4-6 children per group. Identified children were

evenly distributed among groups and the non-identified children were randomly placed in groups.

All groups were led by the same leader. The group leader, a Caucasian female, was a school psychology doctoral student at the time. A co-leader (one of three school psychology graduate students) was also present during sessions in order to help the leader with activities and to reduce discipline issues (primarily through non-verbal means such as eye contact, sitting close to a disruptive member or between two disruptive members, etc). Co-leader interventions were not coded in the current study due to the very small role the co-leader plays in the STORIES program. For instance, there are whole sessions in which the co-leader never even makes a comment on the written transcripts.

The current study utilized data from 4 of the 12 groups. Groups were selected according to cognitive treatment response. Treatment responses codes ranged from one to three. A score of one indicated a poor treatment response (many verbalizations that missed the mark and little or no improvement over sessions). A score of two indicated a moderate treatment response (few, but appropriate verbal responses and/or little improvement). A score of three indicated good treatment response (high quality of responding throughout and/or improved quality of responses). As stated in the introduction, treatment response codes were given to individual group members by the group leader at the conclusion of treatment and by an independent rater, who coded group member verbalizations using session transcriptions. Ratings were averaged for individual group members (the interrater reliability between these sources was 86.4%). The two groups with the highest treatment response average and the two groups with the lowest treatment response average were selected for comparison in this study (see Table 4).

Table 4
Group Descriptions According to Treatment Response

<p><u>High treatment response groups:</u></p> <p>Group 1 (H1) = Treatment response (TR) average = 2.3 5th grade 3 males and 3 females - 2 identified males (TR average = 1.5) - 4 non-identified (TR average = 2.8)</p> <p>Group 2 (H2) = TR average = 2.6 5th grade 1 male and 5 females - 1 identified female (TR = 1) - 5 non-identified (TR average = 3)</p>
<p><u>Low treatment response groups:</u></p> <p>Group 1 (L1) = TR average = 1.7 4th grade 2 males and 2 females - 1 identified male (TR = 1) - 1 identified female (TR = 3) - 2 non-identified (TR average = 2.5)</p> <p>Group 2 (L2) = TR average = 1.5 4th grade 4 males - 2 identified males (TR average = 1) - 2 non-identified males (TR average 2)</p>

Twelve sessions were chosen for the study; three sessions from each group.

Sessions were chosen according to group stage so that one early treatment session, one middle treatment session and one late treatment session were chosen for each group. It is important to note that while all groups completed 15 sessions, and while all groups followed the same session outlines, not all sessions were transcribed due to audio tape difficulties (some of the tapes did not come out or were of poor quality and could not be transcribed). In order to minimize the effects that individual sessions may have had on

the findings, sessions that were available were counterbalanced across high and low treatment groups (see Table 5).

Table 5
Counseling Sessions Chosen According to Group

Group	Early Session	Middle Session	Late Session
H1	2 nd	7 th	12 th
H2	3 rd	6 th	13 th
L1	2 nd	7 th	12 th
L2	3 rd	6 th	13 th

Dividing Sessions into Thought Units

As stated earlier, a thought unit is defined as verbalizations by the interventionist that together make a cohesive idea. A thought unit may be one sentence, or it may be a group of related sentences. One speaking turn may contain several thought units or it may take several speaking turns to comprise a thought unit. See Appendix A for a detailed account of dividing sessions into thought units.

In order to ensure that the division of sessions into thought units was reliable, three of the 12 sessions were divided into thought units independently by this researcher and a fellow doctoral student. Following the division of each session into thought units, the two researchers reconciled discrepancies and clarified the rules contained in the coding manual. Percent agreement for the division of sessions into thought units across the three sessions was 87% (Session 1 = 82.5%; Session 2 = 86.9%; Session 3 = 91.5%).

Establishing the Interrater Reliability of the GLIS

The same three sessions that were jointly divided into thought units were coded using the Group Leader Intervention Scales. Three raters participated in coding the sessions: this researcher (Rater 1); her advisor (Rater 2); and a fellow doctoral student

(Rater 3). All three raters participated in roughly 20 hours of practice coding during the final revisions of the scale, during which time definitions were clarified and examples were added to the coding manual.

The following process was used to code all sessions:

1. The session was coded independently by all three raters using Microsoft Excel -based coding sheets (see Table 6). Raters indicated their choices by placing a “1” in one cell per category. All other cells were left blank.

Table 6
Sample Coding Sheet for the Group Process Scale

Thought Unit	Structure	Group Cohesion	Modeling	Information	Exploration	Feedback
1						
2						
3						
4						
5						
6						
7						

2. Rater 1 tallied the coded session across all combinations of raters (i.e., Raters 1 & 2; Raters 1 & 3, and Raters 2 & 3) using transparencies of the coding sheets.
3. Percent agreement across combinations of raters, and an overall percent agreement, was then calculated for each category.
4. Rater 1 created a list of thought units that contained discrepant ratings, including which raters disagreed and which categories were given. Discrepant ratings were

- defined as ratings in which all raters disagreed. If two out of three raters agreed, the code was not considered discrepant.
5. A summary of the percent agreement results and a list of all discrepancies were then sent via email to the other two coders. Discrepancies were reconciled and rules and/or definitions within the scales were clarified as necessary. Examples that helped clarify the new rules also were added to the coding manual at this time.
 6. Finally, Rater 1 sent the updated coding manual to the other two raters prior to coding the next session.

After the three sessions had been coded, the combined percent agreement across raters and sessions was calculated for each category. In addition, a kappa was calculated for each set of raters by category (Cohen, 1960). The kappa statistic is the most appropriate measure of interrater reliability since it reflects percent agreement that has been adjusted for the number of agreements that would have been expected by chance. The kappa for each set of raters was then averaged as recommended by Hill and O'Brien (1999).

In accordance with the recommendation made by Gardner (1995), only scales whose kappa's were .70 or higher were considered to be reliable and eligible for additional data analysis. Scales that were not found to be reliable were discontinued from the instrument at this time. Therefore, only scales that received a kappa of .70 or higher on the first three sessions were coded by Rater 1 on the remaining nine sessions.

A reliability check was performed by Rater 2 on the remaining nine sessions to help ensure that coding continued to be reliable throughout the study. One transcript page from each of the nine sessions was randomly selected to be coded by Rater 2. The

interrater reliability of these codes was compared with Rater 1's codes using the same process outlined previously.

Establishing the Validity of the GLIS

Chi-square analyses and an examination of the adjusted standardized residuals in each cell of the chi-square contingency tables were used to determine the validity of the scale in relation to the following predictions:

Predictions Regarding Structure:

1. High treatment response groups will have proportionately fewer behavioral management interventions than the low treatment response groups. According to the treatment response scores and to the group leader's qualitative notes, acting out behaviors were negatively related to degree of group engagement/cognitive responsiveness.
2. In terms of group stage, the literature predicts that sessions occurring in the early stage of the program will have proportionately more interventions focused on structure compared to middle and late stages (Dies, 1994; Yalom, 1995).

Prediction Regarding Exploration:

3. In terms of group stage, the topic of exploration will differ. Specifically, fewer reading related interventions and more group-event related interventions are expected in later sessions due to the format of the program. While reading story books takes place across sessions in the STORIES program, there tends to be less reading in later sessions and more group activities, such as art projects, designed to further explore previously covered concepts. Because of this, fewer reading

related interventions and more group-event related interventions in later sessions were expected.

Predictions Regarding Feedback:

4. In the valence of feedback category, high treatment response groups will have fewer negative feedback interventions than low treatment response groups. Since members of high treatment response groups are more likely to provide fairly accurate/plausible comments to begin with, interventions containing negative feedback are less likely.
5. In the type of feedback category, high treatment response groups will have more elaborative feedback and less reframing interventions compared to low treatment response groups. Members of high treatment response groups are more capable, cognitively and behaviorally, of receiving detailed elaboration of topics/previous comments. Additionally, since their initial contributions are likely to be fairly accurate to begin with, group leader responses would contain less reframing feedback.

Prediction Regarding the Direction of Interventions:

6. High treatment response groups will have more interventions directed toward the whole group and more interventions directed at both (i.e., interventions that are prompted by, or directed toward, an individual but generally apply to the whole group), and fewer interventions directed toward individuals compared to low treatment response groups. This prediction seemed likely since individual group members within low treatment response groups require more individual scaffolding and more behavioral interventions.

Prediction Regarding Interrupted Interventions:

7. Interrupted group leader interventions are more likely to occur in lower treatment response groups than in higher treatment response groups. This was a supplementary prediction and is not part of the GLIS. The tendency for low treatment response groups to interrupt the group leader was noted while coding the first two sessions selected for this study (one high treatment response session and one low treatment response session). Interruptions were evaluated by counting the number of thought units within transcripts that are linked together with a mark of “continued” following an interruption by a group member. Additional chi square analyses examining relationships among the remaining GLIS variables also were conducted for exploratory purposes.

Riva and Smith (1997) assert that it is not appropriate to combine data from members of different counseling groups even if they have the same leader because they can differ based on personality, interaction among members and other variables. Instead, they argue that separate groups should be the unit of measurement. To help ensure that significant findings were truly the result of the concepts within the predictions and were not simply due to individual group differences, high (H1 to H2) and low (H1 to H2) treatment groups also were compared to each other throughout the analyses. If chi squares analyses within high and low treatment response groups failed to find significant differences, the findings of the aggregated data were perceived to be strengthened. When in fact there were significant differences within high or low treatment groups, this researcher further explored why these group differences may have been present.

RESULTS

Interrater Reliability

The degree of interrater reliability was established across three raters for 3 of the 12 sessions (799 total thought units). Scales that did not obtain a minimum .70 kappa were discontinued from the instrument. Acceptable kappa levels were achieved for the main categories in the Group Process Scale (see Table 7). Acceptable kappa levels also were achieved for all of the subscales contained in the Group Process Scale except two: (a) type of information, and (b) level of exploration. These two scales were not examined further in the chi-square analyses and were removed from the final version of the GLIS. Two additional scales within the Group Process Scale, initiation of exploration and valence of feedback did not achieve an acceptable kappa when the three sessions were combined, but were kept in the final version of the GLIS since an acceptable kappa level was reached by the third session and on the reliability check performed between Rater 1 and Rater 2.

Regarding the Framework Building Scale, acceptable kappa levels were not achieved for the main categories or its two subcategories (see Table 7), and therefore, these scales also were not examined further in the chi-square analyses and were removed from the final version of the GLIS.

A reliability check between Rater 1 and Rater 2 was conducted during coding of the remaining nine sessions. For the reliability check, one transcript page was randomly selected to be coded by Rater 2 for each of the nine sessions for a total of 137 thought units. Results of the reliability check demonstrated that levels of interrater reliability had been maintained throughout the coding process (see Table 7).

Table 7
Reliability across Raters

Variable	Session 1 (1E)	Session 2 (3M)	Session 3 (3L)	Combined Sessions	Reliability Check by Coders 1 & 2
Main Categories in the Group Process Scale	91.5%	85.4%	87.4%	88.6% .85 kappa	88.3% .85 kappa
<u>Type of Structure</u> - Long-term - Routine management - Behavioral management	88.6%	83.6%	91.7%	87.7% .75 kappa	100% 1.0 kappa
<u>Type of Group Cohesion</u> - Team building - Emotional engagement	84.6%	100%	100%	94.4% .83 kappa	100% 1.0 kappa
<u>Type of Modeling</u> - Self-disclosure - Interaction w/other	91.1%	100%	100%	96.7% .93 kappa	100% 1.0 kappa
<u>Type of Information</u> - Factual information - Reason or explanation - Review information	58.2%	57.5%	76.0%	67.8% .49 kappa	Scale discontinued
<u>Initiation of Information</u> - Spontaneous - Member initiated	68.6%	96.1%	95.1%	89.2% .74 kappa	91.7% .83 kappa
<u>Topic of Exploration</u> - Readings - Group event - Generalize	85.4%	90.3%	89.3%	87.9% .76 kappa	85.7% .71 kappa

Variable	Session 1 (1E)	Session 2 (3M)	Session 3 (3L)	Combined Sessions	Reliability Check by Coders 1 & 2
<u>Initiation of Exploration</u> - Spontaneous - Member initiated	70.0%	88.3%	87.0% .71 kappa	79.7% .58 kappa	86.2% .71 kappa
<u>Level of Exploration</u> - Abstract question - Concrete question	73.8%	82.5%	74.9%	76.6% .52 kappa	Scale discontinued
<u>Valence of Feedback</u> - Positive - Negative - Neutral	69.7%	67.9%	92.5% .82 kappa	74.6 % .53 kappa	96.4% .93 kappa
<u>Type of Feedback</u> - Acknowledge - Paraphrase - Reframe - Elaborate	73.4%	85.3%	88.0%	80.5% .73 kappa	85.2% .80 kappa
Direction of Interventions	81.2%	77.9%	84.4%	80.9% .73 kappa	89.1% .83 kappa
Main Categories in the Framework Building Scale	67.2%	74.9%	76.2%	72.2% .59 kappa	Scale discontinued
Internal, external, or integrative focus	59.6%	76.6%	76.6%	69.7% .49 kappa	Scale discontinued
Story, experience-based, or combination	74.2%	82.4%	89.6%	81.2 % .67 kappa	Scale discontinued

Overview of Chi Square Analyses

The remainder of this chapter is devoted to Pearson chi square analyses. The analyses first examine how the GLIS variables varied according to treatment response, and then examine the GLIS variables according to group stage. In order to examine the data thoroughly, analyses within both sections begin by examining the main categories within the GLIS, and then work through each of the subscales, noting when findings pertain to specific predictions. In addition, when significant differences were found for high and low treatment response groups, additional analyses were performed to determine if the two high groups and/or if the two low groups were statistically different from one another. The chapter concludes with a summary of the predicted findings and a summary of the significant findings that were discovered through the exploratory analyses.

Across the 12 sessions included in the study, a total of 2,899 thought units were coded. The overall distribution across the main categories of the GLIS was significantly different than expected by chance ($\chi^2 = 761.58$, 5df, $p < .001$). Adjusted residuals showed that the interventionist utilized more feedback ($z = 10.9$, $p < .01$), more structure ($z = 10.0$, $p < .01$), and more exploration ($z = 5.21$, $p < .01$) than would be expected by chance. Additionally, there were fewer interventions focused on modeling ($z = -18.3$, $p < .01$) or on group cohesion ($z = -17.2$, $p < .01$) than would be expected by chance.

Table 8
Number of Interventions across the Main Categories of the GLIS

Structure	Group Cohesion	Model	Inform	Explore	Feedback	Total
800 27.6%	92 3.2%	73 2.5%	462 15.9%	640 22.1%	832 28.7%	2899 100%

Of the overall 2,899 thought units, 1367 thought units occurred in high treatment response groups and 1532 thought units occurred in low treatment response groups (see Table 9). This was a significant difference, with high treatment response groups receiving significantly fewer group leader interventions compared to low treatment response groups ($\chi^2 = -4.70$, 1df, $p < .05$; with Yates correction $\chi^2 = -4.60$, 1df, $p < 0.05$).

Table 9
Number of Interventions across Treatment Response and Stage of Group

Treatment Response	Early Sessions	Middle Sessions	Late Sessions	Total
High Groups	508 49.0%	457 46.0%	402 46.4%	1367 47.2%
Low Groups	531 51.1%	536 54.0%	465 53.6%	1532 52.8%
Total	1039 35.8%	993 34.3%	867 29.9%	2899

There also were differences in the overall number of interventions according to stage of group ($\chi^2 = 8.38$, 2df, $p < .05$). Adjusted standardized residuals showed that early sessions received more group leader interventions than expected by chance ($z = 2.01$, $p < .05$), and late sessions received fewer interventions than expected by chance ($z = -2.81$, $p < .01$). The chi-square distribution comparing number of interventions based on treatment response across stage of group was not significantly different than expected by chance.

Treatment Response Analyses

Main Categories of the GLIS

For the main categories of the GLIS, the distribution of interventions based on level of treatment response was not significantly different than expected by chance ($\chi^2 = 5.32$, 5df, $p < 1$; see Table 10). Directly comparing the two high treatment response groups with each other, and directly comparing the two low treatment response groups

with each other, did not show significant differences either, lending support that group differences within level of treatment response did not impact the overall finding.

Table 10
Main Categories of the GLIS across Level of Treatment Response

Treatment Response	Structure	Group Cohesion	Model	Inform	Explore	Feedback	Total
High Groups	363 26.6%	38 2.8%	34 2.5%	215 15.7%	300 21.9%	417 30.5%	1367 47.2%
Low Groups	437 28.5%	54 3.5%	39 2.5%	247 16.1%	340 22.2%	415 27.1%	1532 52.8%
Total	800 27.6%	92 3.2%	73 2.5%	462 15.9%	640 22.1%	832 28.7%	2899

Type of Structure

The distribution of interventions for type of structure based on level of treatment response was significantly different than expected by chance ($\chi^2 = 72.71$, 2df, $p < 0.001$; see Table 11). As predicted, adjusted standardized residuals show that high treatment response groups had proportionately fewer behavioral management interventions compared to low treatment response groups ($z = -8.5$, $p < .01$). Additionally, high treatment response groups had proportionately more routine management interventions compared to low treatment response groups ($z = 6.7$, $p < .01$).

Table 11
Type of Structure across Level of Treatment Response

Treatment Response	Long-term	Routine Management	Behavioral Management	Total
High Groups	30 8.3%	309 85.1%	24 6.6%	363 45.4%
Low Groups	23 5.3%	280 64.1%	134 30.7%	437 54.6%
Total	53 6.6%	589 73.6%	158 19.8%	800

It is important to note that the distribution comparing type of structure for the two high treatment response groups was significantly different than expected by chance ($\chi^2 = 9.01, 2df, p = 0.05$; see Table 12). Among the structure interventions, Group H1 had proportionately fewer routine management interventions than group H2 ($z = -2.41, p < .01$), and group H2 had relatively fewer behavioral management interventions than group H1 ($z = -2.92, p < .01$). In addition, overall, Group H1 ($n = 215$) had more structure interventions than group H2 ($n = 148$; $\chi^2 = 6.23, 1df, p = 0.05$; with Yates correction $\chi^2 = 6.23, 1df, p < 0.05$). The distribution comparing interventions for type of structure for the two low treatment response groups was not significantly different.

Table 12
Type of Structure across High Treatment Response Groups

Group	Long-term	Routine Management	Behavioral Management
H1	8.8% (19)	81.4% (175)	9.8% (21)
H2	7.4% (11)	90.5% (134)	2.0% (3)

Type of Group Cohesion

For type of group cohesion, the distribution of interventions based on level of treatment response was not significantly different than expected by chance ($\chi^2 = 1.55, 1df, p < 1$; with Yates correction $\chi^2 = 1.06, 1df, p < 1$; see Table 13).

Table 13
Level of Treatment Response across Type of Group Cohesion

Treatment Response	Team Building	Emotional Engagement	Total
High Groups	12 31.6%	26 68.4%	38 41.3%
Low Groups	24 44.4%	30 55.6%	54 58.7%
Total	36 39.1%	56 60.9%	92

Type of Modeling

For type of modeling, the distribution of interventions based on level of treatment response was not significantly different than expected by chance ($\chi^2 = 2.14$, 1df, $p < 1$; with Yates correction $\chi^2 = 1.51$, 1df, $p < 1$; see Table 14).

Table 14
Level of Treatment Response across Type of Modeling

Treatment Response	Self-Disclosure	Interaction with Others	Total
High Groups	18 52.9%	16 47.1%	34 46.6%
Low Groups	14 35.9%	25 64.1%	39 53.4%
Total	32 43.8%	41 56.2%	73

Initiation of Information

For initiation of information, the distribution of interventions based on level of treatment response was not significantly different than expected by chance ($\chi^2 = 1.82$, 1df, $p < 1$; with Yates correction $\chi^2 = 1.58$, 1df, $p < 1$; see Table 15).

Table 15
Initiation of Information across Level of Treatment Response

Treatment Response	Spontaneously-offered	Member-initiated	Total
High Groups	136 63.3%	79 36.7%	215 46.5%
Low Groups	141 57.1%	106 42.9%	247 53.5%
Total	277 60.0%	185 40.0%	462

Topic of Exploration

For topic of exploration, the distribution of interventions based on level of treatment response was significantly different than expected by chance ($\chi^2 = 15.32$, 2df, p

< 0.001; see Table 16). Adjusted standardized residuals show that low treatment response groups had proportionately more member-experience focused interventions than expected when compared to high treatment response groups ($z = 3.4, p < .01$).

Additionally, low treatment response groups had proportionately fewer reading focused interventions than expected when compared to high treatment response groups ($z = -3.7, p < .01$).

Table 16
Topic of Exploration across Level of Treatment Response

Treatment Response	Readings	Group-event	Member-Experience	Total
High Groups	173 57.7%	56 18.7%	71 23.7%	300 46.9%
Low Groups	146 42.9%	72 21.2%	122 35.9%	340 53.1%
Total	319 49.8%	128 20.0%	193 30.2%	640

It is important to note that the distribution comparing topic of exploration for the two high treatment response groups was significantly different than expected by chance ($\chi^2 = 24.09, 2df, p < 0.001$; see Table 17), and that the distribution comparing the low treatment response groups was significantly different than expected by chance as well ($\chi^2 = 36.34, 2df, p < 0.001$).

Table 17
Topic of Exploration across High and Low Treatment Response Groups

Group	Reading	Group-event	Member Experience
H1	71.1% (108)	10.5% (16)	18.4% (28)
H2	43.9% (65)	27.0% (40)	29.1% (43)

Group	Reading	Group-event	Member Experience
L1	47.2% (83)	8.5% (15)	44.3% (78)
L2	38.4% (63)	34.8% (57)	26.8% (44)

The differences within high and low groups are linked to session effects, and to a lesser extent, individual group differences. As previously described, sessions were counterbalanced across high and low treatment response groups. Sessions 2, 7, and 12 were coded for Group H1 and Group L1. Sessions 3, 6, and 13 were coded for Group H2 and Group L2. Group H2 had more interventions focused on group events compared to H1 ($z = 3.7, p < .01$). Group L2 also had more interventions focused on group events compared to L1 ($z = 5.9, p < .01$). Session 13 was coded as the late session for both H2 and L2. The focus of session 13 was an art activity, which necessitated more group event exploration (H2 had 30 group event related interventions in the late session and L2 had 36). The art activity in session 13 also helps to explain why H2 had relatively fewer interventions focused on the readings ($z = -4.8, p < .01$), although this finding was also impacted by H1's relatively high frequency of interventions focused on readings. In addition, Group L1 had significantly more interventions focused on member-experience than Group L2 ($z = 3.4, p < .01$). This finding is linked to Group L1's early session, which had an unusually high number of member-experience interventions ($n = 42$). Finally, Group H1 had relatively fewer interventions focused on member-experience ($z = -2.2, p < .01$), although there is no clear explanation for why this was the case.

The impact session effects had on this category can be seen more directly by comparing Group H1 with L1, and Group H2 with L2. When Groups H2 and L2 are compared, there are no significant differences. When Groups H1 and L1 are compared,

there are two significant group differences ($\chi^2 = 25.27$, 2df, $p < 0.001$). First, Group L1 had relatively more interventions focused on member-experience than Group H1 ($z = 5.0$, $p < .01$), which can be accounted for by Group L1's early session having an unusually high number of member-experience interventions and Group H1 having an unusually low number of member-experience interventions overall. Second, Group H1 one had relatively more interventions focused on readings ($z = 4.38$, $p < .01$), for which there is no clear explanation.

Initiation of Exploration

For initiation of exploration, the distribution of interventions based on level of treatment response was significantly different than expected by chance ($\chi^2 = 10.4$, 2df, $p < 0.001$; with Yates correction $\chi^2 = 9.93$, 2df, $p < 0.01$; see Table 18). Adjusted standardized residuals show that high treatment response groups had proportionately more spontaneously-generated interventions than expected ($z = 3.2$, $p < .01$), and relatively fewer member-initiated interventions ($z = -3.2$, $p < .01$) compared to low treatment response groups. The two high treatment response groups were not significantly different than expected, and the two low treatment response groups were not significantly different than expected.

Table 18
Initiation of Exploration across Level of Treatment Response

Treatment Response	Spontaneously-generated	Member-initiated	Total
High Groups	194 64.7%	106 35.3%	300 46.9%
Low Groups	177 52.1%	163 47.9%	340 53.1%
Total	371 58.0%	269 42.0%	640

Valence of Feedback

For valence of feedback, the distribution of interventions based on level of treatment response was significantly different than expected by chance ($\chi^2 = 21.16$, 2df, $p < 0.001$; see Table 19). Adjusted standardized residuals show that, as predicted, high treatment response groups had relatively fewer negative feedback interventions ($z = -3.7$, $p < .01$) compared to low treatment response groups. High treatment response groups also had proportionately more positive feedback intervention than expected by chance ($z = 3.5$, $p < .01$). The two high treatment response groups were not significantly different than expected, and the two low treatment response groups were not significantly different than expected.

Table 19
Valence of Feedback across Level of Treatment Response

Treatment Response	Positive	Negative	Neutral	Total
High Groups	207 49.6%	11 2.6%	199 47.7%	417 50.1%
Low Groups	156 37.6%	35 8.4%	224 54.0%	415 49.9%
Total	363 43.6%	46 5.5%	423 50.8%	832

Type of Feedback

For type of feedback, the distribution of interventions based on level of treatment response was significantly different than expected by chance ($\chi^2 = 32.42$, 3df, $p < 0.001$; see Table 20). Adjusted standardized residuals show that, as predicted, high treatment response groups had proportionately more elaborative feedback interventions than expected ($z = 5.1$, $p < .01$). High treatment response groups also had fewer feedback interventions providing simple acknowledgement or disagreement than expected ($z = -$

3.2, $p < .01$) and fewer paraphrases than expected ($z = -2.6$, $p < .01$) compared to low treatment response groups. High treatment response groups did not have proportionately fewer reframing interventions than low cognitive groups as predicted.

Table 20
Type of Feedback across Level of Treatment Response

Treatment Response	Simple Acknowledgement or Disagreement	Paraphrase	Reframe	Elaborate	Total
High Groups	118 28.3%	81 19.4%	83 19.9%	135 32.4%	417 50.1%
Low Groups	161 38.8%	112 27.0%	71 17.1%	71 17.1%	415 49.9%
Total	279 33.5%	193 23.2%	154 18.5%	206 24.8%	832

It is important to note that, while the distribution comparing the low treatment response groups was not significantly different, the distribution comparing type of feedback for the two high treatment response groups was significantly different than expected than chance ($\chi^2 = 15.5$, 3df, $p < 0.001$; see Table 21). Compared to Group H1, Group H2 received proportionately more simple acknowledgement or disagreement interventions than expected ($z = 3.8$, $p < .01$), and fewer paraphrase interventions than expected ($z = -2.1$, $p < .05$).

Table 21
Type of Feedback across High Treatment Response Groups

Group	Acknowledge	Paraphrase	Reframe	Elaborate
H1	20.1% (43)	23.4% (50)	21.5% (46)	35.0% (75)
H2	36.9% (75)	15.3% (31)	18.2% (40)	29.6% (60)

Direction of Interventions

For direction of interventions, the distribution of interventions based on level of treatment response was significantly different than expected by chance ($\chi^2 = 63.91$, 2df, p

< 0.001; see Table 22). As predicted, adjusted standardized residuals show that high treatment response groups had proportionately more group-directed interventions than expected ($z = 5.0, p < .01$), more both-directed interventions than expected ($z = 4.0, p < .01$), and fewer individually-directed interventions than expected ($z = -8.0, p < .01$) compared with low treatment response groups. The two high treatment response groups were not significantly different than expected, and the two low treatment response groups were not significantly different than expected.

Table 22
Direction of Interventions across Level of Treatment Response

Treatment Response	Toward Group	Toward Individual	Both (combination)	Total
High Groups	520 38.0%	515 37.7%	332 24.3%	1367 47.2%
Low Groups	448 29.2%	804 52.5%	280 18.3%	1532 58.8%
Total	968 33.4%	1319 45.5%	612 21.1%	2899

Stage of Group Analyses

Main Categories of the GLIS

For the main categories of the GLIS, the distribution of interventions based on stage of group was significantly different than expected by chance ($\chi^2 = 57.35, 10df, p < 0.001$; see Table 23). As predicted, adjusted standardized residuals showed proportionately more interventions focused on structure in early sessions than expected by chance ($z = 3.0, p < .01$), and proportionality fewer interventions focused on structure in later sessions than expected ($z = -4.8, p < .01$).

Adjusted analyses showed also proportionately more interventions focused on group cohesion in middle sessions than expected by chance ($z = 2.1, p < .05$). In

addition, there were proportionately fewer interventions focused on providing information in earlier sessions than expected ($z = -4.9, p < .01$), and proportionately more interventions providing information in later sessions than expected ($z = 4.2, p < .01$). Finally, there were proportionately fewer interventions focused on exploration in middle sessions than expected ($z = -3.2, p < .01$).

Table 23
Main Categories of the GLIS across Group Stage

Stage	Structure	Group Cohesion	Model	Inform	Explore	Feedback	Total
Early	321 30.9%	30 2.9%	30 2.9%	119 11.5%	249 24.0%	290 27.9%	1039 35.8%
Middle	293 29.5%	41 4.1%	18 1.8%	167 16.8%	185 18.6%	289 29.1%	993 34.3%
Late	186 21.5%	21 2.4%	25 2.9%	176 20.3%	206 23.8%	253 29.2%	867 29.9%
Total	800 27.6%	92 3.2%	73 2.5%	462 15.9%	640 22.1%	832 28.7%	2899

Type of Structure

For type of structure, the distribution of interventions based on stage of group was significantly different than expected by chance ($\chi^2 = 11.64, 4df, p < 0.05$; see Table 24). Adjusted standardized residuals showed that later sessions contained proportionately more routine management related interventions than expected by chance ($z = 2.5, p < .01$), and fewer behavioral management related interventions than expected ($z = -3.3, p < .01$). When interpreting these results it is important to keep in mind that, at the main scale level, later sessions had fewer interventions focused in structure than would be expected by chance.

Table 24
Type of Structure across Group Stage

Stage	Long-term	Routine	Behavioral	Total
Early	22 6.9%	228 71.0%	71 22.1%	321 40.1%
Middle	16 5.5%	211 72.0%	66 22.5%	293 36.6%
Late	15 8.1%	150 80.6%	21 11.3%	186 23.3%
Total	53 6.6%	589 73.6%	158 19.8%	800

Group Cohesion

For type of group cohesion, the distribution of interventions based on stage of group was not significantly different than expected by chance ($\chi^2 = 5.14$, 2df, $p < .10$; see Table 25).

Table 25
Type of Group Cohesion across Group Stage

Stage	Team Building	Emotional Engagement	Total
Early	15 50%	15 50%	30 32.6%
Middle	17 41.5%	24 58.5%	41 44.6%
Late	4 19.0%	17 81.0%	21 22.8%
Total	36 39.1%	56 60.9%	92

Modeling

For type of modeling, the distribution of interventions based on stage of group was not significantly different than expected by chance ($\chi^2 = .10$, 2df, $p < 1$; see Table 26).

Table 26
Type of Modeling across Group Stage

Stage	Self-Disclosure	Interaction with Others	Total
Early	14 46.7%	16 53.3%	30 41.1%
Middle	9 50.0%	9 50.0%	18 24.7%
Late	9 36.0%	16 64.0%	25 34.2%
Total	32 43.8%	41 56.2%	73

Initiation of Information

For initiation of information, the distribution of interventions based on stage of group was not significantly different than expected by chance ($\chi^2 = 2.15$, 2df, $p < 1$; see Table 27).

Table 27
Initiation of Information across Group Stage

Stage	Spontaneous	Member-initiated	Total
Early	78 65.5%	41 34.5%	119 25.8%
Middle	98 58.7%	69 41.3%	167 36.1%
Late	101 57.4%	75 42.6%	176 38.1%
Total	277 60.0%	185 40.0%	462

Topic of Exploration

For topic of exploration, the distribution of interventions based on stage of group was significantly different than expected by chance ($\chi^2 = 107.00$, 4df, $p < 0.001$; see Table 28). As predicted, adjusted standardized residuals show that group leader interventions in late sessions were less likely than expected by chance to focus exploration on readings ($z = -5.7$, $p < .01$). Also as predicted, late sessions were more

likely to focus interventions on group-events ($z = 8.8, p < .01$), compared to middle sessions ($z = -2.8, p < .01$) and early sessions ($z = -5.8, p < .01$), which both had fewer exploration interventions focused on group-events than expected.

There were two additional findings. First, there were more exploration interventions focused on member-experiences than expected by chance in early sessions ($z = 5.1, p < .01$). Second, exploration interventions in middle sessions were more likely than expected by chance to focus on readings ($z = 5.9, p < .01$), and less likely to focus on member-experiences than expected by chance ($z = -3.9, p < .01$).

Table 28
Topic of Exploration across Group Stage

Stage	Readings	Group-event	Member Experience	Total
Early	124 49.8%	21 8.4%	104 41.8%	249 38.9
Middle	126 68.1%	24 13.0%	35 18.9%	185 28.9%
Late	69 33.5%	83 40.3%	54 26.2%	206 32.2%
Total	319 49.8%	128 20.0%	193 30.2%	640

Initiation of Exploration

For initiation of exploration, the distribution of interventions based on stage of group was not significantly different than expected by chance ($\chi^2 = 1.88, 2df, p < 1$; see Table 29).

Table 29
Initiation of Exploration across Group Stage

Stage	Spontaneous	Member-initiated	Total
Early	140 56.2%	109 43.8%	249 38.9%
Middle	115 62.2%	70 37.8%	185 28.9%
Late	116 56.3%	90 43.7%	206 32.2%
Total	371 58.0%	269 42.0%	640

Valence of Feedback

For valence of feedback, the distribution of interventions based on stage of group was significantly different than expected by chance ($\chi^2 = 25.00$, 4df, $p < 0.001$; see Table 30). Adjusted standardized residuals show that there were more positive feedback interventions in early sessions than would be expected ($z = 3.4$, $p < .01$), and fewer neutral feedback interventions than expected ($z = -3.8$, $p < .01$). However, in late sessions there were fewer positive feedback interventions than expected ($z = -2.3$, $p < .05$), fewer negative feedback interventions ($z = -2.8$, $p < .01$), and more neutral feedback interventions than expected ($z = 3.7$, $p < .01$).

Table 30
Valence of Feedback across Group Stage

Stage	Positive	Negative	Neutral	Total
Early	150 51.7%	19 6.6%	121 41.7%	290 34.9%
Middle	118 40.8%	22 7.6%	149 51.6%	289 34.7%
Late	95 37.5%	5 2.0%	153 60.5%	253 30.4%
Total	363 43.6%	46 5.5%	423 50.8%	832

Type of Feedback

For type of feedback, the distribution of interventions based on stage of group was significantly different than expected by chance ($\chi^2 = 18.71$, 6df, $p < 0.01$; see Table 31).

Adjusted standardized residuals show that there were more feedback interventions focused on elaboration in early sessions than expected ($z = 3.2$, $p < .01$). In late sessions, there were more interventions providing simple acknowledgement than expected ($z = 2.7$, $p < .01$) and fewer interventions containing elaboration than expected ($z = -2.6$, $p < .01$).

Table 31
Type of Feedback across Group Stage

Stage	Acknowledge	Paraphrase	Reframe	Elaborate	Total
Early	87 30.0%	67 23.1%	45 15.5%	91 31.4%	290 34.9%
Middle	90 31.1%	76 26.3%	56 19.4%	67 23.2%	289 34.7%
Late	102 40.3%	50 19.8%	53 20.9%	48 19.0%	253 30.4%
Total	279 33.5%	193 23.2%	154 18.5%	206 24.8%	832

Direction of Interventions

For direction of interventions, the distribution of interventions based on stage of group was not significantly different than expected ($\chi^2 = 9.34$, 4df, $p < .10$; see Table 32).

Table 31
Direction of Interventions across Group Stage

Stage	Group	Individual	Both	Total
Early	367 35.3%	458 44.1%	214 20.6%	1039 35.8%
Middle	313 31.5%	445 44.8%	235 23.7%	993 34.3%
Late	288 33.2%	416 48.0%	163 18.8%	867 29.9%
Total	968 33.4%	1319 45.5%	612 21.1%	2899

Number of Interrupted Interventions

It was predicted that interventions were more likely to be interrupted in low treatment response groups than in high treatment response groups. This prediction was partially supported. The overall frequency of interrupted interventions was not significant according to treatment response ($\chi^2 = 2.67$, 1df, $p = .10$). However, the distribution of interrupted interventions comparing high and low treatment response groups across stage of group was significantly different than expected by chance ($\chi^2 = 20.86$, 2df, $p < 0.001$; see Table 33). Adjusted standardized residuals show that low treatment response groups were more likely to be interrupted in middle sessions compared to high treatment response groups ($z = 4.15$, $p < .01$). For the late sessions, the opposite pattern than predicted emerged, with high treatment response groups having more interrupted interventions than low treatment response groups than expected by chance ($z = 3.9$, $p < .01$). Recall that overall interventions were not statistically different when comparing high and low treatment response groups across stage of group (see Table 9), strengthening the interpretive value of this finding. In addition, when directly compared, the two high treatment response groups and the two low treatment response groups were not significantly different.

Table 33
Number of Interrupted Interventions per Session

Treatment Response	Early Sessions	Middle Sessions	Late Sessions	Total
High Groups	24 41.7%	7 15.2%	39 58.2%	70 40.9%
Low Groups	34 58.6%	39 84.8%	28 41.8%	101 59.1%
Total	58 33.9%	46 26.9%	67 39.2%	171

Table 34
Summary of Predicted Findings

Category	Finding	Differences present within high or low groups	Additional interpretive issues
1. Type of structure according to treatment response	As predicted, high treatment response groups received proportionately fewer behavioral management interventions compared to low treatment response groups.	Overall, Group H1 received more structure-related interventions than Group H2. Within the structure category, Group H1 had proportionately fewer routine management interventions and more behavior management interventions compared to Group H2.	High treatment response groups had proportionately more routine management interventions compared to low treatment response groups.
2. Overall structure according to stage of group	As predicted, across the main categories of the GLIS there were proportionately more interventions focused on structure in early sessions than expected by chance, and fewer interventions focused on structure in later sessions.	N/A	Although not predicted, the type of structure that occurred differed according to stage of group. In later sessions, there were proportionately more (although fewer in number) routine management interventions and fewer behavioral management interventions than expected by chance.
3. Topic of exploration according to stage of group	As predicted, group leader interventions in late sessions were less likely to explore readings compared to group-events or member experiences, and more likely to focus on group events.	N/A	Two related findings were: (a) in early sessions, there were more explorations of member experience and fewer explorations of group events; and (b) in middle sessions, there was more exploration of readings compared to group events and member experiences.

Category	Finding	Differences present within high or low groups	Additional interpretive issues
4. Valence of feedback according to treatment response	As predicted, high treatment response groups received proportionately fewer negative feedback interventions compared to low treatment response groups.	No	A related finding was that high treatment response groups received proportionately more positive feedback than low treatment response groups.
5. Type of feedback according to treatment response	As predicted, high treatment response groups received more elaborative feedback compared to low treatment response groups. However, high treatment response groups did not receive fewer reframing interventions than low treatment response groups as predicted.	Group H2 received proportionately more interventions focused on simple acknowledgement or disagreement, and proportionately fewer interventions focused on paraphrasing compared to Group H1.	A related finding was that high treatment response groups had fewer feedback interventions providing simple acknowledgement compared to low treatment response groups.
6. Direction of interventions according to treatment response	As predicted, high treatment groups received more interventions directed toward the whole group and at both (i.e., directed toward an individual, but applies to the whole group), and fewer interventions directed toward individuals compared to low treatment response groups.	No	

Category	Finding	Differences present within high or low groups	Additional interpretive issues
7. Interrupted interventions according to treatment response	Overall, low treatment response groups were more not more likely to have interrupted interventions within sessions compared to high treatment response groups as predicted. However, there were group stage differences for this variable.	No	Group stage differences: (a) low treatment response groups had more interrupted interventions in middle sessions; (b) high treatment response groups had more interrupted interventions in late sessions.

Table 35
Summary of Additional Findings

Category	Finding	Differences present within high or low groups
1. Main categories of the GLIS according to stage of group	Three other significant findings were present among the main categories according to stage of group besides the different amount of structure addressed in the predicted findings: (a) there were more group cohesion interventions in middle sessions than expected by chance; (b) there were fewer exploration interventions in middle sessions than expected by chance; (c) there were fewer information interventions in middle sessions, and more information interventions in later sessions than expected by chance.	N/A
2. Topic of exploration according to treatment response	Low treatment response groups had proportionately more member-experience focused interventions than expected by chance, and fewer reading-focused interventions compared to high treatment response groups. These findings are questionable given the differences found within high and low treatment response groups.	Group H1 and Group H2 were significantly different from one another. Group L1 and Group L2 also were significantly different from one another. These findings are largely linked to session differences resulting from how the sessions were distributed across high and low treatment response groups (i.e., the same sessions were coded for Groups H1 and L1, and the same sessions were coded for Groups H2 and L2). When directly compared, there are no significant differences for Groups H2 and L2. However, when Groups H1 and L1 are directly compared, there were still two group differences that cannot be explained on the basis of session effects and are likely due to group differences. Group L1 had relatively fewer interventions focused on readings, and relatively more interventions focused on member experiences than Group H1.

Category	Finding	Differences present within high or low groups
3. Initiation of exploration according to treatment response	High treatment response groups had proportionately more spontaneously-generated interventions than expected by chance, and relatively fewer member-initiated interventions compared to low treatment response groups.	No
4. Valence of feedback according to stage of group	In early sessions, there were more positive feedback interventions and fewer neutral feedback interventions than expected by chance. In late sessions, there were fewer positive and negative feedback interventions and more neutral feedback interventions than expected by chance.	N/A
5. Type of feedback according to stage of group	In early sessions, there were more elaborative feedback interventions than expected by chance. In late sessions, there were more interventions providing simple acknowledgment and fewer interventions containing elaborative feedback than expected by chance.	N/A

DISCUSSION

The purpose of this study was to develop and validate an instrument for examining therapist interventions in group counseling settings. The new instrument, named the Group Leader Intervention Scales (GLIS), was developed through examination of written transcripts of sessions from a group intervention for children that used stories and peer group processes to increase social problem-solving. The GLIS codes group leader verbalizations on six main group process variables (structure, group cohesion, modeling, information, exploration and feedback), as well as various subcategories linked to the main categories.

Results of the current study provide support regarding the interrater reliability of the GLIS and provide initial support regarding the validity of the GLIS in terms of its ability to distinguish between groups on the basis of cognitive treatment response (high or low) and on the basis of group stage (early, middle, or late).

Acceptable interrater reliability levels ($\kappa \leq .70$) were achieved for the main categories of the GLIS. Acceptable interrater reliability levels also were achieved for all of the subscales connected to the main categories except two, type of information, and level of exploration, which have been removed from the current version of the scale (see Appendix A). Two additional subcategories within the GLIS, initiation of exploration and valence of feedback, did not achieve acceptable interrater reliability levels when the results of the interrater reliability coding were averaged across sessions, but were kept in the final version of the GLIS since acceptable interrater reliability levels were reached by the third session that was collectively coded by the three raters and on the reliability check performed by two raters as the final nine sessions were coded.

An additional scale, the Framework Building Scale, was developed to assess social problem-solving as conceptualized by Crick and Dodge (1994), taking into account the role emotion plays in the process (Lemerise & Arsenio, 2000). Unfortunately, acceptable interrater reliability levels were not achieved for the main categories of the Framework Building Scale, or for its two subcategories. Therefore, these scales have been removed from the final version of the GLIS. It may be that social problem-solving is too complex a construct to be captured with a nominal scale. Crick and Dodge (1994) argued that the behavioral path in any given situation follows a logical sequence of events (e.g., from a single provocation by a peer to retaliation). However, they also conceded that human information processing does not occur sequentially, but rather it occurs in a simultaneous manner. While the problem-solving steps contained in the Crick and Dodge model were examined in the STORIES program, the steps were not always examined individually. In practice, many group leader interventions dealt with more than one step, making reliable coding using the Framework Building System difficult since it utilized mutually exclusive categories.

Among the main categories of the GLIS, there were no significant differences according to treatment response, although it is important to note that there were differences in the number of interventions that were used across the main categories. Overall, interventions most frequently utilized feedback, structure, exploration and information, while group cohesion and modeling were relatively rare. While the main categories did not distinguish groups based on treatment response, the subcategories within the main categories were able to tease apart differences based on treatment response. The inclusion of subcategories at this level of specificity is unique to the GLIS

and is an important contribution to the study of group processes since current therapist verbal response mode systems that are commonly used only measure interventions at the main category level.

Predicted Findings

Prediction 1: Type of Structure according to Treatment Response

As predicted, high treatment response groups received proportionately fewer behavioral management interventions compared to low treatment response groups. In addition, high treatment response groups had proportionately more routine management interventions compared to low treatment response groups. It is important to note that, within the high treatment response groups, Group H1 received more structure-related interventions overall compared to Group H2. Within the structure subcategories, Group H1 had proportionately fewer routine management interventions and more behavior management interventions compared to Group H2. These results are not surprising since Group H2 was an unusually high functioning group. Of the six members in Group H2, five members received the highest possible score (3) on the treatment response measure (average treatment response = 2.6). For Group H1, only three out of the six members received the highest possible score on the treatment response measure (average treatment response = 2.3). It is plausible that the highest functioning group did not require as much structuring as the other high functioning group. Overall though, the difference between the high groups is minor when the high groups are compared to both of the lower functioning groups (i.e., high treatment response groups still had far fewer behavior management interventions compared to low treatment response groups).

Prediction 2: Overall Structure according to Stage of Group

As predicted, across the main categories of the GLIS there were proportionately more interventions focused on structure in early sessions, and fewer interventions focused on structure in later sessions. The finding that early sessions received more structure is supported by the literature (Dies, 1994; Yalom, 1995). In addition, after reviewing the literature, McGuire, Taylor, Broome, Blau and Abbott (1986) concluded that structure in the early phases of treatment has positive effects on member perception of the group, process involvement and level of communication within the sessions.

Although not predicted, the type of structure that occurred also differed according to stage of group. In later sessions, there were proportionately more (although fewer in overall number) routine management interventions and fewer behavioral management interventions than expected by chance. These findings may be related to the greater number of activities that were conducted in the later sessions of the STORIES program. These findings may also be linked to members having been conditioned thorough earlier sessions to follow the rules, and therefore, they no longer require as many behavioral management interventions from the group leader.

Prediction 3: Topic of Exploration according to Stage of Group

As predicted, group leader interventions in late sessions were less likely to explore readings compared to group events or member experiences, and more likely than expected by chance to focus interventions on group events. These findings are not surprising since activities are more prevalent in later sessions, so it is logical that there are fewer reading-related interventions and more group-event related interventions.

Two additional related findings were: (a) in early sessions, there were more explorations of member experience and fewer explorations of group events; and (b) in middle sessions, there was more exploration of readings compared to group events and member experience. It may be that members' previous experiences are initially used to engage the children in the group, but that the focus shifts in the middle sessions toward a more thorough exploration of the concepts being introduced in the readings.

Prediction 4: Valence of Feedback according to Treatment Response

As predicted, high treatment response groups received proportionately fewer negative feedback interventions compared to low treatment response groups. A related finding was that high treatment response groups received proportionately more positive feedback than low treatment response groups. These differences are not surprising since lower quality responses are more likely to lead to disagreement from the group leader. It is also important to stress that negative feedback interventions within the STORIES program are rare overall compared to positive and neutral feedback interventions, and that they were often quite mild (e.g., "Not exactly," or "No, the character's name was...").

Prediction 5: Type of Feedback according to Treatment Response

As predicted, high treatment response groups received more elaborative feedback compared to low treatment response groups. This finding is intuitive since members of high treatment response groups are more capable, cognitively and behaviorally, of receiving detailed elaboration of topics and previous comments. Also, elaboration implies acceptance of what was said, and provides a starting point to extend the concept further. A related finding was that high treatment response groups had proportionately fewer feedback interventions providing simple acknowledgement compared to low treatment

response groups. However, high treatment response groups did not receive fewer reframing interventions compared to low treatment response groups as predicted. In fact, high groups had more reframing interventions compared to low groups (83 compared to 71), although this difference was not significant. It is possible that the group leader reframed responses for low and high treatment response groups for different purposes. For low treatment response groups, the intent may have been more corrective, whereas the intent in high treatment response groups may have been to help group members view concepts in a more complex manner.

It is important to note that Group H2 received proportionately more interventions focused on simple acknowledgement or disagreement, and proportionately fewer interventions focused on paraphrasing compared to Group H1. As was the case with differences within the high treatment response groups in the structure subcategory, these differences can be viewed in terms of Group H2 being an unusually high functioning group. One could argue that Group H2 only needed group leader support at the level of simple acknowledgment or disagreement compared to Group H1, where paraphrasing would help reinforce the concepts that were being discussed. The differences between the high groups for acknowledgement and paraphrasing, however, should not detract from the main predicted finding that high treatment response groups would have more elaboration focused interventions since both Group H1 ($n = 75$) and Group H2 ($n = 60$) had a much higher frequency of elaboration interventions than either of the low treatment response groups (Group L1 = 38; Group L2 = 33).

Prediction 6: Direction of Interventions according to Treatment Response

As predicted, high treatment groups received more interventions directed toward the whole group and at both (i.e., directed toward an individual, but applied to the whole group), and fewer interventions directed toward individuals compared to low treatment response groups. One interpretation for this finding is that individuals within the low treatment response groups have an increased need for individual structuring (particularly behavioral management) and an increased for the group leader to periodically “check-in” with them for their understanding. A potential confounding variable for this interpretation is that high and low groups differed according to size. Low groups each had four members, whereas high treatment response groups each had six members. Having fewer members in the low treatment response groups may have made directing more interventions toward individuals possible.

Prediction 7: Interrupted Interventions according to Treatment Response

Overall, low treatment response groups were not more likely to have interrupted interventions within sessions compared to high treatment response groups as predicted. However, there were group stage differences for this variable: (a) low treatment response groups had more interrupted interventions in middle sessions compared to high treatment response groups; and (b) high treatment response groups had more interrupted interventions in late sessions compared to low treatment response groups. A qualitative analysis of the types of interruptions that occurred for high and low functioning groups provides a reasonable explanation for these findings. High treatment response group interruptions tended to be linked to the content being covered, and were more conversational in nature. Conversely, low treatment response group interruptions tended

to be more disruptive to the flow of the session and were often accompanied by a structural intervention by the group leader (e.g., a behavioral management intervention such as, “Shh;” or a routine management intervention such as “We are on page 73”). The fact that more interruptions occurred in the late sessions for high treatment response groups compared to low treatment response groups may partly be a function of the increased number of less structured activities that occur in late sessions, and partly due to group members feeling more comfortable spontaneously sharing comments that were on-topic, and therefore, did not require structural intervention by the group leader.

Additional Findings

Main Categories of the GLIS according to Stage of Group

Although there were no main category differences according to treatment response, there were four significant main category differences according to stage of group. First, overall structure differed according to stage of group as previously addressed in the predicted findings section. Second, there were more group cohesion interventions in middle sessions than expected by chance. This finding may be linked with overcoming resistance, which is common in the middle stage of group development (Corey & Corey, 2002; Yalom, 1995). Third, there were fewer exploration interventions in middle sessions than expected by chance. Fourth, there were fewer information interventions in middle sessions and more information interventions in later sessions than expected by chance. It is important to recognize that these stage effects must be viewed as exploratory and interpreted with caution since the sessions within the STORIES program differed in the relative emphasis of readings or activities and not all sessions from the program were included in the present study.

Topic of Exploration according to Treatment Response

Low treatment response groups had proportionately more member-experience focused interventions than expected by chance, and fewer reading-focused interventions compared to high treatment response groups. The interpretation of this finding in terms of the low responding groups needing more concrete discussion of concepts (i.e., personal experience) introduced in the readings is tempting. However, such an interpretation is tenuous given the differences found within high and low treatment response groups.

Group H1 and Group H2 were significantly different from one another based on topic of exploration. Group L1 and Group L2 also were significantly different from one another based on topic of exploration. These findings were linked largely to session differences resulting from how the sessions were counterbalanced across high and low treatment response groups (i.e., the same sessions were coded for Groups H1 and L1, and the same sessions were coded for Groups H2 and L2). For instance, Session 13 was coded as the late session for both Group H2 and Group L2. The focus of session 13 was an art activity, which necessitated more group event exploration. The art activity in session 13 also helps to explain why Group H2 had relatively fewer interventions focused on the readings, although this finding was also impacted by Group H1's relatively high frequency of interventions focused on readings. When directly compared, there are no significant differences between Groups H2 and L2 across the topic of exploration category. However, when Groups H1 and L1 are directly compared, there were still two group differences that cannot be explained on the basis of session effects and are likely due to group differences: Group L1 had relatively fewer interventions focused on readings, and relatively more interventions focused on member experiences. This pattern

of findings is commensurate with the previous interpretation that low responding groups required more concrete discussion of concepts introduced through the readings.

The unclear findings for topic of exploration according to treatment response illustrates the difficulty of examining only a portion of the sessions within a course of treatment since the main difference within this variable was a session difference resulting from how sessions were counterbalanced across groups based on level of treatment response. However, despite not having “clean” results for this subcategory, the GLIS was still able to distinguish between groups; the distinction was just based on a session effect, rather than on level of treatment response or stage of group.

Initiation of Exploration according to Treatment Response

High treatment response groups had proportionately more spontaneously-generated interventions than expected by chance, and relatively fewer member-initiated interventions compared to low treatment response groups. It may be that the group leader felt freer to move the discussions into new areas with the high treatment response groups since the group was capable of covering more conceptual ground, where the group leader let group members in the low treatment response groups initiate further exploration as they were ready to. Also, since there were more interventions focused on member experience in low treatment response groups there would naturally be more follow up of member initiated information.

Valence of Feedback according to Stage of Group

In early sessions, there were more positive feedback interventions and fewer neutral feedback interventions than expected by chance. In late sessions, there were fewer positive and negative feedback interventions and more neutral feedback interventions

than expected by chance. This finding is somewhat difficult to contextualize since the current literature on valence of feedback does not contain studies on stage of treatment. However, it seems as though a higher degree of positive feedback would be important in earlier sessions to encourage group members' participation (i.e., build a level of comfort for expressing their ideas), but that this is less important in later sessions when the group has been working together for awhile. The decreased proportion of negative feedback in later sessions may be related to increased conceptual growth among group members, which is demonstrated through higher quality comments that require less negative feedback. The decrease in the proportion of negative feedback in later sessions may also be linked to less exploration of readings in later sessions, and therefore, fewer "right or wrong" group member comments to respond to.

Type of Feedback according to Stage of Group

In early sessions, there were more elaborative feedback interventions than expected by chance. In late sessions, there were more interventions providing simple acknowledgment and fewer interventions containing elaborative feedback than expected by chance. It may be that elaborative feedback early on helps teach group members a common language that is linked to the concepts being addressed in the STORIES program and/or provides an opportunity to introduce related concepts that have not yet been addressed (e.g., early on "bully" and "victim" must be defined and linked to story content). Once all group members are speaking a "common language," and other group members begin providing relevant feedback to members, it would be less necessary for group leaders to provide elaborative feedback. Although the current study did not examine the quality of group member feedback, results from a study conducted by

Morran, Stockton and Harris (1991) suggested that group member feedback becomes more effective, and more similar to group leader feedback, throughout the life of a group.

The Importance of all of the Main GLIS Categories in Combination

For both level of treatment response and stage of group, there were no significant findings for three of the GLIS subcategories: (a) type of group cohesion, (b) type modeling, and (c) initiation of information. One could argue that group cohesion, as defined by team building and emotional engagement interventions, is an affective scale in essence and that the appropriate level of group cohesion depends on the combination of individuals within any given group and how they mesh together, rather than on the level of cognitive growth groups display. Given this argument, it is not surprising that no pattern for group cohesion was found among the four groups in the current study.

One common factor among the three subcategories that failed to obtain significant findings was the relative infrequency with which those subscales were used relative to the other subscales. This was particularly true for group cohesion and modeling. In other settings, and if examining different research questions, these scales may in fact be more central. Furthermore, if the group cohesion and modeling categories had not been present in the current study, these interventions would have been coded as miscellaneous. So, while these two scales were not chosen as often as the other four main scales, and did not distinguish between level of treatment response or group stage within this study, they serve an important function within the instrument.

While there is a “miscellaneous” category available on the GLIS, it was used only four times throughout the current study; and these instances occurred as sessions were being coded across raters for interrater reliability. Even then, the other two raters did not

agree that they were miscellaneous interventions and any differences that were found among the other two raters regarding specific interventions were settled during the reconciliation process. Therefore, while the original intent was to study miscellaneous codes for patterns in order to improve the instrument, there were too few to find any meaningful patterns and no miscellaneous codes remained by the final analyses.

Limitations of the Study

- 1) The GLIS has face validity and interrater reliability, and may be compared with existing measures for content validity. However, construct validity was not firmly established for the new measure in the current study.
- 2) Having the same group leader for all groups was both an advantage and a disadvantage. It made the differences between groups easier to examine. However, leadership style has many potential interacting variables so findings from the current study that may be perceived as effective leadership may not necessarily generalize to other group leaders.
- 3) Stage of treatment effects, particularly those that were discovered during the exploratory analyses, should be interpreted with caution since not all sessions over the course of treatment were coded in the current study. This was problematic in that it confounded some of the results. For example, the comparison of treatment response on the topic of explorations was confounded by session differences (e.g., i.e., activities or readings).
- 4) Generalizability concerns for the GLIS exist given the small number of groups that were explored and the specific nature of the program from which it was developed. However, the STORIES program, like most interventions for children, follows a

structured curriculum. Hence, results of this study suggest that the GLIS is applicable for examining different groups that also utilize structured curriculums.

Future Directions

The current study was unable to establish an acceptable level of interrater reliability for the abstract/concrete distinction within the exploration category or the type of information that was provided (fact, reason, and review). In large part, the failure to establish an acceptable level for these two subscales appears to be linked to poor operational definitions. For instance, the abstract/concrete distinction within exploration never properly defined the complexity of abstract exploration. Exploration questions which required abstract thought could still be answered yes or no, which by the current definition were coded as a concrete question. It was especially disappointing that the abstract/concrete exploration subcategory was excluded prior to the validity analyses since in theory it would have been useful for categorizing differences between high and low cognitive treatment response groups. A future study could reexamine these two scales to see if clearer operational definitions would increase their reliability.

Further research is needed to see how the GLIS generalizes to less structured interventions. For example, the GLIS could be to examine different types of counseling groups and therapists with different theoretical orientations. Results obtained through this approach would likely be even more dramatic than results in the current study, which examined only one type of program and only one group leader. For instance, Rogerian therapists are likely to utilize more group cohesion, particularly emotional engagement, than cognitive-behavioral therapists, who would likely utilize more reframing feedback.

This study only examined group leader verbalizations. However, future studies could use the GLIS along with a measure of group member verbal responses (e.g., Hill, 1992) to assess the effectiveness of various therapist interventions. This would be an important step in beginning to understand how interactions between group leader and group members impact group processes.

APPENDIX A:

Group Leader Intervention Scales Coding Manual

Rules for using the scales:

- Codes from the GLIS are assigned to thought units. Thought units are defined as verbalizations by the interventionist that together makes a cohesive idea or intention.
- Each thought unit is coded for one main element (category A-G), for subcategories contained within each element, and for whether the intervention was directed toward an individual, the group or both.
- Thought units are indicated in the text by a backslash and are numbered consecutively throughout a session.
- Idle chatter at the beginning or end of a session that is not related to the session and/or group is not broken into thought units or subsequently coded. Likewise, story content covered in readings is not broken into thought units or subsequently coded.
- A thought unit may be one sentence or it may be a group of related sentences. One speaking turn may contain several thought units or it may take several speaking turns to comprise a thought unit.
- If the group leader is interrupted before a thought was finished and he or she picks up on that same thought at the beginning of the next speaking turn it would be considered one thought unit. It is also possible for a separate thought unit to occur in between a continued thought unit within the same speaking turn. For example, the group leader may have to bring group members' attention back to him or her prior to finishing the original thought. Continuations of previous thought units would receive the same number with "(con't)" after it to indicate it is a continued thought unit.
- Two consecutive pieces of information provided by the group leader that are not related and are separated by either a reading or a group member comment are considered separate thought units rather than a continuation of the same "information" thought unit.
- If the group leader poses a general question or provides information to the group then calls on a specific member it would be divided into two thought units. For example, "What's Bobby feeling guilty about? How do you feel inside, when you feel guilty? Yes, Tammy?" would be split into two thought units: "What's Bobby feeling guilty about? How do you feel inside, when you feel guilty?" (exploration) and "Yes, Tammy?" (structure).

- One sentence may contain multiple intents, and as such, would be broken into more than one thought unit. For instance, a sentence such as, “That is a good idea, what do other group members think about this?” would be split into two thought units: “That is a good idea” (feedback) and “what do other group members think about this?” (exploration). This may also occur if the group leader begins the speaking turn by repeating or agreeing with a group member previous comment then moves on to a new thought unit within the same sentence. For example, “Yeah (feedback), and she’s embarrassed by that (information).”
- When a question is raised and immediately answered by the group leader that is clearly intended to set the stage for an upcoming session or group event, such as “Now what do you think is going to happen, we’re going to find out next week,” it will not be separated into separate thought units. This is in contrast to, “Who do you think might be Isaac from the front cover? Probably him right?” where it is harder to tell from the written transcript if the information was subsequently given because the group members did not respond to the question.
- Determining at what point reframing or elaborative feedback becomes exploration or information. Rule of thumb: when feedback veers substantially from the content included in group members’ previous comments (within several speaking turns) it would be considered information or exploration and, therefore, would require separate thought units.

Scenario 1: The group is discussing the meaning of “war” and one of the members says “a bunch of different people get together and fight.” The group leader response being coded is, “They fight over something. Usually it is different countries, isn’t it?” This speaking turn would remain one thought unit (an elaborative feedback) because the information provided by the group leader is actually an extension of the content already provided by the group member and, therefore, is linked to the feedback.

Scenario 2: Following a reading in a book, the group leader asks group members what information they just learned about a character. A group member replied, “His father left him with his uncle.” The group leader response being coded is, “Yeah, so his parents left him and she said that after that happened, he always...?” This speaking turn would be separated into two thought units, “Yeah, so his parents left him” (paraphrasing feedback) and “and she said that after that happened, he always...?” (exploration) because the group leader is pulling for additional content that has not been brought up by the group member.

Scenario 3: The group member tells a story about breaking a glass. The group leader responds by asking her if she felt guilty about it. The group member replied, “Yes, but she [her mother] didn’t do anything about it.” The group leader response being coded is, “Ok, she understood that it was an accident. Because sometimes if you do something wrong and you don’t tell anyone you feel guilty inside meaning you feel bad.” Since the second portion of the thought unit veers from the content

contained in the group member's comment, it would be divided into 2 thought units, "Ok, she understood that it was an accident (elaborative feedback) and, "Because sometimes if you do something wrong and you don't tell anyone you feel guilty inside meaning you feel bad" (information).

- Because the co-leader's verbal role is minimal in the STORIES program, co-leader interventions are not coded.
- Interventions that cannot be classified within any categories (i.e., miscellaneous ones) will be tallied and examined for implications in revising the measure.

A. Structure

Definition: used to manage the flow of sessions.

Types of structure:

1. Long-term - provide an advanced organizer for what to expect later in the session or for future sessions. This may include the presentation of possible ideas/topics to be explored during the next session. Examples:
 - "Before we start, let me tell you about the tape recorder"
 - "Next week we will start a new story"
 - "Before we read today, let's review from last time."
 - "Today we will meet a new character in our book who is..."
 - "We can start filling out our character web now."
 - "We'll see [following a prediction]. Chapter 3 everybody."
 - "Think about these questions and we'll talk about them next time"
2. Routine management - redirect comment or topic, facilitating turn taking by responding to verbal or nonverbal initiatives.

Examples:

 - "Yes, Donte?"
 - "Hold your thoughts {for now}."
 - "Please keep your books open to page 3."
 - "Let's skip this part and go to the bottom of page 14."
 - "Thank you for waiting" (said to call on group member)
 - "You are the leader this time, who do you chose as sweep?"

Note: If the group leader is repeating a previously asked question as a way of calling on another member, it would more appropriately fall under this category rather than exploration because the main intent is facilitating turn taking.
3. Behavioral management - enforce a rule, redirect behavior. These interventions tend to be reactive in nature. Examples:
 - "Remember we are not supposed to look ahead in the book."
 - "Stop talking."
 - "Please sit back down in your chair."

“Close your books for a minute.”

“Please put that away.”

B. Group Cohesion

Definition: efforts to engage members in the group and to foster a sense of group identity or belonging (i.e., individuals are valued by the group; the group is special). A code of group cohesion is appropriate if the group leader offers support/encouragement spontaneously.

Types of group cohesion:

1. Team building - creating an atmosphere or building traditions that foster group members' identification as a team. This includes coming up with the group name, establishing common rules, etc. Examples:

“Our group is very special.”

“I will call your group name when I come to get you.”

“To help our group, it is important to listen to each other.”

“We will always pick a leader and a sweep.”

“She has already been the leader. Please pick someone else.”

2. Emotional engagement or support - fostering investment in relationships among group members; demonstrating the importance of each individual and the value of their contribution to the group; expressing that the group is a safe place to share. Examples:

“I'm so excited we finally got to start our group.”

“I am so happy to see all of you.”

“We missed you when you were absent last week.”

“You don't have to tell, only if you want to.”

“It's okay if you don't remember what you were going to say. When you think of it we'll stop and you can tell us.”

“You may not be friends in the classroom, but we are going to get to know each other pretty well and you will learn to help each other.”

“It's okay” (following an attempt to quiet the other members so the speaker could continue with their comment).

Note: If the support is solicited through previous comments from a group member(s), it would be coded as positive feedback. For instance, if a group member first says, “I love coming to group” and the group leader responds, “And I love having you here.”

C. Modeling

Definition: the demonstration of how to perform an action or express an idea or emotion.

Types of modeling:

1. Self-disclosure - sharing a personal thought, feeling or experience. Sometimes explanations or ideas are expressed in the context of self-disclosure. Examples:
 - “I was in a dark mood today, I don’t know why. I just was.”
 - “I never liked pop quizzes myself.”
 - “Sometimes I get mad over silly things that have nothing to do with what is really bothering me...then when I say what is really bothering me I feel better.”
 - “Well, I remember starting a new school when I was your age and I was pretty nervous.”
 - “Sometimes I walk around with a big smile on my face and you may think I’m happy, but I’m not always happy even when I’m smiling.”

Note: Disclosures that follow a group member statement or question such as, “I can see that to” or, in response to a previous comment, “You know what that tells me? That tells me...” would be coded as feedback rather than self-disclosure.

2. Interaction with others - the demonstration of prosocial behavior. Examples:
 - “Can I color in your picture?”
 - “Thank you for reminding me. I had forgotten.”
 - “Let me help you...”
 - “You’re very welcome.”
 - “Thank you for your response.”

D. Information

Definition: providing known facts, clarifications, reasons or explanations for new or previously covered readings or events that occur within or outside the group (e.g., popular culture, historical references). This includes reviewing story content to ensure group member understanding prior to moving on with the readings.

Note: Group leader responses such as “yes” or “no” that occur following a group member question are coded as information rather than feedback—simple acknowledgment. (e.g., a group member asks, “Can you do that?” and the group leader responds, “Sure.”)

Initiation of information:

1. Spontaneously offered by the group leader - the focus of the intervention is on providing information about a new concept. The intervention is not directly related to the content contained in a group member’s question or statement immediately preceding the intervention.

2. Given in response to a group member’s statement or question

E. Exploration

Definition: inviting or engaging the group members to think about an idea, feeling or event in order to further clarify or extend the lesson beyond the readings or known facts. This rubric includes working with the concept through discussion, connecting an occurrence/activity in the group to ideas from the story, or exploring lessons that have been learned from the stories or life experiences.

Note: Repetition of a previously asked question that is clearly a method of facilitating turn taking rather than exploration should be coded as structure.

Topic of exploration:

1. Further exploration of the readings - questions about intentions or predictions for what will happen next in the story or what a character would like to do.

Examples:

“What did the character want?”

“Why do you think the character asked for help?”

“What is the character feeling when she broke her pencil?”

“What’s a pop quiz?” - a concept just introduced in the text.

“How do you think the character will react?”

“What would happen if the character decides to...”

“Is the character’s blood really boiling?”

“Now we know that the character is also...?” – hanging question that is pulling for specific information.

2. Further exploration of a group event - explore the reasons for, or determine the implications of, an event that occurred within the group/among group members. Examples:

“Did it hurt your feelings when he told you your idea was stupid?”

“Why do you disrupt other group members when they are talking?”

“Were you going to say that too?”

“Who was the leader on the way here?”

3. Further exploration of group members’ experiences outside the group - determine implications of how an idea or experience (generated either through readings or group events) relates to one’s life outside the group or broader society. Examples:

“What would happen if you told one of your friends that?”

“Has that ever happened to you?”

“How do you feel inside when you feel guilty?”

“Have you always been a good singer or ...practice a lot?”

“Did it catch on fire?” - in response to a group member’s story about Styrofoam that was placed by a heater.

“What would you do in that situation?”

Initiation of exploration:

1. Spontaneously generated by the group leader - the focus of the intervention is on exploring a new concept. The intervention is not directly related to the content contained in a group member's question or statement immediately preceding the intervention.
2. Given in response to a group member's statement or question

F. Feedback

Definition: comments or reactions to a group member's idea, feeling or behavior that stemmed from the readings or an experience within or outside the group.

Valence of feedback:

1. Positive feedback – explicit comment or reaction that indicates approval or acceptance of a group member's response. Examples:
 - “Yes, that is one way to handle the problem.”
 - “Yeah, we know it's a name like a boy.”
 - “I think that is a very good example.”
 - “You're right Brittney.”
 - “Smart, like Bobbie” (indicates agreement though the extension of the idea).
2. Negative feedback - comment or reaction that indicates disagreement or disapproval of a group member's response. Examples:
 - “Not exactly”
 - “No, the character's name was...”
 - “You're not listening well today.”
3. Neutral feedback – a group member's response is reflected, repeated or acknowledged without an indication of acceptance or disagreement (the inability to take into account nods and other non-verbal forms of communication is a limitation of coding written transcripts). Examples:
 - “Ah.”
 - “Okay, so she understood that...(repeat member's response).”
 - “Well, not exactly, but I can see why you would say that.”

Types of feedback:

1. Simple acknowledgement or disagreement. Examples:
 - “Yes” or “No”
 - “Maybe” or “Probably”
 - “Ah”
 - “Okay” or “Alright”

Note: If a previous member comment is not related to this comment, routine structure would be more appropriate since it was likely said to shift attention back to the discussion or to a different segment of the session.

2. Paraphrase or restatement - a group member's response is repeated or rephrased without changing the meaning of the statement or adding any additional information. Examples:

“Oh, so you already study anyway,” in response to a group member comment, “I already study.”

“Mira,” in response to a group member indicating Mira was the character being referred to in the discussion.

3. Reframing - a group member's response is altered to a more accurate or appropriate answer and/or false information is corrected. Examples:

“Yeah, or they might just think you're a show off, right?”

“Well, it is a little different than that. It is more like...”

“You're right, it does move in that way, but it doesn't sink.”

“I guess so, but sometimes it is hard to remember, isn't it?”

Note: If there is a clear intent to provide information that is not linked to the content contained in the group member's response, a code in the information category would be more appropriate.

4. Elaboration – a group member's response is extended or connected to an additional interpretation/explanation, but is not contradicted or altered.

Examples:

“Yeah, and that tells us that...(extension of member response)”

“You would do that, you wouldn't let yourself get pushed around, right?” following a group member comment that he would have stood up for himself.

Note: If there is a clear intent to further explore the group member's response beyond the content originally contained in the group member's response, a code in the exploration category would more appropriate. For instance, the speaking turn, “Ah, so did you feel guilty about that?” would be divided into two thought units: “Ah,” which is actually neutral feedback using simple acknowledgment and “so did you feel guilty,” which would be coded in the exploration category.

G. Miscellaneous

Used if the above categories are not applicable to the thought unit.

Direction of Intervention: After choosing the type of intervention above and the appropriate subcategories, identify whether the intervention was:

1. Directed toward the entire group - introduction of a new concept or providing an advanced organizer to the group that is not in response to an individual group member's statement or question or directed to a particular group member. Examples:

“Today we are going to start a new book.”
 “What do you all think about...?”
 “Please open your books to page 12.”

2. Directed toward an individual - direct reply or feedback (including paraphrases and restatements) to an individual group member’s statement or question; calling on a particular member to answer a question or provide an opinion. Examples:
 - “Andrea, did you have something to add?”
 - “I don’t think so either.”
 - “Wow. What did you do about it?”

3. Directed toward both - an intervention that was prompted by, or directed toward, an individual but generally applies to the whole group. Information or exploration that immediately follows feedback to an individual, as well as elaborative or reframing feedback initially directed to an individual, fall under this category since the comment may still be directed toward the individual, but is being provided for the benefit of all group members. Examples:
 - “Ah, so you think that..., but couldn’t it also be...”
 - “I understand what you mean and you know what? It’s similar to what Tamesha was talking about earlier.”
 - “His mom’s suspicious, isn’t she Donte?” Immediately following a reading about a character’s mom displaying suspicious behavior.
 - “The speech teacher is using that room today. That is why we are meeting in here.” Following a member’s question, “How come we’re not in the speech room?”

APPENDIX B:
Frequencies and Within-Category Percentages for High and Low Treatment Response Groups across Sessions

Group Process Scale												
CATEGORY	High I Early	High 1 Middle	High 1 Late	High 2 Early	High 2 Middle	High 2 Late	Low I Early	Low 1 Middle	Low 1 Late	Low 2 Early	Low 2 Middle	Low 2 Late
Structure	99 31.7%	69 30.7%	47 24.7%	54 27.6%	54 23.3%	40 18.9%	81 29.7%	72 26.6%	59 27.3%	87 33.7%	98 37.0%	40 16.1%
Group cohesion	11 3.5%	6 2.7%	5 2.6%	4 2.0%	4 1.7%	8 3.8%	7 2.6%	15 5.5%	3 1.4%	8 3.1%	16 6.0%	5 2.0%
Modeling	9 2.9%	4 1.8%	4 2.1%	7 3.6%	2 .9%	8 3.8%	5 1.8%	9 3.3%	5 2.3%	9 3.5%	3 1.1%	8 3.2%
Information	29 9.3%	31 13.8%	47 24.7%	32 16.3%	40 17.2%	36 17.0%	28 10.2%	44 16.2%	41 19.0%	30 11.6%	52 19.6%	52 20.9%
Exploration	68 21.8%	47 20.9%	37 19.5%	45 23.0%	51 22.0%	52 24.5%	77 28.2 %	50 18.5%	49 22.7 %	59 22.9 %	37 14.0%	68 27.3%
Feedback	96 30.8%	68 30.2%	50 26.3%	54 27.6%	81 34.9%	68 32.1%	75 27.5 %	81 29.9%	59 27.3%	65 25.2%	59 22.3%	76 30.5%
Type of Structure												
CATEGORY	High I Early	High 1 Middle	High 1 Late	High 2 Early	High 2 Middle	High 2 Late	Low I Early	Low 1 Middle	Low 1 Late	Low 2 Early	Low 2 Middle	Low 2 Late
Long-term	12 12.1%	1 1.4%	6 12.8%	5 9.3%	5 9.3%	1 2.5%	3 3.7%	10 13.9%	3 5.1%	2 2.3%	0 0%	5 12.5%
Routine management	74 74.7%	62 89.9%	39 83.0%	49 90.7%	48 88.9%	37 92.5%	46 56.8%	43 59.7%	39 66.1%	59 67.8%	58 59.2%	35 87.5%
Behavioral management	13 13.1%	6 8.7%	2 4.2%	0 0%	1 1.8%	2 5.0%	32 39.5%	19 26.4%	17 28.8%	26 29.9%	40 40.8%	0 0%

Type of Group Cohesion												
CATEGORY	High I Early	High 1 Middle	High 1 Late	High 2 Early	High 2 Middle	High 2 Late	Low I Early	Low 1 Middle	Low 1 Late	Low 2 Early	Low 2 Middle	Low 2 Late
Team Building	6 54.5%	3 50.0%	0 0%	1 25.0%	1 25.0%	1 12.5%	2 28.6%	3 20.0%	1 33.3%	6 75.0%	10 62.5%	2 40.0%
Emotional engagement or support	5 45.5%	3 50.0%	5 100%	3 75.0%	3 75.0%	7 87.5%	5 71.4%	12 80.0%	2 66.6%	2 25.0%	6 37.5%	3 60.0%
Type of Modeling												
CATEGORY	High I Early	High 1 Middle	High 1 Late	High 2 Early	High 2 Middle	High 2 Late	Low I Early	Low 1 Middle	Low 1 Late	Low 2 Early	Low 2 Middle	Low 2 Late
Self-disclosure	5 55.6%	2 50.0%	3 75.0%	3 42.9%	1 50.0%	4 50.0%	4 80.0%	5 55.6%	0 0%	2 22.2%	1 33.6%	2 25.0%
Interaction with others	4 44.4%	2 50.0%	1 25.0%	4 57.1%	1 50.0%	4 50.0%	1 20.0%	4 44.4%	5 100%	7 77.8%	2 66.6%	6 75.0%
Initiation of Information												
CATEGORY	High I Early	High 1 Middle	High 1 Late	High 2 Early	High 2 Middle	High 2 Late	Low I Early	Low 1 Middle	Low 1 Late	Low 2 Early	Low 2 Middle	Low 2 Late
Spontaneously offered by the group leader	22 75.9%	21 67.7%	29 61.7%	25 78.1%	22 55.0%	17 47.2%	17 60.7%	28 63.6%	28 68.3%	14 46.7%	27 51.9%	27 51.9%
Provided in response to member comment	7 24.1%	10 32.3%	18 38.3%	7 21.9%	18 45.0%	19 52.8%	11 39.3%	16 36.4%	13 31.7%	16 53.3%	25 48.1%	25 48.1%

Topic of Exploration												
CATEGORY	High I Early	High 1 Middle	High 1 Late	High 2 Early	High 2 Middle	High 2 Late	Low I Early	Low 1 Middle	Low 1 Late	Low 2 Early	Low 2 Middle	Low 2 Late
Readings	44 64.7%	42 89.4%	22 59.5%	27 60.0%	32 62.7%	6 11.5%	27 35.1%	37 74.0%	19 38.8%	26 44.1%	15 40.5%	22 32.4%
Group event	0 0%	5 10.6%	11 29.7%	4 8.9%	6 11.8%	30 57.7%	8 10.4%	1 2.0%	6 12.2%	9 15.2%	12 32.4%	36 52.9%
Member experience	24 35.3%	0 0%	4 10.8%	14 31.1%	13 25.5%	16 30.8%	42 54.5%	12 24.0%	24 49.0%	24 40.7%	10 27.0%	10 14.7%
Initiation of Exploration												
CATEGORY	High I Early	High 1 Middle	High 1 Late	High 2 Early	High 2 Middle	High 2 Late	Low I Early	Low 1 Middle	Low 1 Late	Low 2 Early	Low 2 Middle	Low 2 Late
Spontaneously generated by the group leader	49 72.1%	30 63.8%	27 73.0%	29 64.4%	36 70.6%	23 44.2%	29 37.7%	29 58.0%	28 57.1%	33 55.9%	20 54.1%	38 55.9%
Provided in response to member comment	19 27.9%	17 36.2%	10 27.0%	16 35.6%	15 29.4%	29 55.8%	48 62.3%	21 42.0%	21 42.9%	26 44.1%	17 45.9%	30 44.1%
Valence of Feedback												
CATEGORY	High I Early	High 1 Middle	High 1 Late	High 2 Early	High 2 Middle	High 2 Late	Low I Early	Low 1 Middle	Low 1 Late	Low 2 Early	Low 2 Middle	Low 2 Late
Positive	66 68.8%	27 39.7%	23 46.0%	32 59.3%	37 45.7%	22 32.4%	28 37.3%	34 42.0%	16 27.1%	24 36.9%	20 33.9%	34 44.7%
Negative	1 1.0%	4 5.9%	1 2.0%	1 1.8%	1 1.2%	3 4.4%	9 12.0%	9 11.1%	0 0%	8 12.3%	8 13.6%	1 1.3%
Neutral	29 30.2%	37 54.4%	26 52.0%	21 38.9%	43 53.1%	43 63.2%	38 50.7%	38 46.9%	43 72.9%	33 50.8%	31 52.5%	41 53.9%

Type of Feedback												
CATEGORY	High I Early	High 1 Middle	High 1 Late	High 2 Early	High 2 Middle	High 2 Late	Low I Early	Low 1 Middle	Low 1 Late	Low 2 Early	Low 2 Middle	Low 2 Late
Simple acknowledgement or disagreement	13 13.5%	13 19.1%	17 34.0%	11 20.4%	34 42.0%	30 44.1%	35 46.7%	20 24.7%	20 33.9%	28 43.1%	23 39.0%	35 46.1%
Paraphrase or restatement	25 26.0%	18 26.5%	7 14.0%	9 16.7%	14 17.3%	8 11.8%	15 20.0%	33 40.7%	18 30.5%	18 27.7%	11 18.6%	17 22.4%
Reframing	14 14.6%	17 25.0%	15 30.0%	13 24.1%	9 11.1%	15 22.1%	10 13.3%	16 19.8%	10 16.9%	8 12.3%	14 23.7%	13 17.1%
Elaboration	44 45.8%	20 29.4%	11 22.0%	21 38.9%	24 29.6%	15 22.1%	15 20.0%	12 14.8%	11 18.6%	11 16.9%	11 18.6%	11 14.5%
Direction of Intervention												
CATEGORY	High I Early	High 1 Middle	High 1 Late	High 2 Early	High 2 Middle	High 2 Late	Low I Early	Low 1 Middle	Low 1 Late	Low 2 Early	Low 2 Middle	Low 2 Late
Toward the group	132 42.3%	72 32.0%	80 42.1%	88 44.9%	87 37.5%	61 28.8%	68 24.9%	75 27.7%	77 35.6%	79 30.6%	79 29.8%	70 28.1%
Toward an individual	111 35.5%	85 37.8%	68 35.8%	60 30.6%	84 36.2%	107 50.5%	158 57.9%	133 49.1%	103 47.7%	129 50.0%	143 54.0%	138 55.4%
Both - toward an individual, but intended for group too	69 22.1%	68 30.2%	42 22.1%	48 24.5%	61 26.3%	44 20.7%	47 17.2%	63 23.2%	36 16.7%	50 19.4%	43 16.2%	41 16.5%

REFERENCES

- Beck, A. P., & Lewis, C. M. (2000). Introduction. In A. P. Beck & C. M. Lewis (Eds.), *The process of group psychotherapy: Systems for analyzing change* (pp. 3-19). Washington, DC: American Psychological Association.
- Bednar, R. L., & Kaul, T. J. (1994). Experiential group research: Can the canon fire? In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 631-633). New York: John Wiley & Sons.
- Bergin, A. E., & Garfield, S. L. (1994). Overview, trends, and future trends. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 821-830). New York: John Wiley & Sons.
- Beutler, L. E., Machado, P. P., & Allstetter-Neufeldt, S. (1994). Therapist variables. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 229-269). New York: John Wiley & Sons.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Corey, M. S., & Corey, G. (2002). *Groups: Process and practice* (6th ed.). Pacific Grove, CA: Brooks/Cole.
- Crick, N.R., & Dodge, K. A. (1994). A review and reformulation of social information processing mechanisms in children's social adjustment. *Psychological Bulletin*, 115, 74-101.
- Dagley, J. C., Gazda, G. M., Eppinger, S. J., & Stewart, A. (1994). Group psychotherapy research with children, preadolescents, and adolescents. In A. Fuhrman & G. M.

- Burlingame (Eds.), *Handbook of group psychotherapy: An empirical and clinical synthesis* (pp. 340-369). New York: John Wiley & Sons.
- Dies, R. R. (1994). Therapist variables in group psychotherapy research. In A. Fuhriman & G. M. Burlingame (Eds.), *Handbook of group psychotherapy* (pp. 114-154). New York: Wiley.
- Dwivedi, K. N. (1993). Introduction. In K. N. Dwivedi (Ed.), *Group Work with Children and Adolescents: A Handbook* (pp. 3-13). Bristol, PA: Jessica Kingsley.
- Elliott, R., Hill, C. E., Stiles, W. B., Friedlander, M. L., Mahrer, A. R., & Margison, F. R. (1987). Primary response modes: Comparison of six rating systems. *Journal of Consulting and Clinical Psychology, 58*, 265-272.
- Elliott, R., Hill, C. E., Stiles, W. B., Friedlander, M. L., Mahrer, A. R., & Margison, F. R. (1987). Primary response modes: Comparison of six rating systems. *Journal of Consulting and Clinical Psychology, 58*, 265-272.
- Fuhriman, A., & Burlingame, G. M. (1994a). Group Psychotherapy: Process and Practice. In A. Fuhriman & G. M. Burlingame (Eds.), *Handbook of group psychotherapy: An empirical and clinical synthesis* (pp. 3-39). New York: John Wiley & Sons.
- Fuhriman, A., & Burlingame, G. M. (1994b). Measuring small group process: A methodological application of chaos theory. *Small Group Research, 25*, 502-519.
- Gardner, W. (1995) On the reliability of sequential data: measurement, meaning, and correction. In J. M. Gottman (Ed.) *The analysis of change*. Mahwah, NJ: Erlbaum.

- Green, L. R. (2000). Process analysis of group interaction in therapeutic groups. In A. P. Beck & C. M. Lewis (Eds.), *The process of group psychotherapy: Systems for analyzing change* (pp. 23-47). Washington, DC: American Psychological Association.
- Hill, C. E. (1990a). Exploratory in-session process research in individual psychotherapy: A review. *Journal of Consulting and Clinical Psychology, 58*, 288-294.
- Hill, C. E. (1990b). Is individual therapy process really different from group therapy process? The jury is still out. *The Counseling Psychologist, 18*, 126-130.
- Hill, C. E. (1991). Almost everything you ever wanted to know about how to do process research on counseling and psychotherapy but didn't know who to ask. In C. E. Watkins, Jr. & L. J. Schneider (Eds.), *Research in counseling* (pp. 85-118). Hillsdale, NJ: Lawrence Erlbaum.
- Hill, C. E. (1992). An overview of four measures developed to test the Hill Process Model: Therapist intentions, therapist response modes, client reactions, and client behaviors. *Journal of Counseling and Development, 70*, 728-739.
- Hill, C. E., & O'Brien, K. M. (1999). *Helping skills: Facilitating exploration, insight, and action*. Washington, DC: American Psychological Association.
- Hoag, M. J., & Burlingame, G. M. (1997). Child and adolescent group psychotherapy: A narrative review of effectiveness and the case for meta-analysis. *Journal of Child and Adolescent Group Therapy, 7*, 51-68.
- Holmes, S. E., & Kivlighan, D. M., Jr. (2000). Comparison of therapeutic factors in group and individual treatment processes. *Journal of Counseling Psychology, 47*, 478-484.

- Horne, A. M., & Rosenthal, R. (1997). Research in group work: How did we get where we are? *The Journal for Specialists in Group Work, 22*, 228-240.
- Kivlighan, D. M., Jr., Multon, K. D., & Brossart, D. F. (1996). Helping impacts in group counseling: Development of a multidimensional rating system. *Journal of Counseling Psychology, 47*, 43, 347-355.
- Lambert, M. J., & Hill, C. E. (1994). Assessing psychotherapy outcomes and processes. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 72-113). New York: John Wiley & Sons.
- Leichtentritt, J., & Shechtman, Z. (1998). Therapist, trainee, and child verbal response modes in child group therapy. *Group Dynamics: Theory, Research, and Practice, 2*, 36-47.
- Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Development, 71*, 107-118.
- Marmar, C. R. (1990). Psychotherapy process research: Progress, dilemmas, and future directions. *Journal of Consulting and Clinical Psychology, 58*, 265-272.
- Marshall, S. P. (1995). *Schemas in Problem Solving*. New York: Cambridge University.
- McGrath, P. (1993). Historical Development of group psychotherapy. In K. N. Dwivedi (Ed.), *Group Work with Children and Adolescents: A Handbook* (pp. 100-114). Bristol, PA: Jessica Kingsley.
- McGrath, P. (1993). Historical Development of group psychotherapy. In K. N. Dwivedi (Ed.), *Group Work with Children and Adolescents: A Handbook* (pp. 100-114). Bristol, PA: Jessica Kingsley.

- McGuire, J. M., Taylor, D. R., Broome, D. H., Blau, B.I., & Abbott, D. W. (1986). Group structuring techniques and their influence on process involvement in a group counseling training program. *Journal of Counseling Psychology, 33*, 270-275.
- McLeod, J. (1994). *Doing counseling research*. London: Sage Publications.
- Morran, D. K., Stockton, R., & Harris, M. (1991). Analysis of group leader and member feedback messages. *Journal of Group Psychotherapy, Psychodrama, and Sociometry, 44*, 126-135.
- Posthuma, B. W. (1996). *Small groups in counseling and therapy: Process and leadership* (2nd ed.). Needham Heights, MA: Allyn and Bacon.
- Rahill, S., & Teglassi, H. (2003). Processes and outcomes of story based and skill-based social competency programs for children with emotional disabilities. *Journal of School Psychology, 41*, 413-429.
- Reynolds, C. R., & Kamphaus, R. W. (1992). *The Behavior Assessment System for Children*. Circle Pines, MN: American Guidance Service.
- Riva, M. T., & Smith, R. D. (1997). Looking into the future of group research: Where do we go from here? *The Journal for Specialists in Group Work, 22*, 266-267.
- Shechtman, Z. (1999). Bibliotherapy: An indirect approach to treatment of childhood aggression. *Child Psychiatry and Human Development, 30*, 39-53.
- Shechtman, Z. (2003). Therapeutic factors and outcomes in group and individual therapy of aggressive boys. *Group Dynamics: Theory, Research and Practice, 7*, 225-237.
- Shechtman, Z., & Ben-David, M. (1999). Individual and group psychotherapy of childhood aggression: A comparison of outcomes and process. *Group Dynamics: Theory, Research, and Practice, 3*, 263-274.

- Shechtman, Z., & Yanov, H. (2001). Interpretives (Confrontation, Interpretation, and Feedback) in preadolescent counseling groups. *Group Dynamics: Theory, Research, and Practice*, 5, 124-135.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Teglasi, H. (2001). Essentials of TAT and other storytelling techniques assessment. In A.S. Kaufman & N. L. Kaufman (Series Eds.), *Essentials of psychological assessment series*. New York: John Wiley & Sons.
- Teglasi, H., & Rothman, L. (2001). STORIES: A classroom-based program to reduce aggressive behavior. *Journal of School Psychology*, 39, 71-94.
- Trotzer, J. P. (1999). *The counselor and the group: Integrating theory, training, and practice* (3rd ed.). Ann Arbor, MI: Edwards Brothers.
- Wheelan, S. A. (1994). *Group processes: A developmental perspective*. Boston: Allyn and Bacon.
- Yalom, I. D. (1995). *The theory and practice of group psychotherapy* (4th ed.). New York: Basic Books.