

## Diffusion of treatment interventions: exploration of ‘secondary’ treatment diffusion

Knut Gundersen<sup>ab\*</sup> and Frode Svartdal<sup>ab</sup>

<sup>a</sup>*Diakonhjemmet Høgskole Rogaland, ART-department, Vaagsgt. 40, 4306 Sandnes, Sandnes, Norway;* <sup>b</sup>*University of Tromsø, Norway*

*(Received 7 October 2008; final version received 5 November 2008)*

The present report presents outcome results from a randomized controlled effect study on extended *Aggression Replacement Training* (ART). In a pre–post design, a 30-hour ART intervention was found to significantly reduce behavioral problems and increase social skills. The control group did not demonstrate comparable changes, but still indicated improvement. Such control-group improvement may be caused by improper treatment and control group implementation (diffusion of treatment) and/or ‘secondary’ diffusion caused by participants in the treatment group affecting control group subjects by demonstrating changed behavior. Both mechanisms were explored, and it is concluded that the improvement observed in the control group was due to such ‘secondary diffusion’. Implications of these findings are discussed.

**Keywords:** Aggression Replacement Training; peer intervention; social competence; social skills; anger control

### Introduction

When interventions are implemented in randomized control group studies, interventions intended for participants in the treatment group sometimes also affect participants in the control group. In a research context this affects the internal validity of the study, as conclusions about treatment efficacy are threatened. This problem, often named ‘diffusion of treatment’, is an experimenter-related issue because it is an effect of the improper implementation of the intervention (e.g. Kazdin, 1998).

Diffusion of treatment implies that improvement caused by the intervention in the treatment condition is also observed in the control condition because elements of the treatment are found in both conditions. For example, if a teacher participating in an intervention program adopts elements of the program in classes not participating in the program, control group subjects may be inadvertently affected by the intervention. However, treatment ‘effects’ may be observed in control groups through other mechanisms as well. For example, if participants in the experimental group demonstrate extensive improvement and thus act as role models for participants in the control group, it is possible that control participants may change their behavior accordingly. Under such circumstances, interventions may have been

---

\*Corresponding author. Email: knut.gundersen@diakonhjemmeths.no

implemented with no diffusion between conditions, but ‘diffusion’ at a secondary level may still occur in that participants in the treatment group demonstrate changed behavior that in turn affect control group participants. This effect may be pronounced in intervention programs carried out in settings like schools and institutions with extensive interaction between participants, and especially so if the intervention has a relatively long duration (Conrad, Conrad, & Walcott-McQuigg, 1991).

The present paper explores these issues in our efforts to evaluate intervention effects of a program based on *Aggression Replacement Training* (ART), a program for training of social skills and aggression management. In a previous study (Gundersen & Svartdal, 2006), groups of youth participants were randomly allocated to an ART intervention or to a control condition. Standard instruments measured social skills and behavior problems before and after treatment. Participants in the ART group demonstrated improvement on 9 of 10 measures, whereas participants in the control group improved on 2 of 10 measures and also tended to improve on three additional measures. This clearly indicated an effect of the ART intervention.

In the context of the present investigation, the finding that control group participants tended to demonstrate improved scores on some measures is of particular interest. This ‘effect’ could in principle be attributed to a number of variables. For example, as at least some of the control group participants received other interventions between measurements, this may have caused a positive change. Other potential explanations include regression to mean effects, and a test–retest effect. We believe that these variables most probably can be excluded as explanations of the changes observed in the control group in the Gundersen and Svartdal study. First, the members of the control group received three sessions per week on academic subjects (e.g. math, Norwegian) instead of the ART training. It is unlikely that this caused significant changes in deviant behavior. Second, although we do not have data on alternative interventions, random allocation to ART and control group implies that such interventions were equally distributed in the two groups (the ART intervention did not exclude other interventions). Third, only five of the control group participants were classified at problem level 3 (i.e. the most behavioral problems; Gundersen & Svartdal, 2006, p. 71). For this reason we also consider the regression to mean effect as unlikely in explaining the observed changes. Finally, the possibility that the test–retest procedure could generate positive changes is not supported by available data. For example, a Norwegian study applying Social Skills Rating System (SSRS) in a pre–post comparison did not observe a change in control group participants (Manger, Eikeland, & Asbjørnsen, 2002).

We believe, therefore, that it is more likely that the positive change observed in the control group of the Gundersen and Svartdal (2006) study could be attributed to diffusion of treatment effects, primary or secondary. A majority of the participants in that study (63%) were recruited from schools, and control group participants were almost exclusively from schools (83%). Moreover, as a majority of the school participants was recruited from the same school or even the same class and the intervention was distributed over 13 weeks (Gundersen & Svartdal, 2006, p. 68), it is likely that diffusion of treatment effects did play a role in creating positive change in the control group. Unfortunately, Gundersen and Svartdal (2006) did not collect data that could illuminate these issues. The investigation reported in the present paper attempted to do so. Participants were randomly allocated to treatment (ART)

or control conditions. Behavior problems and social skills were measured twice, first at the outset of the investigation and then following the ART intervention. In addition a number of variables were included that attempted to measure treatment fidelity and the context in which treatment was administered. Given that the intervention of this second study in fact creates reliable positive changes in participants in the treatment group, we then ask whether these changes are also observed in the control group, and if so, if it is possible to trace these control group 'effects' to specific mechanisms.

### ***ART (Aggression Replacement Training)***

ART is a social competence intervention method consisting of anger control, prosocial skills, and moral reasoning training. A specific intervention typically consists of 30 sessions, 10 on each component. The components are practiced every week during the intervention period. The training, which is led by two coaches, is conducted in groups of four to eight participants. Each session has a firm structure, and in addition to the specific components mentioned there is also room for games, recognition, home work, etc.

#### *Social skills training*

The core of the social skills training component is 40 social skills for preschool (McGinnis & Goldstein, 1990), 60 skills for primary school (McGinnis & Goldstein, 1984), and 50 skills for junior-high/high school and adults (Goldstein, Glick, & Gibbs, 1998). Each session consists of training one specific skill, the skill being described in four or five steps. The skills are adopted through a firm structure (i.e. definition, demonstration, discussing its function, selecting a main role player and a co-player, planning and performing the role play, and obtaining feedback from participants and trainers). Examples of skills are *listening*, *giving a compliment*, *dealing with group pressure*, and *negotiating*.

#### *Moral reasoning training*

Moral reasoning training comprises the value component of ART. This part of the program was primarily developed by John Gibbs. Based on Kohlberg's (1984) six moral stages, Gibbs proposed four stages of moral development, from egocentric and immature thinking to more empathetic and value-based thinking. Examples of immature thinking include *blaming others*, *assuming the worst*, and *minimizing*. Such thinking is, according to Gibbs, cognitive distortions that may be changed. In dilemma discussions, in which mature viewpoints are reinforced and cognitive distortions are exposed and discussed, a more empathetic understanding of reality is developed.

#### *Anger control training*

The Anger Control program was initially developed by Eva Feindler (Feindler, 1995). The program focuses on physiological responses, cognitive processes, and behavioral responses. Focus on physiological responses help participants identify

external anger triggers as well as own anger cues, and specific techniques help participants master anger reactions. The cognitive component emphasizes the typical thought patterns found in persons with aggressive and impulsive actions. Through cognitive restructuring strategies, participants are helped to identify irrational thought patterns and replace them with a more rational situation analysis. Participants are encouraged to develop alternative thought patterns and/or self-instructions that help reduce the conflict and create 'mental distance' to the anger triggers (Feindler, 1995, p. 33). The behavioral component involves establishing new prosocial actions that can replace previous performance patterns. Note that anger control not only focuses on anger control *per se*, but aims at providing more functional behavioral alternatives. Thus, self-assertion techniques and communication skills are important parts of the program. When conducting the program, students were allowed to include the concept of character education into the moral reasoning program (Salmon, 2004), and the concept of setting events into the anger control training (Gundersen, Olsen, & Finne, 2006). Also, rehearsals on selected pedagogical techniques described by Gundersen and Moynahan (2006) were included. The program can thus be seen as a form of extended ART.

### *Diffusion of treatment interventions*

In the present context we distinguish between two variants of the problem of diffusion of treatment interventions.

(a) Elements of the manipulation can be found in the control condition. As an example, Kazdin (1998) discusses the study by Austin and colleagues (Austin, Lieberman, King, & DeRisi, 1976), who compared behavioral oriented social skill training program and an eclectic-milieu therapy approach applied to psychiatric patients in a day hospital program. The analyses showed a slight, but not significant, advantage to the behavioral program. By the end of the study, it was discovered that one of the therapists in the eclectic-milieu therapy approach was participating in behavioral training and used these techniques in his therapy. When reanalyzing the data, excluding the contribution from this therapist, the behavioral program appeared to be significantly better than the other program. Thus, because the treatment provided in the two conditions was not as distinct as planned, a conclusion about a possible treatment effect was threatened.

Such diffusion of treatment effects can take many forms (see Kazdin, 1998, p. 24, for discussion), all of which represent threats to the internal validity of the investigation. As discussed by Kazdin (1998), diffusion of treatment may be traced to the improper implementation and non-implementation of the intervention in the treatment and control groups, respectively. In the context of intervention evaluation, improper implementation of the treatment may cause a reduced effect because the control group may change in a positive direction. However, the reverse is also possible: the regime of the no treatment condition may spread into the treatment condition (Kazdin, 2003, p. 33).

The quality of program implementation is therefore of great importance for evaluation of intervention effects. Thus, meta-analyses on the effectiveness of school-based programs for reducing aggressive behavior demonstrate a relation between inadequately implemented programs and lack of effect (Wilson & Lipsey, 2007; Payne, Gottfredson, & Gottfredson, 2006). Further, programs that are initiated by

researchers and programs that are implemented as part of a demonstration project (where great attention is given to program fidelity) are much more successful compared to programs implemented under routine conditions (Wilson, Lipsey, & Derzon, 2003). Gottfredson and Gottfredson (2002) identified several school and program characteristics related to high quality implementation, including program standardization (clear and explicit materials and procedures), organizational capacity (staff support, image to community), organizational support (program providers, training quality, time for preparation of sessions), and integration into normal school operation (how enthusiastic and widespread the program is adopted). In their survey of 3691 school-based prevention programs, these authors found that typical implementation of prevention programs are generally not well integrated into normal school operations and that the school staff implementing these programs is in need of better training, support, and supervision. In the context of ART, these issues have been given much attention (Gundersen, 2005).

(b) Diffusion of treatment may also occur even if treatment procedures are properly implemented in the program, as when participants in intervention and control groups interact during the course of the study. If participants in the treatment group demonstrate improvement and act as role models for participants in the control group, it is possible that control participants may change their behavior accordingly. Here, interventions may have been properly implemented, but diffusion at a secondary level may still occur and thus represent threats to the internal validity (e.g. Madden, Bucholz, Todorov, & Grant, 2002). This issue may be particularly important in Norway, as pupils in the same age group will belong to, and thus interact with, a large group defined by age rather than by membership in specific school classes. Subgroups of pupils with behavior problems may selectively interact (Bagwell, 1999; Estell, 2007), and deviant behavior may be developed and maintained by reinforcement from others in the subgroup. This may explain how behavior problems develop, but the reverse is also possible: if pupils with behavior problems participating in the ART program demonstrate a positive change, that change may affect members of the control group through similar mechanisms and cause a 'secondary' diffusion of treatment effect.

There are a number of factors that increases the likelihood of 'secondary' diffusion. When randomizing a small group of pupils with deviant behavior as common trait into two groups, it is probable that there will be close friendships across members of the two groups (Estell, 2007) which will increase the tendency of mutual influence. Also when interventions extend over time like in school setting, participants in the intervention interact with members of the control group daily in class or recess and can be influenced by general changes in class milieu. Such 'secondary' diffusion may of course exist in conjunction with diffusion of treatment traceable to improper implementation of interventions. Thus, teachers responsible for implementing the ART training might also be involved in teaching control group participants and bring techniques from intervention into the control group classroom.

There is probably no practical way of separating the effects of these two forms of diffusion. In the present paper we nevertheless attempted to assess the potential effect of diffusion of both types in analyses of treatment effects observed in treatment and control groups receiving ART or no treatment. One approach would be to correlate treatment effects to environmental variables relevant to the intervention. For example, one would predict a high degree of diffusion between

treatment and control conditions if participants in both were recruited from samples with high interaction; if participants were recruited from samples with low interaction, less diffusion would be expected. Further, one might also expect that interventions demonstrating a high degree of positive change will affect control group participants positively, whereas a lower effect is expected in interventions demonstrating a low degree of positive change.

## Method

### *Participants*

Participants were 140 children and youths (mean age = 11.0 years; 56 girls) from elementary and intermediate schools, kindergartens and special schools in Norway. Table 1 shows the actual distribution of participants.

### *ART trainers*

The ART trainers were students enrolled in a 60 credits postgraduate study in 'Training of social competence' at Diakonhjemmet University College, Rogaland. All students had as basic 3 years of college education as teachers or authorized social educators, and all were employed full time at local schools and institutions. Fourteen of the trainers were educated as ART trainers before they started the postgraduate program, and they also had practice running ART groups. Prior to the ART intervention, all students had theoretical and practical training in ART and expanded ART corresponding to 30 credits education. This included also leading at least 18 sessions of ART.

### *Consent and approval by ethic committee*

Parents, teachers and youths involved in the study received extended written and verbal information about the investigation. The information was presented by the ART students in meetings with groups and/or individually. Participants signed their consent to participate. The protocol for the project was approved by regional ethics committee.

### *ART and control groups*

Potential participants were recruited from the ART trainers' workplaces (schools or institutions), and were initially assessed with a simple checklist (Gundersen, 2003)

Table 1. Participants and mean ages.

	Age (years)
Elementary schools ( $n = 90$ )	10.4
Institutions ( $n = 15$ )	14.9
Kindergartens ( $n = 12$ )	4.5
Intermediate schools ( $n = 13$ )	13.9
Special schools ( $n = 10$ )	14.6

to differentiate three general levels of behavior problems: (1) *None* (the youth may need to train prosocial skills, but has not developed significant behavior problems); (2) *Mild* (the youth displays some degree of behavior problems); (3) *Some* (the youth displays some degree of serious behavior problems). For each school or institution, two groups of participants were composed with a pre-defined distribution of the three levels of behavior problems. Then, one of these groups was randomly allocated to the intervention (ART) group and the other to the control condition. Overall, there were 12 student projects and 140 participants in total. The ART group consisted of 77 participants (mean age = 11.1 years, SD = 3.12) and the control group of 63 participants (mean age = 10.9 years, SD = 3.20). We calculated an index for social competence and problem behavior derived from SSRS scores at pre-test.<sup>1</sup> There was a marginal group difference between the ART and control groups on the social competence index (the respective means were 24.8 and 26.1,  $F(1,123) = 3.85$ ,  $p > 0.05$ ), but a nonsignificant difference on the problem behavior index (mean ART = 30.1, mean control = 27.9,  $F(1,124) = 2.08$ ,  $p > 0.15$ ). Thus, the groups were comparable prior to the intervention phase, but the ART group tended to display lower social skills scores and somewhat higher problem behavior scores.

### ***Intervention***

The ART group received a 30-hour ART extended intervention with 10 sessions from each of the components skills training, anger training, moral reasoning training (Goldstein et al., 1998). The very young participants received ART sessions adapted to their age levels. Overall, participants attended most of the training sessions (mean absence = 3.04 hours). The ART students received information and supervision before and during the implementation of the ART interventions with the purpose of achieving optimal treatment fidelity of the interventions. Instructors from the college visited and supervised at least once during the intervention.

Participants in the control groups did not receive any services beyond that provided by the various schools and institutions. The control group members received ordinary academic subjects in class during ART sessions.

### ***Instruments***

Two instruments were selected to assess behavioral problems and social skills, using multiple informants (i.e. youth, parent, teacher), CADBI and SSRS. In addition the students filled in a questionnaire with items that provided information about key variables of each project (Gundersen, Moynahan, & Svartdal, 2003).

#### ***CADBI***

*The Child and Adolescent Disruptive Behavior Inventory 2.3* (CADBI) has a parent and teacher version (Burns, Taylor, & Rusby, 2001a,b). Both versions contain assertions relating to various forms of behaviors and behavior problems towards adults (parents, teachers) and peers, activity level, and attentional skills (e.g. 'Starts physical fights with peers'; 'Does not seem to show concern for the feelings of others'). Each of the assertions is scored on an eight-point frequency of occurrence

scale (never in the past month, one or two times in the past month, three or four times in past month, two to six times per week, one time per day, two to five times per day, six to nine times per day, and 10 or more times per day). In part 6, parents and teachers rate the child's academic competence on 11 items. These items are rated on a seven-point scale (from severe difficulty to excellent performance). CADBI has achieved good internal consistency 3-month test-retest reliability (Fitzgerald, 2002; Iredale, 2000; Skansgaard & Burns, 1998; Skansgaard & Burns, 1998). In a previous investigation (Gundersen & Svartdal, 2006) it was found that CADBI and the Child Behaviour Check List (Achenbach & Rescola, 2001) correlated rather well.

### *SSRS*

Social Skills Rating System (SSRS) (Gresham & Elliot, 1990), Norwegian version (Ogden, 2003) is a standardized norm-referenced scale measuring children and adolescents social skills, problem behavior and academic competences. The SSRS was chosen because it has demonstrated validity in several studies (Demaray et al., 1995; Gresham & Elliot, 1990; Ogden, 2003) and its relevance to the ART program. Respondents (teacher, parent and youth) rate 30 statements which describe examples of social behavior and mark how often this behavior is demonstrated. Originally, ratings were done on a three-points scale, but Ogden (2003) increased this to four (1=never, 2=sometimes, 3=often, 4=very often). The reason for this change was that the expanded scale was found to better discriminate between the most socially skilled students (Ogden, 2003). The different items are combined in four subscales, *Cooperation*, *Assertion*, *Self-control*, and *Responsibility* (the subscale *Responsibility* is not included in the teacher scale).

### *Factors relevant for the implementation and outcome*

A questionnaire (Gundersen et al., 2003) was administered that attempted to quantify information related to the implementation of the ART intervention. Among the variables measured were age and motivational characteristics of the participants; contact with parents; contact with teachers; interaction between ART and control groups; absence from training; interest in, and support of, program from parents, teachers, and institution; and presence of an identifiable role model. The questionnaire was filled in by the student ART trainers. Each student was evaluated with a Community Juvenile Accountability Act (CJAA) checklist developed by Washington State Institute of Public Policy in connection with the investigation from Washington State's *Community Juvenile Accountability Acts Program* (see Barnoski & Aos, 2004).

### *Procedure*

The student ART trainers were divided into 16 sub-groups. The projects took place in normal schools (10), residential homes for children with behavior problems (three), special school (two), and kindergarten (one). One of the special school groups was excluded because of improper implementation, and in one of the school projects and two of the residential home projects it was not possible to establish a control group. Overall, 12 projects with 77 participants were included in the

intervention group and 63 in the control group. Five of the groups were lead by two or three students and the intervention took place at a location different from the students' ordinary work place. Seven of the groups consisted of one student and one staff member at the student's ordinary work place.

### *Design and statistical analyses*

The overall design was a mixed design with treatment (ART vs no treatment) as the group factor and pretest–test as the repeated measures factor. Analysis of variance (ANOVA) was used to compare differences in scores on the individual instruments between pretest and test, and between the ART and control groups. The alpha level was set to 0.05

### **Results and discussion**

We first present the results related to effects of the ART intervention. Then we discuss results related to a possible 'secondary' diffusion effect.

#### *SSRS – parents*

The results are summarized in Table 2. As can be seen from the table, five of five predicted pre–post differences were observed in the ART group. In the control group, one of the pre–post comparisons (*Problem behavior*) was significantly different between pre and post.

#### *SSRS – teachers*

The pre–post comparisons of the teacher scores are shown in Table 3. Three out of four predicted differences were observed in the ART group; in the control condition,

Table 2. Pre–post comparisons of SSRS subscales – parents.

	Pre	Post	
<i>ART</i>			
Cooperation	22.66	24.16	$F(1,102) = 9.44, p < 0.005$
Assertion	27.71	28.86	$F(1,102) = 6.13, p < 0.025$
Responsibility	28.36	29.79	$F(1,102) = 6.97, p < 0.01$
Self control	24.71	26.82	$F(1,102) = 13.62, p < 0.001$
Problem	28.83	26.74	$F(1,102) = 13.18, p < 0.001$
<i>Control</i>			
Cooperation	23.07	23.61	$F(1,102) = 2.37, p > 0.12$
Assertion	28.85	29.29	$F(1,102) = 0.77$
Responsibility	29.31	29.10	$F(1,102) = 0.13$
Self control	25.13	26.29	$F(1,102) = 3.58, p = 0.06$
Problem	27.81	25.75	$F(1,102) = 10.66, p < 0.005$

Table 3. Pre–post comparisons of SSRS subscales – teachers.

	Pre	Post	
<i>ART</i>			
Cooperation	25.10	25.40	$F(1,116) = 0.36$
Assertion	24.26	25.37	$F(1,116) = 4.90, p < 0.05$
Responsibility	–	–	–
Self control	23.38	24.89	$F(1,116) = 8.10, p < 0.001$
Problem	31.73	28.06	$F(1,116) = 13.88, p < 0.001$
<i>Control</i>			
Cooperation	26.41	27.43	$F(1,116) = 3.60, p = 0.06$
Assertion	24.84	25.11	NS
Responsibility	–	–	–
Self control	24.30	25.88	$F(1,116) = 8.03, p < 0.01$
Problem	29.11	27.34	$F(1,116) = 2.92, p = 0.09$

one comparison was different (*Self control*), two were marginally different, and one was not significant.

#### *CADBI – parents*

Table 4 shows the pre–post comparisons in the ART and control groups. Four out of five predicted differences were observed in the ART group. In contrast, and as predicted, none of the pre–post comparisons in the control group were significant.

#### *CADBI – teachers*

Table 5 shows the pre–post comparisons in the ART and control groups. Here, only two of five predicted differences appeared in the ART group. As predicted, none of the comparisons in the control group were significantly different.

Table 4. Pre–post comparisons, CADBI subscales – parents.

	Pre	Post	
<i>ART</i>			
Part 1	19.28	16.72	$F(1,105) = 5.53, p < 0.05$
Part 2	15.74	15.29	NS
Part 3	20.54	16.80	$F(1,105) = 10.62, p < 0.005$
Part 4	25.21	21.38	$F(1,92) = 8.26, p < 0.005$
Part 5	33.23	28.04	$F(1,103) = 13.44, p < 0.001$
<i>Control</i>			
Part 1	18.47	16.84	NS
Part 2	17.12	16.27	NS
Part 3	18.66	17.66	NS
Part 4	24.29	22.55	NS
Part 5	30.56	29.17	NS

Table 5. Pre–post comparisons, CADBI subscales – teachers.

	Pre	Post	
<i>ART</i>			
Part 1	17.53	16.15	NS
Part 2	21.04	17.95	$F(1,124) = 8.83, p < 0.005$
Part 3	24.42	22.63	NS ( $p = 0.14$ )
Part 4	28.44	28.56	NS
Part 5	40.29	35.92	$F(1,119) = 4.32, p < 0.05$
<i>Control</i>			
Part 1	15.17	14.81	NS
Part 2	17.27	16.16	NS
Part 3	20.28	18.48	NS ( $p = 0.16$ )
Part 4	25.48	22.81	$F(1,111) = 2.87, p = 0.09$
Part 5	33.64	31.70	NS

### Summary of outcome results

Overall, in the ART condition, 14 of the 19 comparisons of individual subscales turned out to be significantly different. In contrast, two of the 19 comparisons in the control group were significant. We conclude, therefore, that it is highly likely that the ART intervention had effect in increasing social skills and reducing behavior problems.

The most reliable changes in the ART condition seemed to occur in the parent evaluations. Thus, nine of 10 comparisons in the pre–post analyses of the parent evaluations were significantly different. In contrast, four of nine pre–post comparisons in the teacher evaluations were significantly different. The cause of this difference is not obvious, but it is well known that parents and teachers provide different evaluations of social skills and problem behavior (e.g. Achenbach, McConaughy, & Howell, 1987), and those differences may at least in part explain the differences in parent and teacher evaluations. For example, in the SSRS data, parents and teachers provided markedly different scores on the subscales *Cooperation* (overall mean, pre-test scores parents = 23.44; mean teachers = 26.28;  $F(1,123) = 25.63, p < 0.001$ ) and *Assertion* (overall mean, pre-test scores parents = 28.45; mean teachers = 24.99;  $F(1,123) = 39.37, p < 0.001$ ). These differences between informants are larger than the effects of the ART intervention and must be taken into account in explaining differences in treatment effects between parents and teachers.

### Secondary treatment diffusion

Another feature of the present data is that participants in the control conditions tended to demonstrate a positive change in pre and post comparisons. To illustrate this, two SSRS indexes were created, one for social skills and one for problem behavior. The social skills index was a combined score for the three subscales that parents and teachers have in common (i.e. *Cooperation*, *Assertion*, and *Self Control*). Figure 1 shows the net change (post minus pre) in the parents and teachers evaluations. As the figure indicates, all participants tended to improve, but the improvement was (as expected) larger in the intervention group.

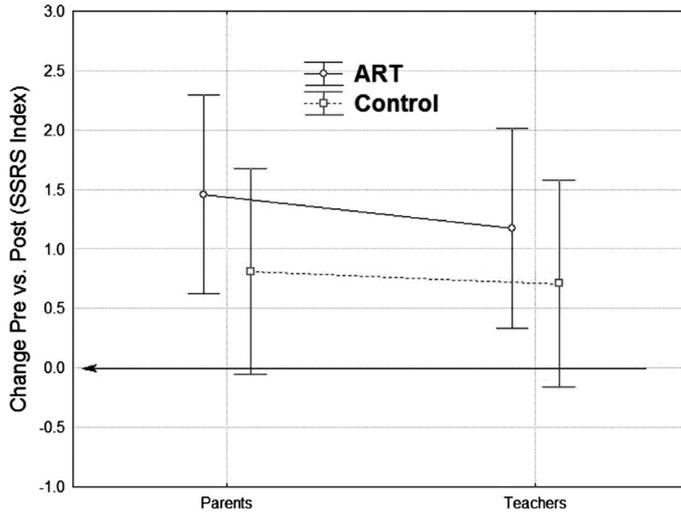


Figure 1. The net change, pre vs post, social skills index (0 indicates no change).

Similarly, comparison between the parent and teacher problem scores (SSRS subscale *Problem behavior*) demonstrates that both the ART and control groups changed in a positive direction. In fact, in the parent evaluations, there was no difference in improvement between the ART and control groups. Figure 2 displays the problem behavior data.

These observations indicate a relation between a positive change in the ART condition and a corresponding (but weaker) change in the control group. To obtain a better impression of this relation, we compared the ART and control groups in the 11 student projects that were included in the present analysis. The SSRS social skills index previously mentioned served as a measure of improvement. As can be seen

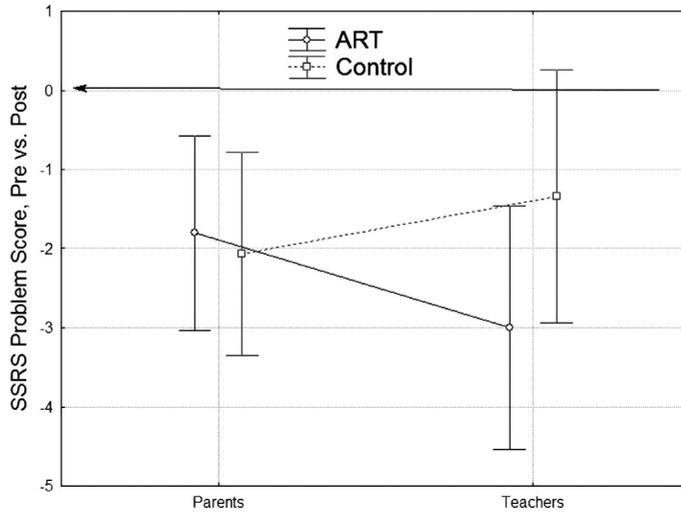


Figure 2. The net change, pre vs post, problem behavior (0 indicates no change).

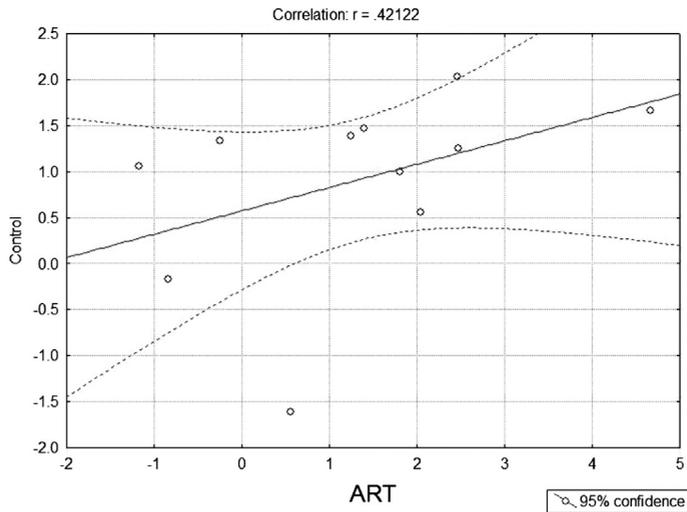


Figure 3. Correlation between pre and post changes in the 11 individual projects included in the present project. Dotted lines indicate 95% confidence intervals.

from Figure 3, there was a relatively high correlation ( $r = 0.42$ ) between improvement in the ART groups and a corresponding improvement in the control groups. This result indicates that participants in the control groups are affected by the intervention through some undisclosed mechanism.

The fact that a 'treatment effect' is observed in the control groups can be traced to several possible mechanisms. As previously discussed, regression to mean effects and effects of other interventions are unlikely as explanations of the observed changes. In the present context, note that the ART and control groups were essentially comparable prior to the intervention, but the control group demonstrated a somewhat better overall social skills score and also a higher problem score compared to the intervention group. A regression effect for the second measurement would then imply a pattern of change in the control group that would be opposite to that actually observed. The potential effect of other interventions cannot be excluded. However, we doubt that such interventions for only some participants in the control group would appear at the group level.

Second, the change observed in the control group might simply be due to a test-retest effect. For example, administering the test twice might cause a more favorable score on the second occasion due to increased attention to the relevant skills and problem behaviors. However, data from another Norwegian study using SSRS did not observe a change in SSRS scores between pre and post administrations in the control groups (Manger et al., 2000). Furthermore, Sæstad and Kyrrestad (2007) administered SSRS on two occasions to test a possible test-retest effect and found no evidence for such an effect. These observations indicate, therefore, that the change observed in the control group was not due to a simple test-retest effect.

A third possible mechanism is related to diffusion between conditions, i.e. poor implementation quality. We attempted to test this possibility by using a questionnaire that compared the groups on the following measures: (1) type of institution; (2) the presence of an identifiable role model; (3) support and interest

at the institution; (4) contact with the parents; (5) contact with teachers. The change index was subjected to ANOVA, with these scores (variables 2–5, scores 0 = low to 3 = high) as continuous independent variables. Only one of the variables, support and interest at the institution, was significantly associated with the change scores,  $F(1,89) = 7.42$ ,  $p < 0.001$ . There was, however, no difference between the ART and control groups on any of the implementation quality measures. The results from supervision also indicated that the program was carried out according to the design. This is also supported by the fact that absence (3,04) was spread over the three components and thus implies no serious threat to treatment integrity. Hence it is difficult to draw any specific conclusion from the implementation scores about a possible effect on control group performance.

A fourth possible mechanism is diffusion of treatment elements between ART and control group. In seven of the 12 projects, training was conducted by students and staff members at the organization where training took place. It is possible that staff members teaching control group participants could have adopted elements of the ART training into ordinary class and thus contributed to the positive results in the control group. However, since such contact with control group participants would have occurred in classes involving academic subjects, we consider this possibility to be low.

A fifth possible mechanism involved in control group changes is a 'secondary' diffusion effect. Because the ART participants demonstrated a larger degree of improvement compared to the control group participants, a secondary diffusion implying improvement should be expected to be initiated by ART participants and affect control group participants. This assumption is supported by the fact that there was extensive contact between members of the two groups in 10 of the 11 projects. It may be noteworthy that scores on the variable *presence of an identifiable role model* was low for both groups. Clearly, these low scores weaken such an explanation. On the other hand it must be remembered that this score was given by the ART trainers, not by the youths themselves. Furthermore, it is possible that new and more productive skills in ART group participants develop and affect others at the group level, and not via significant role models.

### General discussion

The present research indicates that the ART intervention has a reliable effect in enhancing social skills and reducing behavior problems. Our results also indicate that the positive changes observed in the control group may have resulted as a consequence of interaction between treatment and control group participants. We cannot pinpoint the exact mechanism involved in such a control group 'treatment effect'. Logically, interaction between treatment and control group participants works in two directions. For ART group participants, the skills trained in the ART sessions compete with skills observed in non-ART participants; for control group participants, the new skills observed in ART group participants compete with old skills still observed in control group participants. In this perspective, the positive 'effect' observed in control group participants may be increased from interaction with ART group participants, but the effect observed in ART participants may also be reduced by the influence of the control group participants. Obviously, the overall net effect was positive.

Research indicates that deviant peer affiliation is a strong predictor of delinquent behavior (Dodge, Lansford, & Dishion, 2006), and a high proportion of delinquent behaviors such as violence and drug use are committed in groups of adolescents who are known to be obsessed with peer culture (Warr, 1996). Elliott and Menard (1996) consider deviant peer affiliation to be an even stronger predictor of delinquent behavior than variables such as family, school, and community characteristics. For interventions, there is a substantial risk for negative outcome when aggregating delinquent youths in treatment groups (Dishion, McCord, & Poulin, 1996), and the normal recommendation is that programs for training social competencies should be carried out either individually or as a part of a school-wide program (Dodge et al., 2006). However, some programs with deviant youth have been found effective, but these typically include just one or very few deviant members (Dodge & Sherrill, 2006). To increase the probability of a prosocial culture in the group, Dodge and Sherrill (2006) therefore recommend populating the group with a majority of well-adjusted youths, imposing a high degree of structure that does not allow deviancy training to occur, employing a high ratio of leaders to youths, and employing leaders who have sufficient expertise to maintain control over group behavior.

In the current investigation the participants in each group represented different levels of behavior problems, from those with some degree of serious behavior problems to ordinary pupils without behavior problems. Even if the average student in the intervention group had slightly more behavior problems than the average class, the composition of the groups can rather be characterized as a vertical average of the class members in contrast to a horizontal group where students with equal degree of deviancy are put together. Combined with a firm structure, treatment fidelity and skilled instructors the different groups serve to keep a prosocial culture in the group.

The positive outcome observed in the intervention group in the current study is as expected both when comparing with other programs and also with similar ART studies. As previously discussed, deviant peer interaction is a strong predictor for delinquent behavior. The processes through which deviant peer contagion occur include factors like deviant training, modeling, labeling enhanced exposure to opportunities for deviance, and identity formation (Dishion, Dodge, & Lansford, 2006). It is also more enticement and social reinforcement for rule-breaking behavior in deviant groups; while typical outcome of aggregation in nondeviant group is enhanced prosocial competencies (Bagwell, 1999). As the current study was carried out in small schools where members of the control group and members of the intervention group knew each other and thereby also interacted with their old friends outside the intervention room the consequence could either be that the members of the intervention group were negatively influenced by the members of the control group or the opposite. The results of this study indicate that the social competencies learned in the intervention group also affect the control group in a positive way. As discussed, earlier recommendations for programs for deviant youths have been either individual treatment or treatment including all students in a school. Based on this study we will also suggest that programs should be directed towards an average of the students, which includes both deviant and nondeviant youths. The positive effects of the members of the control group in this study indicate a positive peer contagion outside the investigation room during the sessions. However, it is also suggested that the ART program in turn also should be directed towards members of the control group to strengthen the effect.

## Note

1. The social skills index was the mean of the three SSRS subscales that are common to parents and teachers, i.e. *Cooperation*, *Assertion*, and *Self Control*. The problem index was the mean of the parents and teacher problem scores.

## References

- Achenbach, T.M., McConaughy, S.H., & Howell, C.T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213–232.
- Achenbach, T.M., & Rescola, L.A. (2001). *Manual for the ASEBA School-Age Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.
- Austin, N.K., Lieberman, R.P., King, L.W., & DeRisi, W.J. (1976). A comparative evaluation of two day hospitals: Goal attainment scaling of behavior therapy vs. milieu therapy. *Journal of Nervous and Mental Disease*, *163*, 253–262.
- Bagwell, C.L. (1999). The friendship relations of antisocial boys. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, Vol. 60(1-B), 0356.
- Barnoski, R., & Aos, S. (2004). *Outcome evaluations of Washington States' research-based programs for juvenile offenders*. Report no. 04-01-1201. Olympia, WA: Washington State Institute for Public Policy.
- Burns, G.L., Taylor, T.K., & Rusby, J.C. (2001a). *Child and adolescent disruptive behavior inventory 2.3: Parent version*. Pullman, WA: Washington State University, Department of Psychology.
- Burns, G.L., Taylor, T.K., & Rusby, J.C. (2001b). *Child and adolescent disruptive behavior inventory 2.3: Teachers' version*. Pullman, DC: Washington State University, Department of Psychology.
- Conrad, K.M., Conrad, K.J., & Walcott-McQuigg, J. (1991). Threats to internal validity in worksite health promotion program research: Common problems and possible solutions. *American Journal of Health Promotion*, *6*(2), 112–122.
- Demaray, E., Maray, M.K., Ruffalo, S.L., Carlson, J., Burse, R.T., & Olsen, A.E. (1995). Social skills assessment. A comparative evaluation of six published rating scales. *School Psychology Review*, *4*, 648–671.
- Dishion, T.J., Dodge, K.A., & Lansford, J.E. (2006). Findings and recommendations. A blueprint to minimize deviant peer influence in youth interventions and program. In T.J. Dishion, K.A. Dodge, & J.E. Lansford (Eds.), *Deviant peer influences in programs for youth* (pp. 366–394). New York: The Guilford Press.
- Dishion, T.J., McCord, J., & Poulin, F. (1996). When interventions harm: Peer groups and problem behavior. *American Psychologist*, *54*, 755–764.
- Dodge, K.A., Lansford, J.E., & Dishion, T.J. (2006). The problem of deviant peer influences in intervention programs. In K.A. Dodge, T.J. Dishion, & J.E. Lansford (Eds.), *Deviant peer influences in programs for youth*. New York: Guilford.
- Dodge, K.A., & Sherrill, M.R. (2006). Deviant peer group effects in youth mental health intervention. In K.A. Dodge, T.J. Dishion, & J.E. Lansford (Eds.), *Deviant peer influences in programs for youth* (pp. 97–122). New York: Guilford.
- Elliott, D.S., & Menard, S. (1996). Delinquent friends and delinquent behavior: Temporal and developmental patterns. In J.D. Hawkins (Ed.), *Delinquency and crime: Current theories* (pp. 28–67). New York: Cambridge University Press.
- Feindler, E.L. (1995). Ideal treatment package for children and adolescents with anger disorders. *Issues in Comprehensive Pediatric Nursing*, *18*, 233–260.
- Fitzgerald, J. (2002). Internal validity of constructs within the conduct problem behavior domain. Unpublished.
- Gresham, N.J., & Elliot, S.N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.
- Goldstein, A.P., Glick, B., & Gibbs, J.C. (1998). *Aggression replacement training – A comprehensive intervention for aggressive youth* (rev. ed.) Champaign, IL: Research Press.

- Gottfredson, D.C., & Gottfredson, G.D. (2002). Quality of school-based prevention program. Results from a national survey. *Journal of Research in Crime and Delinquency*, 39(1), 3–36.
- Gundersen, K. (2003). Presjekk, en grovsortering for klassifisering av problematferd. Unpublished.
- Gundersen, K. (2005) Implementering av ART. In L. Moynahan, B. Strømgren, & K. Gundersen (Eds), *Erstatt Aggresjonen* (pp. 140–153). Oslo: Universitetsforlaget.
- Gundersen, K., & Moynahan, L. (2005). *Nettverk og Sosial Kompetanse*. Oslo: Universitetsforlaget.
- Gundersen, K., & Svartdal, F. (2006). Aggression Replacement Training in Norway: Outcome evaluation of 11 Norwegian student projects. *Scandinavian Journal of Educational Research*, 50, 63–81.
- Gundersen, K., Moynahan, L., & Svartdal, F. (2006). Rammebetingelser for betydning av ART prosjektene. Unpublished.
- Gundersen, K., Michell Olsen, T.M., & Finne, J. (2005). *ART – En metode for trening av sosial kompetanse. Trenermanual*. Sandnes: Diakonhjemmet University College.
- Iredale, T. (2000). Internal validity of the DSM-IV oppositional defiant disorder, attention-deficit/hyperactivity disorder, and conduct behavior problem symptoms. Unpublished.
- Kazdin, A. (1998). *Research design in clinical psychology* (3rd ed). Boston, MA: Allyn and Bacon.
- Kohlberg, L. (1984). *Essays on moral development: The psychology of moral development*. San Francisco, CA: Harper & Row.
- Madden, P.A.F., Bucholz, K.K., Todorov, A., & Grant, J.D.H.A.C. (2002). The assessment of peer selection and peer environmental influences on behavior using pairs of siblings or twins. *Twin Research*, 5, 38–43.
- Manger, T. (2000). Evaluering av prosjekt “KREPS”. Sosialt-kognitivt opplæringsprogram i ungdomsskolen. *Spesialpedagogikk*, 9–16.
- Manger, T., Eikeland, O.-J., & Asbjørnsen, A. (2002). Effects of training on pupils' social skills. *Journal of Research in Education*, 69, 80–92.
- McGinnis, E., & Goldstein, A.P. (1984). *Skillstreaming the elementary school child – A guide for teaching prosocial skills*. Champaign, IL: Research Press.
- McGinnis, E., & Goldstein, A.P. (1990). *Skillstreaming in early childhood – Teaching prosocial skills to the preschool and kindergarten child*. Champaign, IL: Research Press.
- Ogden, T. (2003). The validity of teacher ratings of adolescents' social skills. *Scandinavian Journal of Educational Research*, 47, 63–75.
- Payne, A.A., Gottfredson, D.C., & Gottfredson, G.D. (2006). Implementation of school-based prevention programs: Results from a national study. *Prevention Science*, 7, 225–237.
- Sæstad, T., & Kyrrestad, H. (2007). SSRS questionnaires administered over the internet: A viable alternative for assessing children's social skills, Master thesis. Tromsø: University of Tromsø.
- Salmon, S. (2004). The peace curriculum: Expanded aggression replacement training. In A. Goldstein, R. Nensen, B. Daleflod, & M. Kalt (Eds.), *New perspectives on Aggression Replacement Training* (pp. 171–189). Chichester: John Wiley & Sons.
- Skansgaard, E.P., & Burns, G.L. (1998). Comparison of DSM-IV ADHD combined type and predominantly inattentive types: Correspondence between ratings and direct observations of inattention hyperactivity/impulsivity, slow cognitive tempo, oppositional defiant, and overt conduct disorder symptoms. *Child and Family Behavior Therapy*, 20, 1–14.
- Warr, M. (1996). Organization and instigation in delinquent groups. *Criminology*, 34, 11–37.
- Wilson, S.J.M., & Lipsey, M.V. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine*, 33, 130–143.
- Wilson, S.J., Lipsey, M.W., & Derzon, J.H. (2003). The effects of school-based intervention programs on aggressive behavior: A meta-analysis. *Journal of Consulting & Clinical Psychology*, 71, 136–149.