

Intervention Groups for Adolescents With Conduct Problems: Is Aggregation Harmful or Helpful?

Wendy Mager,¹ Richard Milich,^{1,2} Monica J. Harris,¹ and Anne Howard¹

Received January 13, 2004; revision received August 20, 2004; accepted August 27, 2004

Past research has suggested that the aggregation of deviant peers during treatment may cause harmful effects (T. J. Dishion, J. McCord, & F. Poulin, 1999). This study compared the effectiveness of problem-solving skills training groups in which all members had conduct problems (“pure” group condition) with groups that consisted of adolescents with and without behavior problems (“mixed” group condition). Participants were 139 sixth- and seventh-graders (mean age = 12.7; 63% male; 55% White) enrolled at public middle schools. Pre-, post-, and 6-month follow-up data were collected, along with intervention process variables. Results showed that, contrary to the deviancy training hypothesis, adolescents in the pure-group condition engaged in more adaptive in-session behavior and received lower scores on the parent and teacher ratings of externalizing behavior at postintervention than those in the mixed-group condition. Mediation analyses showed that the deviancy training that occurred in the mixed-group condition accounted for their worse postintervention scores, findings consistent with the deviancy training hypothesis.

KEY WORDS: conduct problems; group intervention; iatrogenic effects; secondary prevention; deviancy training.

Aggressive and antisocial behavior problems exact a high price from individuals and society at large. Although past research has been successful in documenting a myriad of factors that contribute to the development of antisocial behavior in childhood and/or adolescence (Dishion, French, & Patterson, 1995), less scientific data are available to answer the question of how best to design interventions to prevent or reduce antisocial behavior. One hypothesis that has attracted considerable attention in the literature is that social skills interventions for at-risk adolescents that occur within small groups may be contraindicated or lead to iatrogenic effects (Dishion, McCord, & Poulin, 1999). The purpose of the current study is to provide an empirical test of this hypothesis.

Skills-training programs are predicated on theories that children and adolescents with conduct problems have deficiencies in social skills, anger control, interpersonal

problem-solving, and moral reasoning and that remediation of these skills will lead to improved overall functioning and prognosis. By and large, recent research supports the efficacy of such programs (McMahon & Wells, 1998). Skills-training can be provided in small groups of high-risk youth (“pure” or “homogeneous” groups), small groups consisting of both low- and high-risk youths (“mixed” or “heterogeneous” groups), or individual training sessions between a high-risk youth and a therapist. There is a general consensus that the small-group format for skills-training interventions tends to be more efficacious than individual training sessions (Landau, Milich, & Diener, 1998). This is because of the cost-effectiveness of group treatment and the value in providing a social context for the mastery of social and emotional skills (i.e., more opportunities for role-playing activities, modeling, and peer reinforcement of adaptive behavior). It may be counterproductive, however, to aggregate high-risk youths with externalizing problems into pure treatment groups. This concern is based on the “acting out” nature of externalizing problems, combined with the great importance placed on peers during the adolescent stage of development. Some researchers (e.g., Dishion et al., 1999)

¹Department of Psychology, University of Kentucky, Lexington, Kentucky.

²Address all correspondence to Richard Milich, PhD, Department of Psychology, University of Kentucky, 202A Kastle Hall, Lexington, Kentucky 40506-0044; e-mail: milich@uky.edu.

argue that the deviant peer environment provided in pure treatment groups may interfere with the prosocial skills training process, resulting in the continuation or even exacerbation of conduct problems for the youth involved.

THE IATROGENIC EFFECTS HYPOTHESIS

This argument of possible “iatrogenic” effects of group treatment for adolescent conduct problems has been raised occasionally throughout the history of skills-training programs (e.g., Arnold & Hughes, 1999; Feldman, Caplinger, & Wodarski, 1983; McMahon & Wells, 1998). In fact, the notion of negative group treatment effects for this population is largely accepted, even taken somewhat for granted, by many within the scientific community. For example, Alan Kazdin, a leading figure in the treatment of conduct disorders, was recently quoted as saying, “A number of treatments that harm children are being used; group therapy for behavior disordered children is an absolute no-no. There have been a number of studies that show that if you place these children in groups, those children who receive the treatment do much worse” (Weaver, 2000).

However, only recently has the iatrogenic hypothesis been carefully articulated and the relevant research reviewed (Dishion et al., 1999). In an *American Psychologist* article, Dishion et al. (1999) examined some of the social developmental and intervention literatures that bear on the popular notion of iatrogenic effects of skills-training treatments that aggregate high-risk youth. Dishion et al. (1999) concluded that there is considerable cause for concern. Specifically, they argued that in treatment groups made up entirely of aggressive, deviant adolescents, there may be more modeling and reinforcement of antisocial behavior than prosocial behavior. They asserted that this, in turn, seems to lead to the further development of conduct problems in youth, both within and outside the treatment group setting (Dishion et al., 1999). The evidence for this negative treatment process and outcome stems primarily from their own research on “deviancy training” and the outcome of their Adolescent Transitions Program intervention study.

Deviancy training refers to “the process of contingent positive reactions to rule-breaking discussions,” which Dishion and colleagues argue perpetuates or “trains” deviancy in one adolescent by another (Dishion et al., 1999, p. 756). An example of deviancy training might be an adolescent repeatedly gaining positive peer attention (e.g., expressions of interest, laughter) in response to talking about rebelling against parents, using drugs, or engaging in aggression or violence. Dishion, Spracklen, Andrews,

& Patterson (1996) studied this process by analyzing the topics and response patterns of 206 adolescent male friendship dyads during 25-min videotaped discussions. They found that social reinforcement within the dyad (i.e., positive affect or gesture) was strongly associated with conversation topic. For nondelinquent youth dyads, social reinforcement tended to follow normative discussion; for delinquent dyads, the reinforcement occurred more often for deviant, rule-breaking conversation. Furthermore, in follow-up studies of these youth, they found strong relations between the nature of youths’ dyadic interactions at age 13–14 years and the rate of problem behaviors reported 2 years later. Boys whose conversations were characterized by deviancy training had a higher probability of substance use initiation, higher self-reported delinquency, and higher self-reported and police-reported violent behavior than dyads characterized by peer reinforcement of normative discussions (Dishion, Capaldi, Spracklen, & Li, 1995; Dishion, Eddy, Haas, Li, & Spracklen, 1997; Dishion et al., 1996).

This body of research seems adequate to suggest that deviancy training may occur within the friendships of high-risk male adolescents. Dishion et al. (1999) went on to suggest that this process also occurs within group skills-training interventions and that it interferes with the therapeutic effectiveness of such interventions. Dishion and Andrews (1995) provided evidence of this effect from the results of the Adolescent Transitions Program (ATP), an intervention aimed at reducing the problem behavior of early adolescents. Participants were middle-school students (mean age = 12 years) who were identified as “high-risk” on the basis of a 10-question screening instrument completed by a parent. Researchers randomly assigned the families of these high-risk adolescents to one of the following conditions: (1) parent-focus group, (2) teen-focus group, (3) both parent- and teen-focus groups, (4) self-directed change (provided intervention materials only), or (5) quasi-experimental, no-treatment control.

Results indicated that there were limited, immediate postintervention benefits to participation in the parent and teen groups (i.e., less coercive family interactions and improved school behavior) and limited long-term benefits to the parent intervention (reduced teen tobacco use). However, there were a few long-term, iatrogenic effects associated with teen participation in the intervention (i.e., either teen-only or teen- and parent-focus groups). These adolescents had higher teacher ratings of externalizing behavior problems, more positive attitudes toward drug use, and more self-reported tobacco use at 1- and 3-year follow-ups (Poulin, Dishion, & Burraston, 2001).

Although the evidence for iatrogenic effects of group intervention sounds convincing at first glance, there are

limitations to the conclusions that can be drawn from the Dishion and Andrews (1995) study. Specifically, the iatrogenic effects were limited to a small minority of dependent variables. So, a more conservative interpretation of the Dishion and Andrews study is that there were small positive effects for some variables, small negative effects for others, and no effects for many others. Thus, the case for clinically significant, iatrogenic effects due to peer aggregation in group-based skills-training interventions is far from proven.

In summary, then, the hypothesis put forth by Dishion et al. (1999) of possible iatrogenic effects of all-deviant group skills-training programs for externalizing youth is provocative. If true, it could have far-reaching implications for clinical practice. However, a critical view of the literature cited by Dishion et al. (1999) indicates that the evidence for risk may be overstated. In fact, other researchers have subsequently examined the literature and concluded that it does not support the iatrogenic effects hypothesis (Handwerk, 2000).

A ROLE FOR PROSOCIAL PEERS IN INTERVENTION GROUPS

The deviancy training hypothesis suggests that pure groups of antisocial adolescents may be counterproductive. This raises the question of whether mixed groups of antisocial and prosocial participants would be more effective. That is, perhaps high-risk youths show greater gains in prosocial behavior and reduction in problem behavior if they participate in skills-training curricula alongside socially competent, prosocial youths. This represents the flip side of the deviancy training hypothesis. Youths may experience "prosocial training" as opposed to deviancy training, by way of more peer modeling and reinforcement of adaptive behavior.

We were able to locate only one published study that attempted a direct comparison of mixed versus pure groups for adolescents with conduct problems (Feldman et al., 1983). In "the St. Louis Experiment," a total of 701 boys (age 7–15 years) participated in activity groups of 9–15 members at a local Jewish community center. Boys with conduct problems ($n = 263$) were recruited from social service agencies. A boy met criteria for study inclusion if both the referral agent and the boy's parent indicated on a checklist that the boy had engaged in at least 21 antisocial behaviors in the past 7 days. "Antisocial behaviors" were defined as those intended to hurt, disrupt, or annoy others. They included antisocial motor behaviors, physical contacts, verbalizations, and distracting behaviors. Participants were randomly assigned to

either all-deviant treatment groups or to groups in which they were the only high-risk boy. Groups met once a week, for a period of 2–3 hr, for approximately 1 school year (mean number of sessions = 22). The primary finding related to the effect of group composition was that adolescents who had participated in the mixed group condition showed less antisocial in-session behavior (according to nonparticipant group observers) than adolescents who had participated in the pure group condition. Also, mixed group participants self-reported that their out-of-session behavior improved over the course of treatment, whereas all-deviant group participants reported increased antisocial behavior over this same time period. Finally, there was an interaction between group composition and leader experience, such that participants in pure groups with inexperienced leaders fared the worst, according to ratings by group observers and the boys.

Although the Feldman et al. (1983) study is noteworthy because it is the only empirical work to date that directly tested mixed versus pure intervention groups, there are reasons to be cautious in interpreting the results of this study. First, the study only had sufficient outcome data from the perspective of the boys, themselves, and the group observers. These data may have suffered from demand characteristics or other biases. The results would be more compelling if they included outside confirmation of behavioral changes (e.g., by parent or teacher report of the child's general, out-of-group behavior). Second, the composition of Feldman et al.'s mixed groups was not representative of typical intervention studies. From a practical standpoint, few treatment settings can advocate the use of 7–14 prosocial youth for the treatment of every one aggressive youth. Finally, Feldman et al.'s intervention modality is better described as "activity groups" than as "psychological treatment groups." Children had a wide range of group activity options, group leaders received very little training in treatment techniques, results of observer ratings demonstrated that there was considerable overlap between techniques used across the three treatment conditions, and treatment was not manualized. In these ways, the intervention conducted by Feldman et al. was very different from the standard of practice in current psychological treatment settings.

Thus, it seems clear that more research is needed to address the question of the comparative efficacy of pure versus mixed groups for skills-training interventions with adolescents at high-risk for antisocial behavior problems (e.g., Arnold & Hughes, 1999; Dishion et al., 1999; McMahon & Wells, 1998). This study provides such a test and represents an important step toward understanding whether group composition is a treatment moderator. The following hypotheses were set forth: (1) The

mixed skills-training condition will be more effective for high-risk youth than the pure group skills-training condition; (2) Group processes of peer modeling and reinforcement will be more adaptive in the mixed group condition than in the pure group condition; and (3) Group processes will mediate the effects of group composition on outcomes.

METHOD

Design

This study employed a mixed design, including both between- and within-participants factors. Children with externalizing behavior problems (“at-risk”) were randomly assigned to one of two conditions: a problem-solving group intervention with only at-risk members (“pure” group condition) or one with mixed membership (i.e., both at-risk children and children without conduct problems; “mixed” group condition). Psychosocial functioning was assessed within 3 months prior to the start of the intervention (Time 1 or preintervention), 3–5 months later, within 1 month following the end of the intervention (Time 2 or postintervention), and 6 months after the end of intervention (Time 3 or follow up).

Participants

Participants were 139 sixth- and seventh-grade students (63.1% male and 36.9% female) enrolled at four public middle schools in Lexington, KY. Sixty-five of the participants comprised the at-risk sample of interest, and the remaining 74 participants made up the prosocial sample. The age range for the at-risk sample was 11.0–17.8 years ($M = 12.7$, $SD = .97$).³ The racial composition was 55.4% White, 35.4% Black, and 9.2% other. The age range for the prosocial sample was 11.4–14.3 years ($M = 12.5$, $SD = .64$). The racial composition was 87.8% White, 9.5% Black, and 2.8% other. The gender composition was 52.7% male and 47.3% female. There was no difference in age or gender between the at-risk and prosocial samples ($t = 1.46$, $df = 137$, $p = .15$; $\chi^2 = 2.65$, $df = 1$, $p = .10$). The at-risk sample consisted of more Black participants than the prosocial sample ($\chi^2 = 57.76$, $df = 2$, $p < .01$).

³All but one of the participants was less than 14.5 years of age; however, one participant was randomized into the study who was 17-years old. Inspection of this individual’s data indicated that his or her responses were not outliers; thus, we decided to preserve the integrity of our randomization procedures and retain this individual in analyses.

Procedure

For logistic reasons, the study was conducted across two cycles. The first cycle began at two public middle schools in the Spring 2001 semester; the second cycle began at two other public middle schools in the Fall 2001 semester. A five-step process for conducting the study at the host schools was established. The first step was screening and recruiting appropriate children. Host schools mailed a letter describing the study to the parents of all sixth- and seventh-grade students ($N = 1926$) enrolled at their schools. Interested parents were asked to sign an enclosed reply postcard and mail it to the researcher. A total of 357 postcards were received, for an overall response rate of 19%, with comparable rates across the four schools.

One teacher was selected for each student and was asked to complete a brief screening questionnaire about the student. The screening questionnaire consisted of 13 items drawn from the Externalizing and Internalizing Composites of the Child Behavior Checklist—Teacher Report Form (CBCL-TRF; Achenbach, 1991b; see Measures section for more detail). A total of 329 teacher questionnaires were completed. Teachers received \$2 per questionnaire completed.

Children’s eligibility for study participation was based on teacher ratings. To qualify as at-risk, a child’s score on the teacher-rated scale for externalizing behavior had to correspond to a T -score of ≥ 58 (78th percentile). Prosocial children had teacher-rated externalizing scale T -scores of ≤ 54 (66th percentile) and internalizing T -scores of ≤ 56 (72nd percentile). These criteria resulted in a total of 215 students being identified as eligible for the study (83 at-risk and 132 prosocial).

The parents of eligible children were contacted by telephone, invited to attend a small group meeting with the researcher. At these group meetings, parents and children were told the general purpose of the study, were invited to ask questions, then completed consent/assent forms and Time 1 questionnaires if they decided to participate. Parents and children were each paid \$5 for completing the questionnaires, which took approximately 15–30 min (see Measures section for more details about the questionnaires).

A power analysis was conducted in order to estimate the ability to detect true effects of group composition using this research design. Results suggested that, for a two-group ANOVA, anticipating a medium effect size, with an alpha of .05, the power of the study is estimated at .66.

Next, participants were assigned to experimental conditions. The at-risk children were randomly assigned

to either the mixed group condition ($n = 40$) or the pure group condition ($n = 25$). The disproportionate number of adolescents assigned to the mixed group condition resulted from difficulty recruiting enough at-risk children and a desire to maintain the 2:1 ratio of prosocial to at-risk students within mixed intervention groups while keeping group size relatively constant. Participants were assigned to one of 16 intervention groups (four groups at School 1, four at School 2, five at School 3, and three at School 4). The number of students assigned to a group ranged from 5–11 ($M = 8.69$, $SD = 1.96$). Four of the 16 intervention groups were “pure” groups (i.e., consisted entirely of at-risk adolescents). The remaining 12 intervention groups were “mixed” groups (i.e., consisted of prosocial and at-risk adolescents in a 2:1 ratio).

Intervention groups met once per week, in school classrooms, after school, for 70–90 min, for 9–12 weeks (variations in duration of intervention reflect differences in school schedules). Each group was led by two graduate students in psychology (enrolled in clinical, counseling, or school psychology masters or doctoral programs). Leaders had approximately 6–10 h of training in the intervention prior to the start of the group program. They attended weekly supervision meetings with a licensed clinical psychologist throughout the group program. Each group was also assigned two undergraduate research assistants whose responsibilities included collecting observational data during sessions and providing summary behavior feedback to group members at the end of each session.

The content and structure of the intervention was the same for all intervention groups. Program content centered around teaching social problem-solving skills. The manual used by group leaders was modeled after the “Social Competence Promotion Program” (Weissberg, Caplan, Benetto, & Jackson, 1990), a universal prevention program used in many schools and shown to be effective in reducing problem behavior in young adolescents (Caplan et al., 1992). Direct instruction, structured discussions, games, worksheets, and role-plays were used to teach participants to use a systematic approach for handling a range of social problems such as teasing, peer pressure, and disagreements with teachers, parents, or peers. Group leaders and observers completed checklists at the end of each session, documenting the material that was covered in the session. According to these data, participants were exposed to 80–100% of the program content each session ($M = 93.98\%$, $SD = 7.15\%$).

Behavior management during the sessions was also structured. Groups had five rules for in-session behavior: (1) Participate in group activities; (2) compliment others; (3) don’t disrupt the group; (4) don’t disrespect others; and

(5) follow directions. At the beginning of each session, leaders stated a behavior goal for group members for the session (e.g., “at least 10 participations and no more than two disruptions”). During the session, group leaders provided oral feedback to adolescents about their rule-related behavior, praising rule-following behavior and pointing out rule violations in a neutral manner. Adolescents were rewarded simply for attending group sessions; at the end of each session, they were paid cash for their attendance. The guaranteed rate of pay was \$3 per session; an extra incentive to attend sessions regularly was the possibility of earning \$5 for attending on a “Bonus Day” (these days were not announced in advance).

Postintervention questionnaire data were collected from teachers, parents, and adolescents at the end of the intervention group program. Teachers were again paid \$2 per questionnaire they completed. One teacher questionnaire was obtained for 131 of the 139 adolescent participants in the study (94%), including 59 of the 65 at-risk target participants (91%). Parents were contacted by telephone and scheduled for a time to complete the postintervention questionnaire. Parents were paid \$5 for completing the questionnaire. Time 2 questionnaire data were obtained from parents of 126 of the 139 adolescents (91%), including 58 of the 65 at-risk participants (89%).

Postintervention questionnaires were administered to adolescent participants at the beginning of their last intervention group session. Group leaders read the questionnaire aloud to all groups. Children were paid \$5 for completing the questionnaire. Time 2 questionnaire data were obtained from 125 of the 139 children, including 57 of the 65 at-risk children (88%).

Approximately 6 months after the end of the intervention, a similar procedure was used to gather the Time 3 questionnaire data from teachers, parents, and children. Parents were contacted by telephone and appointments were made for them and their children to complete the questionnaires. Time 3 questionnaire data were obtained from 68 of the 139 parents and children (49%), including 40 of the 65 at-risk adolescents (62%). Time 3 teacher questionnaire data were obtained for 80 of the 139 total participants (58%), including 47 of the 65 at-risk children (72%).

Measures

Externalizing Behavior Problems

The Child Behavior Checklist scales (CBCL; Achenbach, 1991a) were used to assess participants’ aggressive and delinquent behavior. The CBCL scales consist of 138 behaviors that are rated on a 3-point scale

(0 = *not true*; 1 = *somewhat or sometimes true*; 2 = *very true or often true*). Teachers completed an abbreviated version of the Internalizing and Externalizing composites of the Teacher Report Form (CBCL-TRF; Achenbach, 1991b). The nine items with the highest degree of discriminant validity were selected for the Externalizing scale, and the four items with the highest discriminant validity were selected for the Internalizing scale. Parents completed the Externalizing composite of the standard form of the CBCL (Achenbach, 1991a). Children completed the Externalizing composite of the Youth Self-Report (CBCL-YSR; Achenbach & Edelbrock, 1991). Previous research with the CBCL provides evidence of good internal consistency, with median coefficient alphas of .76 for the subscales and .92 for the composites; test-retest reliabilities ranged from .70 to .95 (Achenbach, 1991a). In this study, alphas ranged from .69 to .95.

Attitudes About Alcohol and Drugs

As actual alcohol and drug use is very low among 11–14-year-olds, participants' alcohol- and drug-related attitudes and intentions were measured. An abbreviated version of a scale developed by Hoyle, Donohew, Lorch, and Palmgreen (1997) was used. It consisted of six items that measure the respondent's general attitudes toward drugs and alcohol, and had a 4-point response scale. Sample items include "People my age should not use alcohol or drugs" and "Using drugs or alcohol may make you do things you will later regret." At Time 1, coefficient alpha was .72; at Time 2, it was .65; at Time 3, it was .72.

Attitudes Toward Delinquency

Participants responded to a 15-item scale that asked them to judge the extent to which it was wrong to engage in a variety of antisocial acts. There was a 4-point response scale and a "don't know" option. The scale was originally developed for the Pittsburgh Youth Study (Farrington & Loeber, 1996). Alphas were .87, .93, and .78, respectively.

Social Skills

The Social Skills domain of the Social Skills Rating System (SSRS; Gresham & Elliott, 1990) was used to assess participants' social competencies. The scale consists of 30–38 items that are rated on a 3-point scale (0 = *never*; 1 = *sometimes*; 2 = *very often*). The items yield scores on five scales: Cooperation, Responsibility, Self-control, Assertion, and Empathy. Total scores from the SSRS Par-

ent and Student Forms were used in the analyses. The CBCL and SSRS items were interspersed on the parent and child questionnaires, in an effort to prevent a decline in respondent morale that might be associated with having to answer many "problem behavior" (i.e., CBCL) items consecutively. The SSRS has been found to be adequately reliable and valid (Kamphaus & Frick, 1996). In this study, alphas were as follows: .80 for child Time 1; .76 for child Time 2; .60 for child Time 3; .82 for parent Time 1; .78 for parent Time 2; and .72 for parent Time 3.

School Adjustment

Data from participants' school records were provided to the researcher by staff in the school district's administrative offices. Variables calculated from these data included: number of disciplinary actions, number of unexcused absences, number of times tardy to class, and average grade. These four variables were calculated for all participants for three semesters (prior to, during, and immediately after the intervention group program).

Peer Association

We administered a questionnaire to adolescents at the beginning of the first and the last intervention group sessions that assessed their familiarity with, and liking for, each of the other group members. Participants used a 4-point scale to answer three questions about each group member (how well he/she knew the other group member, how much free time he/she spent with the other group member outside of group time, and the extent to which he/she liked the other group member). Then the average score the respondent provided regarding the at-risk children in his or her group was calculated. This resulted in three variables for each respondent—knowing, liking, and spending time with at-risk group members—at each of the three assessment points.

Behavioral Observations

All intervention sessions were attended by two undergraduate research assistants who independently observed and recorded information about the behavior that was exhibited by each group member. Assistants recorded the number of times each child *violated* three group rules ("follow directions," "respect others," and "don't disrupt group activities") and the number of times each child *followed* the remaining two rules ("participate in group activities" and "compliment others") for the

entire session (excluding pregroup, postgroup, and break time).

Assistants were trained to record children's behavior prior to the start of the group program. Assistants were tested to ensure memorization of behavior category definitions and to allow practice using the definitions to code hypothetical behavior. To obtain one estimate per rule per child, scores were first averaged across all group sessions the child attended. Interrater reliability coefficients for assistants' coding of the five variables were calculated. These correlations ranged from .41 to .92, with a median of .81. Thus, scores were averaged across raters. The three negative behavior variables (i.e., disrespect, disrupt, and failure to follow directions) were strongly correlated (r 's ranging from .48 to .63), so they were collapsed into one composite variable called *negative in-session behavior*. Likewise, the two positive behavior variables (i.e., compliment, participate) were correlated ($r = .36$), so they were collapsed into a composite variable called *positive in-session behavior*.

Rating Scales

Immediately after each group session, group leaders and assistants completed a brief questionnaire designed to assess adolescents' general behavior during the session as well as the type and extent of peer reinforcement that occurred during the session. Group leaders and assistants were asked to provide ratings of each child who attended the session on six variables. The variables were active and appropriate participation in group activities; disruptive and disobedient during group activities; gave peers reinforcement for appropriate behavior; gave peers reinforcement for inappropriate behavior; received peer reinforcement for appropriate behavior; and received peer reinforcement for inappropriate behavior. Response options ranged from 1 (*not at all*) to 5 (*very often*). Ratings were averaged across all sessions that a child attended, then interrater reliability was calculated for the four raters (two group leaders and two assistants). Alphas for these six variables ranged from .89 to .95, with a median of .91. Scores were thus averaged across all four raters.

Giving and receiving peer reinforcement for appropriate behavior were strongly correlated ($r = .92$), so they were averaged to create one variable (peer reinforcement for appropriate behavior). Giving and receiving peer reinforcement for inappropriate behavior were also strongly correlated ($r = .90$), and were averaged to create one variable (peer reinforcement for inappropriate behavior).

Attrition

In considering attrition, we used an *intent to treat* criterion, in which participants were included in the study as long as they had any posttest or follow-up data, regardless of how many treatment sessions they may have attended. This is considered a conservative, although appropriate, criterion because the outcome data generated are most representative of what one could expect from real clinical practice. Using this approach, only two at-risk adolescents had no Time 2 data available, and only seven had no follow-up data available. These numbers are too small to conduct formal attrition analyses.

RESULTS

Overview of Design and Hypotheses

This study used a mixed design to test the effects of group composition on intervention processes and outcomes. Adolescents with conduct problems were randomly assigned to either the pure-group condition (i.e., an intervention group that consisted of only other at-risk adolescents) or the mixed-group condition (i.e., an intervention group made up of adolescents with and without conduct problems, in a 1:2 ratio). Data about how the participants related to one another during the intervention groups (i.e., "intervention process variables") were gathered throughout the course of the training sessions. Data about participants' general psychosocial functioning were gathered prior to the intervention, immediately following the intervention, and at six-month follow-up. The study hypotheses were (1) The mixed-group condition will be more effective for high-risk youth than the pure-group condition; (2) intervention processes of peer modeling and reinforcement will be more adaptive in the mixed-group condition than in the pure-group condition; and (3) group processes will mediate the effects of group composition on outcomes.

Before directly addressing these research questions, other analyses are presented. First, hierarchical linear modeling (HLM) techniques were used to assess the appropriateness of traditional statistical approaches for analyzing this data set. Next, a comparison of the prosocial and at-risk samples on group process variables and psychosocial functioning at preintervention is presented to help document the validity of the group classifications. As a check of randomization, the participants assigned to the mixed- versus the pure-group condition are then compared on all dependent variables at Time 1 (preintervention). Finally, analyses addressing each of the three study hypotheses are described. Except for a few instances, which

are noted otherwise, all analyses are based on data from the at-risk participants (i.e., prosocial participants are excluded). In other words, the same data were collected for the prosocial sample as were collected for the at-risk sample in order for all participants to have the same experiences with regard to the study. However, because the at-risk sample was the target of the study (and thus the only sample subject to an experimental manipulation), data analyses were only conducted on the at-risk sample.⁴

Group- Versus Individual-Level Analyses

Because the children participated in groups, these data were nonindependent and thus could potentially lead to erroneous findings if only traditional statistical analyses were conducted, in which the individual is treated as the unit of analysis. To determine the extent of the nonindependence of the data, hierarchical linear modeling (HLM) was used to test for effects of intervention group on the dependent variables at Time 2. An equation that modeled variation at the intervention group level was generated for each of the 15 dependent variables. Only one equation (teacher externalizing) was significant; however, the equation became nonsignificant when group composition (the IV) was entered into the equation. Therefore, these HLM results indicate that traditional statistical analyses (i.e., those that treat the individual as the unit of analysis) are appropriate for this study, despite the nesting of participants within intervention groups. All subsequent analyses thus adopt individual as the unit of analysis.

Differences Between At-Risk and Prosocial Samples

Independent groups *t*-tests were conducted to determine the extent of differences between the at-risk and prosocial samples on measures of in-session behavior and psychosocial functioning at preintervention. Table I shows all means, standard deviations, *t*-values, and effect sizes. At preintervention, at-risk participants scored worse than prosocial participants on all indicators of psychosocial functioning. In terms of in-session behavior, the at-risk participants had higher rates of negative behavior, including showing more disrespect toward others and being more disruptive. They also had higher scores than the prosocial sample on peer reinforcement for inappropriate behavior, especially for receiving that type

of reinforcement. The at-risk sample had higher rates of positive in-session behavior, including more participation than the prosocial sample. Finally, an examination of school-record data revealed that the at-risk sample had significantly lower grades, more behavioral infractions, more unexcused absences, and they were late to class significantly more often than the prosocial sample. Thus, these data confirm the validity of the at-risk and prosocial assignments made in the study.

Differences Between Mixed- and Pure-Group Participants at Preintervention

T-tests were conducted to determine if there were differences between at-risk participants in the mixed-group condition and at-risk participants in the pure-group condition on the dependent variables at baseline. For none of the 11 analyses undertaken did the two groups differ significantly. Thus, it appears that efforts to randomize the participants were effective and the participants in the two conditions were well-matched on baseline measures.

Tests of Hypothesis #1

The first hypothesis was that the mixed-group condition would result in more improvement in at-risk adolescents' social behavior than the pure-group condition. To address this hypothesis, an ANCOVA was conducted on each of the 11 dependent variables collected at Time 2, using the Time 1 scores as the covariate. At postintervention, there were two significant effects and one marginally significant effect of group composition. Contrary to the hypothesis, after controlling for baseline scores, participants in the pure-group condition had lower parent ratings of externalizing behavior at postintervention than did participants in the mixed-group condition, $F(1, 55) = 7.43, p < .01, r = .34$ (see Fig. 1). Post hoc tests revealed that the mixed-group condition did result in decreased scores on this measure ($t = 2.77, p < .01, r = .42$); however, the decrease was greater for the pure-group ($t = 3.58, p < .01, r = .62$). Similarly, after controlling for baseline measures, participants in the pure-group condition had lower teacher ratings of externalizing behavior problems than participants in the mixed-group condition, $F(1, 59) = 7.96, p < .01, r = .34$ (see Fig. 2). The mixed group did not change from pre to post intervention ($t = 1.35, p = .18, r = .22$); however, the pure group decreased significantly ($t = 3.19, p < .01, r = .57$). A marginally significant ANCOVA indicated that participants in the pure-group condition reported liking the at-risk students in their intervention groups

⁴Analyses were also undertaken comparing the results for years one and two of data collection. Very few analyses yielded significant effects due to interactions with year of study, so this factor is not considered further.

Table I. Differences Between Prosocial and At-Risk Samples

| Dependent variable | Prosocial participants | | At-risk participants | | <i>t</i> | <i>df</i> | Effect size (<i>r</i>) |
|---|------------------------|-----------|----------------------|-----------|----------|-----------|-----------------------------|
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | | | |
| <i>Psychosocial functioning at Time 1</i> | | | | | | | |
| Self-report measures | | | | | | | |
| Externalizing (CBCL) | 11.45 | 5.88 | 17.86 | 7.52 | 5.52*** | 136 | .45 |
| Social skills (SSRS) | 22.62 | 4.14 | 20.33 | 4.69 | 3.05*** | 136 | .25 |
| Delinquency attitudes | 57.47 | 3.01 | 55.11 | 6.71 | 2.60** | 136 | .27 |
| Drug/alcohol attitudes | 22.61 | 2.35 | 21.30 | 3.37 | 2.61** | 136 | .24 |
| Parent-report measures | | | | | | | |
| Externalizing (CBCL) | 10.49 | 6.98 | 23.00 | 9.62 | 8.67*** | 137 | .63 |
| Social skills (SSRS) | 20.09 | 4.90 | 16.78 | 5.97 | 3.59*** | 137 | .29 |
| Teacher-report measures | | | | | | | |
| Externalizing (CBCL) | 0.38 | 0.79 | 10.45 | 3.64 | 21.85*** | 137 | .93 |
| Internalizing (CBCL) | 0.11 | 0.39 | 2.45 | 1.82 | 10.15*** | 137 | .77 |
| In-session behavior | | | | | | | |
| Percentage of sessions attended | 63.35 | 30.33 | 56.99 | 30.24 | 1.22 | 133 | .11 |
| Positive in-session behavior rate | 8.48 | 3.31 | 10.20 | 4.06 | 2.67** | 130 | .23 |
| Participating in group activities | 16.03 | 6.30 | 19.30 | 7.72 | 2.68** | 130 | .23 |
| Complimenting others | .93 | .87 | 1.09 | 0.92 | 1.00 | 130 | .09 |
| Negative in-session behavior rate | 0.19 | 0.39 | 0.46 | 0.85 | 2.27** | 130 | .24 |
| Failure to follow directions | 0.10 | 0.35 | 0.13 | 0.30 | 0.45 | 130 | .04 |
| Disrespecting others | 0.19 | 0.41 | 0.38 | 0.75 | 1.73* | 130 | .15 |
| Disrupting group activities | 0.29 | 0.49 | 0.88 | 1.84 | 2.44** | 130 | .29 |
| Peer reinforcement for | | | | | | | |
| Appropriate behavior | 2.05 | 0.61 | 2.04 | 0.57 | 0.09 | 130 | .01 |
| Inappropriate behavior | 1.24 | 0.31 | 1.37 | 0.43 | 2.06** | 130 | .18 |
| School record data | | | | | | | |
| Overall grade (%) | 88.24 | 5.67 | 72.48 | 11.60 | 9.74*** | 132 | .73 |
| No. of behavior infractions | 0.29 | 0.59 | 4.10 | 4.69 | 6.34*** | 132 | .63 |
| No. of times tardy to class | 1.38 | 1.69 | 2.92 | 3.87 | 2.91*** | 132 | .31 |
| No. of unexcused absences | 0.44 | 0.82 | 1.45 | 2.08 | 3.58*** | 132 | .38 |

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

better than did the at-risk participants in the mixed-group condition, $F(1, 45) = 2.94$, $p = .09$, $r = .25$.

A similar set of ANCOVAs examining effects at Time 3 yielded no significant results, indicating no long-term differences between the mixed- and pure-group conditions on these variables (all $F_s < 2.25$, $p_s > .10$). (See Table II for mean scores and standard deviations for all outcome measures.)

Tests of Hypothesis #2

The second study hypothesis was that intervention process variables would be more adaptive in the mixed-group condition than in the pure-group condition. To test this hypothesis, t -tests were conducted, comparing the pure- and the mixed-group conditions on the in-session behavior variables. All significant findings ran counter to the hypothesis (see Table III). Participants assigned to the pure-group condition had *higher* rates of positive in-session behavior than participants assigned to the mixed-

group condition. Comparison of the component variables (participating and complimenting) shows that pure-group participants had higher average rates of in-session participation in discussion and activities, and higher rates of complimenting others during sessions. There was a marginally significant higher rate of not following directions by at-risk children in the mixed-group condition compared to the pure-group condition. Pure groups were also characterized by lower scores on group leader and observer ratings of the extent to which the participant gave and received peer reinforcement for inappropriate behavior.

Tests of Hypothesis #3

Hypothesis #3 stated that intervention process variables would mediate the effect of group composition on intervention outcome variables. To test this hypothesis, we employed the distribution of products method recommended by MacKinnon, Lockwood, Hoffman, West,

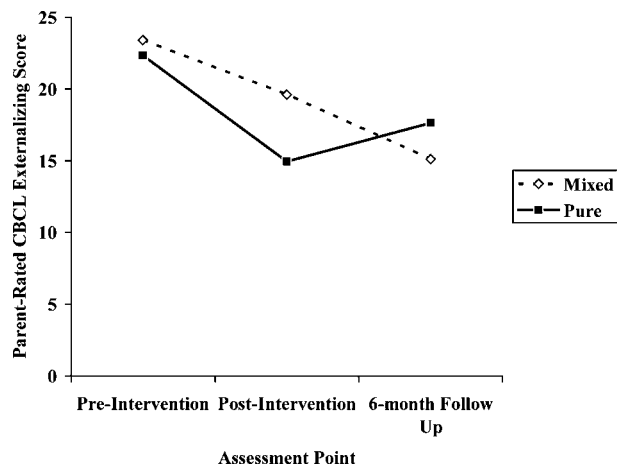


Fig. 1. Effect of group composition on parent ratings of children's externalizing behavior problems.

and Sheets (2002). In their article, MacKinnon and his colleagues ran a series of Monte Carlo simulations comparing 14 methods of testing mediation, including the traditional Baron and Kenny (1986) approach. They concluded that the Baron and Kenny approach had very low power and that the distribution of products method was one of two methods that yielded the most accurate Type I error rates and greatest power. The distribution of products method involves conducting a series of regression models involving the two process variables and the two outcome variables that were shown to be affected by group composition in Hypotheses 1 and 2. The process variables (potential mediators) were positive in-session behavior and peer reinforcement for inappropriate behavior. The outcome

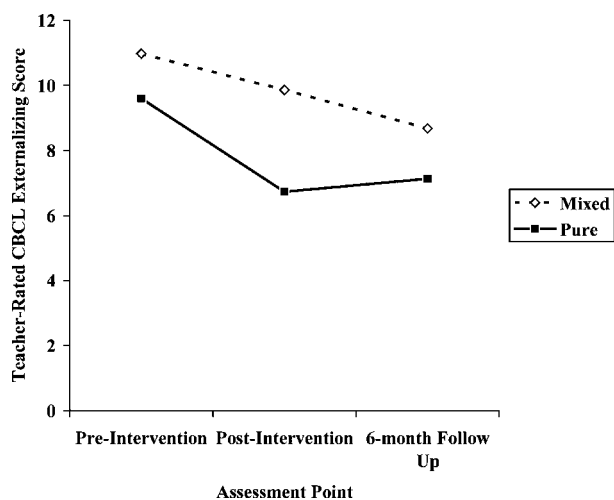


Fig. 2. Effect of group composition on teacher ratings of children's externalizing behavior problems.

variables were parent- and teacher-rated externalizing behavior. First, the mediators were regressed on the independent variable of group composition. Then the dependent variables were regressed on the independent variable and mediator variable, after controlling for Time 1 scores. The Z-score for the relation between group composition and the mediator variable was calculated by dividing the unstandardized regression coefficient by its standard error. Then the Z-score for the relation between the mediator and the dependent variable (when group composition is included as a predictor) was calculated. The two Z-scores were then multiplied and the product (P) was compared to a critical value of 2.18, corresponding to a $p < .05$.

These analyses indicated that for parent-rated externalizing behavior, positive in-session behavior was not a significant mediator ($P = .59$, $p > .05$), but peer reinforcement for inappropriate behavior was a mediator ($P = 3.70$, $p < .05$). For teacher-rated externalizing behavior, positive in-session behavior was a significant mediator ($P = 10.84$, $p < .05$), as was peer reinforcement for inappropriate behavior ($P = 5.82$, $p < .05$). The results of the regression analyses can be found in Tables IV and V.

Auxiliary Analyses

Because the results of the study were contrary to the hypotheses, we conducted further analyses that might help explain the findings. A t -test was used to determine if there was any difference in the intervention group attendance rates of participants in the pure- versus the mixed-group condition. Adolescents in the mixed-group condition attended a mean of 56.7% of sessions; those in the pure-group condition attended a mean of 57.5% of sessions ($t < 1$). Another t -test showed that there was no difference in the mixed- and pure- group conditions with regard to how much of the curriculum they were exposed to in their intervention groups ($t < 1$).

Because group size was confounded with pure/mixed composition, one alternative explanation for our results might be that pure groups, being smaller, produced greater group cohesiveness, thereby accounting for the more positive outcomes. To address this possibility, we created a new variable that averaged participants' liking for each of the other members of the group at the end of intervention. Analyses comparing the liking ratings of the at-risk participants in the pure and mixed groups show no significant difference, $t(52) = 1.35$, $p = .18$, suggesting that group cohesiveness as perceived by the at-risk participants in both conditions was comparable. As a further check, we correlated group size with all of the dependent variables and process variables within each experimental condition.

Table II. Means and Standard Deviations of Pure-Group and Mixed-Group Conditions on Psychosocial Functioning Measures at Pre-, Post-, and Follow-Up Assessments

| Dependent variable | Pure | | | Mixed | | |
|---------------------------------------|--------------|--------------|---------------|--------------|---------------|---------------|
| | Pre | Post | Follow-up | Pre | Post | Follow-up |
| Self-report measures | | | | | | |
| Externalizing (CBCL) | 16.16 (8.14) | 17.33 (8.70) | 16.14 (6.69) | 18.95 (6.98) | 17.64 (7.72) | 16.31 (6.60) |
| Social skills (SSRS) | 20.40 (4.80) | 20.33 (4.94) | 21.43 (2.34) | 20.28 (4.66) | 20.33 (3.83) | 20.46 (3.80) |
| Delinquency attitudes | 55.72 (4.99) | 52.86 (9.17) | 42.86 (15.73) | 54.72 (7.66) | 53.11 (11.40) | 40.04 (16.67) |
| Drug/alcohol attitudes | 21.76 (2.89) | 21.33 (3.28) | 21.50 (3.59) | 21.00 (3.66) | 21.33 (3.21) | 21.29 (3.41) |
| Know at-risk group members | 2.11 (0.70) | 2.64 (0.70) | 2.46 (0.63) | 1.94 (0.95) | 2.48 (0.95) | 2.43 (1.14) |
| Spend time with at-risk group members | 1.72 (0.54) | 1.97 (0.74) | 1.77 (0.44) | 1.71 (0.85) | 2.05 (0.85) | 1.91 (1.08) |
| Like the other at-risk group members | 2.00 (0.78) | 2.78 (0.77) | 2.51 (0.64) | 2.01 (0.99) | 2.58 (0.87) | 2.13 (1.02) |
| Parent-report measures | | | | | | |
| Externalizing (CBCL) | 22.36 (9.80) | 14.95 (8.41) | 17.64 (12.51) | 23.40 (9.60) | 19.61 (8.41) | 15.12 (7.78) |
| Social skills (SSRS) | 17.64 (6.51) | 19.14 (4.62) | 18.43 (5.10) | 16.25 (5.63) | 17.64 (5.25) | 18.19 (4.88) |
| Teacher-report measures | | | | | | |
| Externalizing (CBCL) | 9.60 (3.63) | 6.73 (3.40) | 7.13 (3.69) | 10.98 (3.59) | 9.86 (3.97) | 8.68 (4.83) |
| Internalizing (CBCL) | 2.36 (1.66) | 1.68 (1.32) | 1.88 (1.54) | 2.50 (1.94) | 2.16 (1.89) | 2.77 (2.08) |

Of the 56 correlations, only three were significant, and among these three there was no meaningful pattern. This would suggest that group size is not accounting for the treatment group differences in outcome measures or process variables.

DISCUSSION

Dishion et al.’s (1999) *American Psychologist* article put a prominent voice to the commonly held notion that group treatment for acting-out adolescents may be more harmful than helpful because it provides a forum for the youth to encourage and strengthen each other’s anti-social tendencies. This “deviancy training/iatrogenic effects” hypothesis has intuitive appeal and is quite popular among researchers and clinicians alike. However, strong empirical support for this hypothesis is lacking.

Our study was one of the first to compare directly the efficacy of group interventions in pure and mixed groups. To our surprise, study results show that *more* aggregation of high-risk adolescents in intervention groups (i.e., the pure-group condition) led to better outcomes than less aggregation (i.e., the mixed-group condition). After controlling for baseline scores, participants in the pure-group condition had lower teacher and parent ratings of externalizing behavior on the CBCL at postintervention than the at-risk members of the mixed-group condition. The mixed-group condition did result in a decrease in parent-rated externalizing behavior from pre- to postintervention; however, the decrease was greater for the pure-group condition. For teacher-rated externalizing behavior, there was no change from pre- to postintervention for at-risk participants in the mixed-group condition, whereas scores did decline significantly for participants in the pure-group condition.

Table III. Means and Standard Deviations of Intervention Process Variables for Mixed- Versus Pure-Group Condition

| Dependent variable | Pure | Mixed | t | df | Effect size |
|---|--------------|--------------|---------|----|-------------|
| | | | | | (r) |
| Positive in-session behavior rate | 13.15 (3.56) | 8.41 (3.23) | 5.35*** | 59 | .57 |
| Participating in group activities | 24.89 (6.70) | 15.92 (6.23) | 5.30*** | 59 | .57 |
| Complimenting others | 1.41 (0.90) | 0.89 (0.89) | 2.24** | 59 | .28 |
| Negative in-session behavior rate | 0.34 (0.58) | 0.53 (0.98) | 0.83 | 59 | .11 |
| Failure to follow directions | 0.05 (0.17) | 0.17 (0.35) | 1.80* | 59 | .23 |
| Disrespecting others | 0.21 (.44) | 0.47 (0.88) | 1.53 | 59 | .20 |
| Disrupting group activities | 0.77 (1.66) | 0.95 (1.96) | 0.37 | 59 | .05 |
| Peer reinforcement for appropriate behavior | 2.14 (0.51) | 1.97 (0.61) | 1.15 | 59 | .15 |
| Peer reinforcement for inappropriate behavior | 1.23 (0.23) | 1.46 (0.50) | 2.42** | 59 | .31 |

*p < .10. **p < .05. ***p < .01.

Table IV. Hierarchical Multiple Regression Analysis for Mediation Tests of the Relation Between Group Composition and Parent-Rated Externalizing Behavior

| Variable | <i>B</i> | <i>SE B</i> | Standardized <i>B</i> | <i>p</i> |
|---|----------|-------------|-----------------------|----------|
| Block 1 | | | | |
| Parent externalizing at Time 1 | 0.55 | 0.10 | .61 | <.001 |
| Block 2 | | | | |
| Parent externalizing at Time 1 | 0.55 | 0.09 | .61 | <.001 |
| Group composition | -4.86 | 1.82 | -.27 | <.01 |
| Block 3 (with positive in-session behavior as mediator) | | | | |
| Parent externalizing at Time 1 | 0.55 | 0.09 | .62 | <.001 |
| Group composition | -3.90 | 2.33 | -.22 | <.05 |
| Positive in-session behavior | -0.02 | 0.28 | -.02 | .91 |
| Block 3 (with peer reinforcement of inappropriate behavior as mediator) | | | | |
| Parent externalizing at Time 1 | 0.55 | 0.09 | .62 | <.001 |
| Group composition | -3.99 | 1.85 | -.22 | <.05 |
| Peer reinforcement of inappropriate behavior | 3.59 | 2.04 | .18 | <.10 |

Also contrary to expectations was the finding that participants' *in-session behavior* was more positive in the pure-group condition than in the mixed-group condition. At-risk adolescents in pure groups participated in group activities and discussions more, and they complimented one another more than the at-risk members in mixed groups. Essentially, they were more positive role-models for one another when there was more at-risk peer aggregation than when there was less. They also engaged in less deviancy training (reinforcement of deviant peer behavior) than did the at-risk adolescents in the mixed-group condition. Furthermore, analyses showed that these group process variables served to mediate the effect of intervention group composition on teacher- and parent-rated conduct problems.

Because these findings contradict the widely advocated view that aggregating at-risk adolescents is con-

traindicated, our first step was to consider and rule out problems with the study methodology that might have caused the results. Analyses indicated that neither differences in attendance at the treatment sessions nor differential exposure to the curriculum could account for the obtained group differences. Unfortunately, intervention group size was confounded with group composition. Difficulties with participant recruitment, assignment to condition, and scheduling led to the mixed groups being somewhat larger ($M = 9.5$ children assigned) than the pure groups ($M = 6.3$ children assigned). Although our analyses of the 'liking' variable indicated that group size was not confounded with group cohesiveness, we cannot rule out other possible confounds that may be associated with group size, such as less supervision and fewer opportunities to participate in group discussions in larger groups.

Table V. Hierarchical Multiple Regression Analysis for Mediation Tests of the Relation Between Group Composition and Teacher-Rated Externalizing Behavior

| Variable | <i>B</i> | <i>SE B</i> | Standardized <i>B</i> | <i>p</i> |
|--|----------|-------------|-----------------------|----------|
| Block 1 | | | | |
| Teacher externalizing at Time 1 | 0.56 | 0.13 | .49 | <.001 |
| Block 2 | | | | |
| Teacher externalizing at Time 1 | 0.50 | 0.13 | .44 | <.001 |
| Group composition | -2.58 | 0.92 | -.31 | <.01 |
| Block 3 (with positive in-session behavior as mediator) | | | | |
| Teacher externalizing at Time 1 | 0.51 | 0.12 | .45 | <.001 |
| Group composition | -1.29 | 1.10 | -.16 | .24 |
| Positive in-session behavior | -0.26 | 0.13 | -.27 | <.05 |
| Block 3 (with peer reinforcement for inappropriate behavior as mediator) | | | | |
| Teacher externalizing at Time 1 | 0.44 | 0.12 | .38 | <.001 |
| Group composition | -1.94 | 0.90 | -.23 | <.05 |
| Peer reinforcement for inappropriate behavior | 2.90 | 1.05 | .30 | <.01 |

The next step in understanding the results of this study is to compare them more closely with previous research findings. There is only one previous study, the St. Louis Experiment (Feldman et al., 1983), that set out to address the same basic research question as the present study: comparing treatment processes and outcomes of mixed versus pure treatment groups for adolescent antisocial behavior problems. Feldman et al. reported slightly better results for the mixed-group condition than the pure-group condition. There are many differences between the methodologies used in these two studies, however (e.g., screening and identification of participants, participant age, group size, intervention modality, ratio of prosocial to at-risk youth in mixed groups, and analytic strategies); any one of these differences could be responsible for the somewhat opposite effects found. Furthermore, because the methodology used in the present study reflects more current treatment strategies, it seems likely that it is a more valid test of the effect of deviant peer aggregation in treatment groups as currently implemented.

The Disengagement/Discrimination Hypothesis for Heterogeneous Groups

The current study does lend support for Dishion et al.'s (1999) deviancy training hypothesis. Peer reinforcement of inappropriate behavior during intervention sessions was observed, and it seemed to partially cause attenuated treatment effects. However, the findings diverge from Dishion et al.'s (1999) hypotheses when it comes to linking deviancy training with greater at-risk peer aggregation in treatment groups. That is, Dishion et al. (1999) argued that deviancy training is more likely to occur in settings involving greater aggregation of at-risk adolescents. In contrast, the current study results indicate instead that more deviancy training occurred in the mixed-group condition than in the pure-group condition.

Why should this be the case? Perhaps the mixed-group model tends to lead to disengagement for at-risk members when they compose a minority of the group and when the majority consists of especially high-functioning adolescents. It may be that group leaders tended to focus on the prosocial group members, who were probably easier and more enjoyable to teach. Thus, it might be natural for group leaders to call on the prosocial participants more often, provide more positive reinforcement to them, and provide more opportunities for them to receive positive peer reinforcement. The at-risk adolescents' lack of confidence and/or real lack of ability to compete in this manner with their prosocial counterparts may have led them to withdraw from the group and to form instead a minority out-group. They may have become more de-

fensive, causing a general disengagement from the group and a desire to get reinforcement where they could—from the other deviant adolescents in the group. Unfortunately, we do not have the individual-level data to test this possibility. Future research can examine the factors that may contribute to the development of deviancy training. These may include group composition, group size, and the level of engagement of group members, to name some of the possible factors that are present in the current study.

Implications

The current study supports the continued use of the standard group treatment model for young adolescents with conduct problems: a social cognitive skills-training intervention, conducted in groups that consist entirely of at-risk members. As such, our findings are consistent with reviews of the literature documenting the effectiveness of this treatment approach for children and adolescents with conduct disorder (Brestan & Eyberg, 1998). Practitioners should be aware of the deviancy training phenomenon and the fact that it can occur in intervention group settings. Practitioners can likely minimize the risk of deviancy training by closely supervising group members at all times and by using an effective behavior management system for group sessions.

Study Limitations

This study was helpful in clarifying some issues and in raising new ones with regard to the role of group composition in treating aggressive adolescents. However, two important limitations remain. First, the study had a small sample and only compared two experimental conditions. Future studies with larger samples are needed to determine if our findings can be replicated. Second, our follow up was only 6 months in duration, and some research has suggested that the iatrogenic effect may not show up until several years later (Dishion & Andrews, 1995; Poulin et al., 2001).

Closing Thoughts

Effective intervention with at-risk youth is a complex task with many issues that remain unresolved. The main contribution of this study is to provide evidence for the provocative suggestion that aggregating deviant adolescents in treatment groups may actually be *more* effective than conducting treatment groups consisting mostly of prosocial models. This finding is surprising and significant

because it runs counter to the prevailing opinion among researchers and clinicians. Although psychosocial interventions may well carry some degree of risk for harmful effects and everything possible should be done to minimize the risk, aggregation of at-risk adolescents in group treatment is not necessarily a condition that triggers negative effects.

ACKNOWLEDGMENTS

This research was supported by funding from the National Institute of Mental Health (NRSA Grant # 1 F31 MH63546-01), the Society for the Psychological Study of Social Issues, Division 53 of APA, the University of Kentucky Center on Drug and Alcohol Research, the University of Kentucky Department of Psychology, the University of Kentucky Research Foundation, and the Kentucky Psychological Association. We thank the following people whose assistance made this study possible: Drs Gerald Hutchins, Bobbie Burcham, Kim Hooks, Pam Wilcox, Josh Miller, Kate Flory, and Leah Horvath; Beverly Hollinshed, T.C. Johnson, Steve Carmichael, Robin Potter, and Jennifer Gardner.

REFERENCES

- Achenbach, T. M. (1991a). *Manual for the child behavior checklist/4-18 and 1991 profile*. Burlington: Department of Psychiatry, University of Vermont.
- Achenbach, T. M. (1991b). *Manual for the teacher's report form and 1991 profile*. Burlington: Department of Psychiatry, University of Vermont.
- Achenbach, T. M., & Edelbrock, C. (1991). *Manual for the youth self-report and profile*. Burlington: Department of Psychiatry, University of Vermont.
- Arnold, M. E., & Hughes, J. N. (1999). First do no harm: Adverse effects of grouping deviant youth for skills training. *Journal of School Psychology, 37*, 99-115.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.
- Brestan, E. V., & Eyberg, S. M. (1998). Effective psychosocial treatments of conduct-disordered children and adolescents: 29 years, 82 studies, and 5,272 kids. *Journal of Clinical Child Psychology, 27*, 180-189.
- Caplan, M., Weissberg, R. P., Grober, J. S., Sivo, P. J., Grady, K., & Jacoby, C. (1992). Social competence promotion with inner-city and suburban young adolescents: Effects on social adjustment and alcohol use. *Journal of Consulting and Clinical Psychology, 60*, 56-63.
- Dishion, T. J., & Andrews, D. W. (1995). Preventing escalation in problem behaviors with high-risk young adolescents: Immediate and 1-year outcomes. *Journal of Consulting and Clinical Psychology, 63*, 538-548.
- Dishion, T. J., Capaldi, D. M., Spracklen, K. M., & Li, F. (1995). Peer ecology of male adolescent drug use. *Development and Psychopathology, 7*, 803-824.
- Dishion, T. J., Eddy, J. M., Haas, E., Li, F., & Spracklen, K. M. (1997). Friendships and violent behavior during adolescence. *Social Development, 6*, 207-223.
- Dishion, T. J., French, D. C., & Patterson, G. R. (1995). The development and ecology of antisocial behavior. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Vol. 2. Risk, disorder, and adaptation* (pp. 421-471). New York: Wiley.
- Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm—Peer groups and problem behavior. *American Psychologist, 54*, 755-764.
- Dishion, T. J., Spracklen, K. M., Andrews, D. W., & Patterson, G. R. (1996). Deviancy training in male adolescent friendships. *Behavior Therapy, 27*, 373-390.
- Farrington, D. P., & Loeber, R. (1996). Self-reported delinquency and a combined delinquency seriousness scale based on boys, mothers, and teachers. *Criminology, 34*, 493-518.
- Feldman, R. A., Caplinger, T. E., & Wodarski, J. S. (1983). *The St. Louis conundrum: The effective treatment of antisocial youths*. Englewood Cliffs, NJ: Prentice-Hall.
- Gresham, R. M., & Elliott, S. N. (1990). *Social skills rating system*. Circle Pines, MN: American Guidance Service.
- Handwerk, M. (2000). The iatrogenic effects of group intervention for antisocial youth: Premature extrapolations? *Journal of Behavioral Education, 10*, 223-238.
- Hoyle, R. H., Donohew, R. L., Lorch, E. P., & Palmgreen, P. (1997, May). Sensation-seeking and drug-related attitudes and behaviors among adolescents. Paper presented at the 5th Annual Meeting of the Society for Prevention Research, Baltimore, MD.
- Kamphaus, R. W., & Frick, P. J. (1996). *Clinical assessment of child and adolescent personality and behavior*. Boston: Allyn and Bacon.
- Landau, S., Milich, R., & Diener, M. B. (1998). Peer relations of children with attention-deficit hyperactivity disorder. *Reading and Writing Quarterly: Overcoming Learning Difficulties, 14*, 83-105.
- MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*, 83-104.
- McMahon, R. J., & Wells, K. C. (1998). Conduct problems. In E. J. Mash & R. A. Barkley (Eds.), *Treatment of childhood disorders* (pp. 111-210). New York: Guilford.
- Poulin, F., Dishion, T. J., & Burraston, B. (2001). Three-year iatrogenic effects associated with aggregating high-risk adolescents in cognitive-behavioral preventive interventions. *Applied Developmental Science, 5*, 214-224.
- Weaver, B. (2000, July/August). *APS Observer*. (Available from the American Psychological Society, 1010 Vermont Avenue, NW, Suite 1100, Washington, DC 20005-4907).
- Weissberg, R. P., Caplan, M., Benetto, L., & Jackson, A. S. (1990). *The social competence promotion program for young adolescents: Social problem-solving module*. (Available from CASEL, <http://www.casel.org>)