Effect of ART intervention

Abstract

Eleven groups of students performed a 24-session intervention based on Aggression Replacement Training (ART) (Gottman, 1998) as part of their regular education program. Subjects were 65 children and youth with varying degrees of behavioral problems. 47 subjects received the ART program, 18 received standard social and educational services and served as comparison subjects. Social problems and skills were assessed before and after the ART intervention using Multi-subject Measures (MSM-1, MSNM-2, MSNM-3). Information in the ART group indicated significant improvement following the intervention, both in terms of increased social skills and reduced behavioral problems. In contrast, informants in the comparison group did not generally report improvement.

Key Words: aggression replacement training, social competence, social skills, anger control.
Aggression Replacement Training in Norway:
Outcome evaluation of 31 Norwegian student projects

Children and adolescents with behavior problems can create substantial difficulties for their families, the youth themselves, and society in general. Behavior problems include less serious forms like quarreling and not following rules, as well as more serious forms such as violent behavior against others, destruction of property, and truancy. Kandel (1997) estimated that 10% of children and adolescents exhibit behavior problems, of which 7% are of the more serious forms. Norwegian studies have indicated a similar overall prevalence of behavior problems, but the prevalence of serious behavioral problems is believed to be 1-2% (Dunbar, 2002; Wallerstein, 1996).

Behavior problems are related to a variety of risk and causal factors, and successful prevention and treatment programs should be expected to reflect this complexity (Andrews, 1998). Several studies have evaluated single- and multi-component intervention designs, and there is now an increasing agreement on the factors and combinations of factors that are important in effective prevention and treatment. Interventions that target risk factors associated with behavior problems are often based on social-cognitive, behavioral, and psychological theories. For instance, programs that teach children how to manage stress, improve self-control, and develop social skills have been found to be effective in reducing behavior problems (Andrews, 1998).

Among the intervention methods that address these factors, those with origins in cognitive-behavioral approaches seem to be especially well documented (Dawson, 2000; Gott, 1983; Ugray, 1982). These approaches include moral
measuring training (Gibbs, 1994; Arthure, 1986), anger control training for adults (Noyes, 1975) and adolescent (Miller, 1995), social skills training for adolescents and peer/decisions (Klein et al., 1983) and aggressive children (Kralovick, 1983), and parent management training (Patterson, 1982). Thus, many of these programs prove effective are, directly or indirectly, linked to social competence. Consistent with this, Gottfredson (1977) concluded that the most effective programs are those that include “a range of social competency skills (e.g., developing self-control, stress management, responsible decision-making, social problem solving, and communication skills)” (p. 55). Further, Oliva (2000) concluded that the most promising method for treatment of behavioral problems is the training of a broad spectrum of social skills in relation to peers as well as adults, combined with the rectification of problems in conduct.

At present there are several programs in use in Norway aimed at the prevention and remediation of behavioral problems. The most ambitious enterprise in this respect is the nationwide establishment of Parent Management Training (Patterson, 1982), Webster-Stratman’s program for children (Webster-Stratman, 1981), and Multidisciplinary therapy (Haugland, 1996). ART, Aggression Replacement Training (Tobin, 1999), is also gaining momentum in Norway as well as in other Scandinavian countries. In Norway, ART is approved for prevention and treatment of behavioral problems in schools (Settmaker, 2003) and homes for youth with behavior problems (Andersson, 2003), and is implemented in a number of schools and institutions (e.g., Lindsey Child and Youth Care Centre, Rogaland, several schools in Larvik, Ystefjell, and Hau, Rogaland, and Glance Antun Senior, Ystefjell). A two-year further education ART program is currently offered in Rogaland College.
Although MST, PAT, and ART share many common features, one difference is that whereas MST and PAT are system-oriented in their approach to the understanding and treatment of behavioral problems, ART is focused more on the development of individual social competencies. In the light of the studies reviewed, this may indicate that ART should be very promising in prevention and treatment of problem behavior. A number of outcome studies especially in the U.S. (Davies & Ax, 2004; Snydor, 1999; Goldstein, 1987; Goldstein, 1994) do indeed indicate good effect of ART interventions by increasing social skills and reducing behavioral problems (see also Anderson, 2003).

Evaluation studies of the Norwegian PAT and MST programs are currently being conducted (e.g., Sydor & Holldorf-Brinkhaus, 2006), and the outcomes seem to be positive and agree well with conclusions from previous American studies. As for ART, except for promising results obtained with enhanced ART with autistic children (Meyerson, 2005), no study has so far documented its efficacy when applied to Norwegian or Scandinavian children or youths. The primary goal of the present paper was, thus, to determine whether positive effects observed in American studies could be replicated in a Norwegian sample. The goal is of interest for at least two reasons (Sydor, 2006). First, differences in the social and political climate between the U.S. and Norway regarding the care of youths exhibiting behavioral problems may affect the efficacy of interventions. A youth who commits a criminal act in the U.S. is typically subject to intervention by the juvenile justice authorities, and a likely reaction is punishment and dispositions designed to keep the community safe. In contrast, in Norway, measures for offenders under 15 years cannot be punitive or designed to protect the community (no one under 15 years of age can be...
imposingly, instead, they are made in the best interest of the child. Second, since social services, healthcare, and education are more comprehensive and more universally accessible in Norway than in the U.S., the baseline for intervention is different between these countries. Thus, the expected gain of an intervention for ART, MFT, or PMT over the services already provided to youth in Norway is likely to be smaller compared to the corresponding gains seen for intervention in the U.S. ( disprovid, 2006).

In the present study, intervention effects were assessed using a between-group design with repeated measures. ART subjects received 24 sessions of ART training; comparable comparison subjects received standard social and educational services. Before the ART intervention, a comprehensive battery of instruments was administered to assess behavioral problems and social skills. Following the intervention, the same test battery was repeated. The battery included: Social Skills Rating System (SSRS; Gresham, 1990), Child’s self-Report Behavior Deficit Scale (CBDS) (Benn, 2001, a, and b), Disruptive Behavior (DB) (Gresham, 2001), and Child Behavior Checklist (CBCL) (Achenbach, 2001). In addition, a computerized self-report problem behavior questionnaire was developed (Andresen & Freidman, 2001). Since these instruments varied according to focus (behavioral problems, social skills, mental health), and informant source (parent, teacher, youth), they allowed for specific analyses across referrals and sites.

Overall, the ART intervention as implemented in this study should generate improved social skills and mental health, and reduced behavior problems, no significant change should be observed in the comparison group. Since the ART intervention here is implemented primarily in the school context,
Effect of ART interventions

Teachers’ judgments might be expected to be the most reliable sources for intervention outcome evaluation. But because ART interventions are designed to accrue generalizations of skills by assigning home work, use of realistic life dilemmas, and so on (Goldstein, 2006), we expected that the ART intervention should generalize well, and hence that parents should indicate beneficial effects. However, it is known that teachers and parents do not always agree in their estimates of problem behavior (Gerner, 2005; Schenkraft et al., 1997). This difference is most probably not simply an effect of differences in time spent with the youth, but more likely reflects the fact that parents and teachers observe children in different contexts (e.g., Mesz, 1997). These considerations complicate the outcome predictions for this study, but the most optimistic is one implying positive effects by all informants in the ART group, and no change for informants in the comparison group.

Method

Subjects

Sample participants: The ART trained were 25 students (14 females, 9 males, mean age = 38.1 yrs), participating in a 6-month, further education course in “Training of social competences” at Rogaland College, Norway. All students had no basic 3 years college education in teachers or authorized social education, and all were employed full time at local schools and institutions. Prior to the ART intervention, all students had theoretical and practical training in ART and related topics corresponding to 30 credits education. This included responsibility for 24 sessions ART intervention with colleagues or students.

Children and youth: Subjects were 55 children and youths from schools and institutions in Western and Eastern Norway. Most of them were recruited...
Effect of ART intervention

From the ART-students’ workplace, there were 36 girls (mean age = 14.5 years) and 40 boys (mean age = 15.5 years). Potential participants were stratified with a simple randomization (randomization, 2003) to differentiate these general levels of behavior problems: (1) None (the youth may need to train procedural skills, but has not developed significant behavioral problems); (2) Mild (the youth displays some degree of behavior problems); (3) Severe (the youth displays severe degree of serious behavior problems).

Consent and approval by ethics committees. Parents, teachers, and youth involved in the study received consolidated written and oral information about the investigation. The information was presented by the ART-students in meetings with groups and individually. Participants signed their consent to participate. The protocol for the project was approved by regional ethics committee.

ART intervention

ART (Goldstein, 1998) focuses on the development of social skills by the individual. More specifically, ART consists of three distinct components: addressing training of social skills, anger control, and moral reasoning. Anger control training (the affective or emotional component) enables youths being trained to recognize their external and internal triggers for aggression, aggressive signals, and to control anger using various techniques. Social skills training (the action component) focuses on training a selection of fifty structured skills, some are quite simple (e.g., listening to someone else, starting a conversation), others more complex (e.g., avoiding disruptive, handling group pressure). In moral reasoning training (the thought and values component), participants learn to reason in a more advanced manner to respond to challenging ethical and moral dilemmas, and to handle situations in their own lives in line with their own moral
Training takes place in groups with 5 to 8 participants. Groups are composed with regard to age, similarity of behavioral challenges and friendship between participants. Eftel and Greenberg (1991) also recommended the inclusion of group members with higher level of social competence as positive role models. Two trainers conduct ART sessions. Rules and consequences for infractions are clearly defined. Participation is voluntary and the use of positive reinforcement and small prizes are highly recommended to secure the motivation of trainers.

There is a firm structure in the program including defining the theme of the session, demonstration, role-playing, questioning, where and when to use the skill, feedback evaluation and homework. The purpose of feedback to the youths is primarily to shape proper behavior. All participants have different observation tasks. Goldwater et al. (1998) recommend that all three components (skill training, anger control training, and mood monitoring training) are scheduled for training at least once per week over a period of ten weeks. In order to transfer and maintain skills, it is important to establish contact with important individuals (family members, teachers, skill leaders etc.) in the participants’ environments.

The ART group received a standard 24-hours ART intervention with at least 4 sessions for each component of skills training, anger training, mood monitoring training. The duration of the intervention period was 10 weeks. For practical reasons, the number of sessions was reduced from the recommended number 30 (Goldwater et al., 1998). In average, the intervention had 10.8 sessions.
of social skills training, 8-8 sessions anger control training, and 4-8 sessions arousal
management.

Treatment fidelity. Substantial information and supervision was made
available to the ART students before and during the implementation of the ART
intervention. A web page with detailed instructions was available at all times. To
ensure sufficient trainer quality, each student was required to complete at least 24
sessions as ART trainer before the ART intervention. During the intervention,
supervision (either from videotapes or in vivo) was provided. The students also
evaluated each session during the intervention, and the evaluation was subject to
approval from the college. Following the intervention, students filled out a
questionnaire to identify to which degree they had implemented ART according to
treatment fidelity factors described by Bellin (1995). Finally, overall ART
competence was evaluated by colleagues and superiors in a questionnaire prior to,
and after training. The outcome of this evaluation indicated significant
improvement. Overall, we believe that extensive pre-training and continued
supervision during the intervention secured good treatment fidelity (Bellin, 1995).

Comparison group

The comparison group received the standard child social and educational
services as defined and delivered by the actual school institution. In some cases
this meant no special arrangements. For five of the sub-studies, the comparison
group was enrolled from the same class or school as the ART group. This
represents a potential threat to the design of this study, since the intervention
(which was distributed over 15 weeks) could be defined in that the comparison
groups were affected by the ART program and/or changed behavior in the ART
groups. However, this limitation was occasioned by practical circumstances and could not be remedied.

Instruments

Four instruments were selected to assess behavioral problems and social skills, using multiple informants (youth, parent, teacher). The instruments were mainly chosen among internationally validated instruments. Two were available in Norwegian versions (CBCL, SRS), two were translated as part of this study (CAHRS, HHI). In addition, a custom-made self-report problem behavior checklist was included.

CBCL - Achenbach School-age Forms & Profiles (Achenbach & Rescorla, 2001) comprises an informant-variant or multi-informant questionnaire consisting of the Child Behavior Checklist for Ages 6-18 (parents), Teacher’s Report Form, and Youth Self-Report. CBCL is based on practical experiences and research over 40 years, and is now regarded one of the best-validated tools for assessing behavioral functioning in children and youth. CBCL is translated into Norwegian, and studies have demonstrated acceptable validity (Neculă, 1999).

In the present study, raw scores (rather than T-scores) were used. The internal consistency of the subscales included in the analyses was very high for each informant, r = .80 (parents), .86 (teachers), and .50 (youths).

CAHRS: The Child and Adolescent Hypothetical Rating System (Behavior Inventory-2.3) (CAHRS) has a parent and teacher version (Barnes, 2001). Both versions contain assertions relating to various forms of behavior and behavioral problems towards adults (parents, teachers) and peers, activity level, and emotional skills such as “Enthusiastic physical fights with peers” (“Does not seem to show concern for the feelings of others”). Each of the assertions is scored on an 8-point frequency of
occurrence scale (never in past month, 1-2 times in past month, 3-4 times in past month, 2-4 times per week, 1 times per day, 2-5 times per day, 6 or 9 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 5-point scale (from severe difficulty to excellent performance).

CAMHI has achieved good internal consistency 3-month test-retest reliability (Egeland, 2002; Lundby, 2008; Skaregaard, 1993), but the validity of version 2.5 is not well documented. For the present investigation, CAMHI was translated into Norwegian (Myhland, 2005). The internal consistency of the subscales included in the analysis was very high (—0.84 for the parent version, and 0.80 for the teacher version).

SBS: Social skills inventory (Gottman, 1996). Norwegian version (Oplien, 2003) is a standardized, norm-referenced scale measuring children and adolescents social skills, problem behavior, and academic competence. The validity of SBS has been demonstrated in several studies (Gottman, 1995; Demary, 1999; Oplien, 2003). It was chosen for this study because of its obvious relevance to measurements of social skills since we included measures of other instruments or measures problems behavior, the problem scale of SBS was not used. The SBS items group into four subscales: cooperation, assertion, self-control and responsibility. (Allomitive, social, power, andCADESS scale). This paper includes 30 statements about social behavior in terms of frequency. For example, cooperation includes behaviors such as helping others, sharing materials, and responding to the needs of others (e.g., “helps with tasks,” “completes task on time”). The rating scale for each item was originally a 5-point scale, but Oplien (2003)
increased the to four (1 = never, 2 = sometimes, 3 = often, 4 = very often), which was used in the present study.

In the present sample, the internal consistency for the SRS subscale was high (α = .91 for the parent scores, and .79 for the teacher scores).

HIT. (How I Dead (Gillen, 2003)) is a questionnaire designed to measure self-serving cognitive distortions in clinical youths. It contains 34 items, 39 cognitive distortion items (in addition, there are 6 non-responding forms and 7 positive filler items). Those are divided into four cognitive distortion subscales: self-control, blaming others, restructuring/stereotyping, and asserting the self. The subject’s response to each item statement is marked on a 5-point Likert-type scale from “disagree strongly” (1) to “agree strongly” (5). The 4 HIT items are designed to screen for inaccurate or inaccurate responding. The current version has been validated on four validation samples and exhibited good reliability on every reliability measure. HIT also exhibits high internal consistency, high test-retest reliability, and generally good construct validity (Bergin, 1994). High correlations have been reported between HIT and the self-reported externalizing scale on the CBCL (Achenbach, 1991) after controlling for internalizing disorders.

For the present study, Overall HIT was used for all statistical analyses. This index includes all subscales of HIT. In the present sample, the internal consistency for these subscales was high (α = .80).

Self-Report Questionnaire: Since CADRE has no teacher and parent versions, but no youth version, we developed a questionnaire with 56 statements to be filled out by youth. Several of the items were adapted from a Norwegian MBT study (Storber, 2004). The items were divided in 3 subscales: prosocial
The response to the pro-social statement is marked along a 7-point scale from “does not at all” to “does very well.” Responses to the problem behavior statements are marked on a 9-point scale according to how many times the named behavior has occurred the last month.

**Procedure**

The student ART teams were divided into 31 sub-groups, normally 2 students per sub-group, constituting an ART main team and a control. The formation of these groups was determined primarily by practical considerations (e.g., student working place). From the sub-groups, performed their ART intervention at the workplace of both participating students. Five sub-groups performed the intervention at the workplace of one of the participating students, and the remaining two sub-groups performed the ART intervention at a location different from their own workplace.

Following screening of all potential participants, each sub-study established two client groups representing comparable behavior problems. The group composition was performed according to the recommendations by Achenbach (1991). Each group (6) was required to include at least one youth classified at level 5 (most behavioral problems). Since one of the youth was not of importance here, the sex factor was not manipulated. One of the groups was then, based on a randomized procedure, allocated to the ART condition, the other to the comparison group. The allocation procedure – which was performed per groups rather than on individual participants – was carried out by a superior of the student teams according to instructions provided by the project leaders. For some of the sub-studies, however, randomization was not possible. In one case, the teacher-assistant for the teacher assessed both groups at a school.
included in the ART group. In three sub-studies, training was not to special
schools with small units (youths with Asperger Syndrome, psychiatric problems,
comprehensive behavior problems), and there was not sufficient youths available
to establish two comparable groups. In one sub-study, three of the comparison
groups needed a referral for the project. We performed complete analyses on
the total sample as well as the sub-sample that satisfied the criteria for the
randomization procedure. Since these analyses rendered almost identical
outcomes, we decided to include all subjects in the analyses reported here.

Thus, a total of 47 youths received the ART intervention and 31 youths
were included in the comparison group. The unequal group size was unfortunate,
but we reasoned that this would not be critical to the study. An important
consideration here is that allocation to groups was not influenced by the youths,
teachers or parents, implying that selection or self-selection was not an issue.

Another consideration is that the overall study would compare relatively few
participants in homogeneous sub-groups, implying that the possibility of
statistically valid group comparisons (ART vs. comparison group) would be
limited. Thus, within-group comparisons (pre vs. post) would probably be the
preferred criteria for measurement of change. The latter is supported by the fact
that the study involved a large number of repeated measures, thus increasing
statistical power.

Six of the sub-studies took part in special schools, two in special schools
for adolescent with behavior problems, one in an institution for adolescent with
behavior problems, one in a psychiatric child clinic, and one in a specialized
group for pupils with Asperger Syndrome (7 subjects). Table 1 shows the actual
distribution of youths in the three behavior problem classifications and over types of institutions.

Design and statistical analyses

The overall design was a mixed design with treatment (ART vs. no treatment) as the group factor and pretreatment as the repeated measures factor. To limit the weakness related to the assessment procedure, the three participants with the highest problems in each group were assessed. One way Analysis of Variance (ANOVA) was used to compare differences in scores on the individual instruments between pretreatment and posttreatment and between the ART and comparison groups. As noted, within-group changes from pretreatment to posttreatment were the primary treatment measure. Specific predictions, i.e., tests of change between pretreatment and posttreatment levels, were performed by contrast analysis (Rosenthal & Rosnow, 1991). In addition, we ran analyses to check for moderator variables (e.g., sex, age, motivation). These analyses are reported separately.

Results and discussion

The overall expectation for the present data is that participants in the ART condition should demonstrate improved social skills and reduced occurrence of problem behavior, whereas little or no change should be seen in the comparison group. In the following presentation, this expectation is examined for each of the three areas of focus of the present study, social skills (SSRS), moral thinking (ETI), and problem behavior (CADE, CINC). Since intervention efforts were the main focus of this article, results pertaining to this issue are presented in the text; summary of results are presented in Tables 2 and 3.
Social skills: SRS

Pretreatment data scores on all the four subscales of SRS, Cooperation, Expression, Responsibility, and Self-Control were averaged and subjected to analysis. A 2-way ANOVA of the group and pretreatment data showed no overall significant effects of the Group or pretreatment factors, but a significant Group x pretreatment interaction, F(2, 48) = 5.33, p < .02. Contrast analysis demonstrated that the increase from pretreatment to post was significant within the ART group, F(2, 48) = 10.87, p < .002, whereas the change in the comparison group was nonsignificant, F(2, 48) = .04. These results indicate that the ART and comparison groups changed differently from pretreatment to post, with a significant increase in social skills in the ART group but no change in the comparison group.

Teachers: The teacher checklist includes four subscales, Cooperation, Expression, Responsibility, and Self-Control, which were averaged for the present analysis. The ANOVA showed a significant effect of the pretreatment factor, F(2, 51) = 9.28, p < .001, but nonsignificant effects of Group and of the pretreatment x group interaction. Contrast analysis showed that the change from pretreatment to post within the ART group was significant, F(2, 51) = 11.97, p < .001, whereas the corresponding change within the comparison group was not significant, F(2, 51) = 1.74, p > .26. Thus, the ART group demonstrated a significant increase in SRS scores, but the comparison group demonstrated a nonsignificant change.

Children's ratings. The children's youth version of SRS includes 34 (children) and 39 (youth) questions related to subjective evaluations of social skills. The ANOVA did not indicate significant effects of group, pretreatment, or the group x pretreatment interaction. Moreover, the contrast analysis indicated a
The study included 19 participants and their social skills were assessed using a 10-item measure of social skills. The ANOVA showed a significant effect of the pretest on the group variable, F(1, 2) = 3.75, p < .05, indicating that the pretest scores were significantly different between groups. The effect was not significant when the posttest was used, F(1, 2) = 1.35, p = .22. No significant change was found between the ART group and the control group, F(1, 2) = 1.15, p = .26. However, a significant increase in social skills was observed in both groups, F(1, 12) = 6.99, p < .05.

For the analysis, a mixed-design ANOVA was used. The ANOVA indicated a significant effect of group on the dependent variable, F(1, 50) = 7.86, p = .003, suggesting that the ART group had higher social skills compared to the control group. No significant effect of time was observed, F(1, 50) = 1.35, p = .25. The interaction effect between group and time was not significant, F(1, 50) = 2.37, p = .13. No significant change was observed in the control group, F(1, 2) = 1.15, p = .26.

The results of the analysis showed that the ART group had significantly higher social skills compared to the control group. The main effect of group was significant, indicating that the ART intervention was effective in improving social skills. No significant effect of time was observed, suggesting that the improvement was consistent throughout the intervention.
The ANOVA indicated a significant effect of the two-group factor, $F(1, 46) = 7.57, p < .01$, reflecting a reduction in scores from pretest to test in both groups. The overall group difference and the pretest x group interaction were not significant. Contrast analyses showed that the reduction from pretest to test was significant within the ART group, $F(1, 45) = 4.11, p < .05$, whereas the change within the comparison group was not significant, $F(1, 45) = 2.15, p > .10$.

Children's scores on the CBCL Youth Self-Report Scale (Achenbach & Edelbrock, 1983) were a measure of subjective problems by our caretaker-child self-report form. The form includes 68 questions related to behavior problems. The ANOVA did not indicate significant effects of group, pretest x test, or the pretest x group interaction. However, contrast analyses indicated a significant reduction of problems from pretest to test within the ART group, $F(1, 41) = 5.97, p < .05$, but no significant change in the comparison group.

$\Rightarrow$ Table 3 here $\Rightarrow$

Procedural Behavior: CBCL

Parents. The subscales INT, L/Y, and OTHER of CBCL were included in the analysis. The ANOVA showed a significant effect of the group manipulation, $F(1, 45) = 4.65, p < .05$, of the pretest x test factor, $F(1, 45) = 18.75, p < .001$, as well as for the pretest x group interaction, $F(1, 45) = 3.65, p < .05$. Contrast analyses demonstrated that the reductions from pretest to test within the ART group were significant, $F(1, 40) = 27.56, p < .001$, whereas there was no significant change in the comparison group, $F(1, 40) = 0.20, p > .50$. This indicates that the
ART group demonstrated reduced scores on the subscales following the ART interventions; no change was seen in the comparison group.

Further, for the patients' dems, the subscales INT, EXT, and OTHER of CIRL were included in the analysis. The ANOVA indicated a significant effect of the pretest manipulation, F(1, 47) = 5.10, p = .03, reflecting an overall reduction in scores on the three subscales from pretest to test. The group and the pretest x group interaction were not significant. Contrast analyses showed that the reduction within the ART group was significant, F(1, 47) = 3.14, p = .08, whereas there was no significant change in the comparison group, F(1, 47) = 4.39.

Conversely, again, the subscales INT, EXT, and OTHER of CIRL were included in the analysis. The ANOVA indicated a significant effect of the pretest factor, F(1, 50) = 17.51, p = .0002, reflecting the fact that subscale means were lower on test compared to pretest. The pretest x test and pretest x group interaction were not significant. Contrast analyses showed that the reduction within the ART group was significant, F(3, 50) = 14.91, p < .001, as was the corresponding reduction in the comparison group, F(3, 54) = 5.05, p = .002.

Effect of moderator variables

We also performed analyses to assess the influence of variables that were not manipulated or included in the design. As noted, due to the requirements of group comparability, the sex factor was not manipulated in our design. Seventeen of the subjects were females, nine of which were in the ART condition. Due to the low number of females, we decided not to include sex in post hoc analyses.

Age level. Our subjects spanned from 11 to 77 years of age. Overall, five of seven tests with age level as a covariate indicated that age was not significantly related to outcomes. The exception was CIRL (Parent scale) and CAMH.
Effect of ART intervention

(Teacher ratings) Consistent with existing research (Wilson et al., 2007), the overall conclusions were not affected by the omission.

Analysis: We ran separate analysis with institution (place of intervention) as a covariate. In none of the tests did this covariate indicate that institution has a significant effect on outcome.

General discussion

The primary finding of this report was that a 24-hour ART intervention produced a positive and reliable outcome in that social skills were improved and behavioral problems reduced from pre-test to post-test. Thus, the ART group demonstrated significant improvement on 4 out of 12 items. To our knowledge, this is the first demonstration in Scandinavia of the positive effect of ART in treatment of behavioral problems. Our results are similar to corresponding U.S. studies (Nagert, 1999) and indicate that a positive effect of an ART intervention despite cultural differences between the U.S. and Scandinavia and also despite differences in baselines for intervention.

It is necessary that this positive outcome was achieved despite the fact that the intervention was implemented by students with relatively limited experience in ART training, and intervention also had a limited duration (24 hours). The explanation of this most probably is that the smaller ART teams had extensive knowledge and practical training before the ART intervention commenced, and also that the protocols were supervised throughout the intervention phase.

Although the intervention was carried out in school/institution settings, positive effects were observed also in home settings as evaluated by parents. Thus,
as expected, the effects of ART generalised to home settings. The ART program strongly emphasizes methods to enhance generalization and maintenance of skills, and these were also part of the present projects (e.g., homework, hand signals, and trainee variation in role playing, teaching, and use of real-life situations). Although even better results might have been expected (other factors specifically aimed at enhancing generalization, e.g., incorporating peers as functional mediators to the group (Shamay-Tsoory et al., 2002) and separate training of parents (Cilamari, 2003) had been included in the intervention, our results clearly indicate good generalization of effect. One also needs to take into account that the focus of the five self-reflections provided by the parents in the unilateral condition was from the ART intervention. This indicates that cross-contextual effects confirm the generalization of intervention effects. In this context it should also be noted that the results from the HIT measure were positive, despite the fact that the intervention only contained, on average, 4.5 sessions with social reasoning training (i.e., half the training spent on skills training and aggressive management training). This indicates that training within any one of the three elements of ART (anger control, prosocial skills, and moral reasoning) also generalizes to the others (Korczak, 1996).

Our results indicate that the same parents of positive outcome regardless of informant (parent, teacher, and youth). Previous research has demonstrated that parents and teachers demonstrate differences in their ratings of problem behavior (Korczak, 2003; Ackerman et al., 1987), most likely because parents and teachers observe problems in different settings. Furthermore, the ART intervention is used in the same manner as the teachers do. It should be noted, however, that our analyses do not specifically address multi-
informant agreement in judgments. Such analyses are pursued in a different context (Goldman & Scovel, 2006).

The use in this study of a comprehensive test battery to measure changes in problem behavior and social skills also demonstrated a high degree of consistency in results between tests. This test used all seemed to reflect the positive outcome of the ART intervention. Again, as for informants, we did not measure agreement among the instruments specifically, but we are no reason to assume that any one of the instruments was more sensitive to treatment effects.

It should also be noted that two of the subjective measures of this study (SCL, J scale, and HDL) reflected positive outcomes following the ART intervention, but this outcome occurred also in the control group. This result should be seen in the light of the fact that several of the interventions took place under circumstances that made adequate separation of the ART and comparison subjects quite difficult (i.e., a possible dilution between conditions). The effect did not, however, show up in teachers’ judgments, which most probably would have been the case if dilution between conditions had been a serious problem in this study. The fact that the profile themselves indicated similar improvement irrespective of condition may, however, indicate that some dilution occurred.

Although the overall outcome of the interventions analyzed here was positive, some qualifications must be pointed out. First, our total sample was relatively small, and the standardization procedure did not succeed in establishing proper treatment and control groups. Our conclusions are therefore not based on a proper randomized control group design, but primarily on within-group comparisons before and after treatment. Although proper control group comparisons definitely would have been preferable, it should be noted that
Effect of ART interventions

randomization of participants is complicated by the nature of the ART interventions. Rather than focusing on individuals, the ART interventions focus on a group, and this group should be composed in a specific manner (Goldstein, 1998). As a consequence, randomization must be performed on group level, and this procedure is, as we have experienced, quite vulnerable. Future research could compensate this situation by including larger samples selected from larger pools of potential participants.

A second issue is the role of moderator variables on outcome of intervention. Consistent with earlier research (Wilson et al., 2003), age level did not mediate the effect of intervention. Further, we could not find systematic variations in outcome as a function of institution (place of intervention). From our perspective this is not surprising. The ART trainers were students educated in the same class, they acted as trainers in pairs, and received common instructions and supervision. On the other hand, it is well-known that a variety of moderator variables do affect intervention outcome, and trainer quality is one of a number of relevant variables (e.g., Baroudi & Av, 2004; Hollin, 1995). Among other relevant factors are the frame conditions for implementation (e.g., support in organization, group size, role in society). Analysis of such factors should attempt to quantify the quality of these variables to assess their effect on outcome.

We are currently pursuing this issue in a more detailed analysis of institutional factors related to outcome effects of ART interventions (Gundersen & Soreide, 2004).

Finally, the positive outcome of this study must be interpreted in light of the fact that our ART trainers, being highly selected and recruited from all over Norway, most probably were highly motivated and particularly well-skilled.
Wilson et al. (2003) also found that routine programs showed smaller effects than those introduced and supervised by researchers. Thus, it would be a serious recommendation that future research evaluated ART interventions that were not performed as “promised” projects. Future research on the effect of ART interventions should also plan for systematic follow-up studies. We suggest that such follow-up studies should include subject age levels that secure the participation of all informants in all phases of the intervention. In Norway, a sound follow-up investigation should take into consideration the fact that most juvenile delinquents leave school between the ages of 12 and 14 yrs.
References
Burlington, VT: University of Vermont, Department of Psychiatry.
Behavioral and Emotional Problems: Implications of Cross-Situational
Correlations for Situational Specificity. Psychological Bulletin. Vol. 101, 213-
232.
behavioral and emotional problems: implications of cross-situational correlations
Forms & PopEss. Burlington, VT: University of Vermont, Research Center for
Children, Youth, & Families.
Oslo: Kommunalbygget.
for young offenders: The effect of clinically relevant treatment. Reforming
Psychology: Criminal Justice and Behaviour, Vol 17, No 1, 19-52.


Effect of ART interventions

25

victimization: A meta-analysis. Canadian Journal of Criminology, October, 446-
467.


Anger Disorders. In H. Karimova (Ed.), Anger Disorders: Definition, Diagnosis,
and Treatment. Taylor & Franich.

Freedland, J. (2002). Internal validity of outcomes within the context of problem behavior


American Guidance Service.

school-based counseling in delinquency. Journal of Counseling and Clinical
Psychology, 12, 27-45.

Champaign, IL: Research Press.

Generalization of Goals. Champaign, IL: Research Press.


Unpublished Work.

Unpublished Work.


Effect of ART interventions


## Table 1. Number of youth in the three problem behavior classification and type of institution

<table>
<thead>
<tr>
<th>N</th>
<th>Problem level</th>
<th>Type of institution (distribution of problems levels 3, 2, 1 in parentheses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>7: N=18 (38.3%)</td>
<td>School: N=24 (4, 12, 6)</td>
</tr>
<tr>
<td></td>
<td>2: N=25 (54.7%)</td>
<td>Special school: N=19 (4, 2, 2)</td>
</tr>
<tr>
<td></td>
<td>1: N=4 (8.7%)</td>
<td>Institution: N=14 (4, 7, 3)</td>
</tr>
<tr>
<td>18</td>
<td>7: N=5 (27.8%)</td>
<td>School: N=13 (4, 8, 1)</td>
</tr>
<tr>
<td></td>
<td>2: N=10 (55.5%)</td>
<td>Special School: N=6 (4)</td>
</tr>
<tr>
<td></td>
<td>1: N=3 (16.7%)</td>
<td>Institution: N=3 (1, 2, 0)</td>
</tr>
</tbody>
</table>
Table 2. Social skills. Means and within-groups effects (pret vs. post) for the ART and comparison groups for the SRS and HIT instruments (SE in parentheses). *p < .01. Note that the HIT measure only applies to youths. Also note that subscale scores are provided based on different numbers of items included in each.

<table>
<thead>
<tr>
<th></th>
<th>ART</th>
<th></th>
<th>HIT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
</tr>
<tr>
<td>TMT</td>
<td>22.5(6.17)</td>
<td>21.2(6.33)</td>
<td>11.09*</td>
</tr>
<tr>
<td>Vis</td>
<td>18.7(1.16)</td>
<td>20.7(2.20)</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td>21.2(2.14)</td>
<td>23.9(6.42)</td>
<td>2.72(1.14)</td>
</tr>
<tr>
<td></td>
<td>12.4(1.30)</td>
<td>11.4(0.80)</td>
<td>1.30</td>
</tr>
<tr>
<td>Vis</td>
<td>8.5(2.02)</td>
<td>8.6(2.52)</td>
<td>1.38</td>
</tr>
</tbody>
</table>
Table 3. Social problems. Means and within-groups effects (present vs. past) for the ART and comparison groups for the CADRE and CBCL instruments (SE in parentheses). * = p < .05. Note that the youth measures in the CADRE column are from Gundersen and Sorensen (2013).

<table>
<thead>
<tr>
<th></th>
<th>CADRE</th>
<th>CBCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ART</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past</td>
<td>156.50</td>
<td>38.25</td>
</tr>
<tr>
<td>Std.</td>
<td>42.25</td>
<td>7.85</td>
</tr>
<tr>
<td>Low</td>
<td>25.82</td>
<td>3.90</td>
</tr>
<tr>
<td>Std.</td>
<td>5.25</td>
<td>1.10</td>
</tr>
<tr>
<td>Past</td>
<td>15.82</td>
<td>3.90</td>
</tr>
<tr>
<td>Std.</td>
<td>5.25</td>
<td>1.10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past</td>
</tr>
<tr>
<td>Std.</td>
</tr>
<tr>
<td>Past</td>
</tr>
<tr>
<td>Std.</td>
</tr>
</tbody>
</table>