

The Efficacy of a Universal School-Based Program to Prevent Adolescent Depression

Ian M. Shochet, Mark R. Dadds, Denise Holland, Kathy Whitefield,
Paul H. Harnett, and Susan M. Osgarby

School of Applied Psychology, Griffith University, Queensland, Australia

Evaluated whether a universal school-based program, designed to prevent depression in adolescents, could be effectively implemented within the constraints of the school environment. Participants were 260 Year 9 secondary school students. Students completed measures of depressive symptoms and hopelessness and were then assigned to 1 of 3 groups: (a) Resourceful Adolescent Program—Adolescents (RAP—A), an 11-session school-based resilience building program, as part of the school curriculum; (b) Resourceful Adolescent Program—Family (RAP—F), the same program as in RAP—A, but in which each student's parents were also invited to participate in a 3-session parent program; and (c) Adolescent Watch, a comparison group in which adolescents simply completed the measures. The program was implemented with a high recruitment (88%), low attrition rate (5.8%), and satisfactory adherence to program protocol. Adolescents in either of the RAP programs reported significantly lower levels of depressive symptomatology and hopelessness at post-intervention and 10-month follow-up, compared with those in the comparison group. Adolescents also reported high satisfaction with the program. The study provides evidence for the efficacy of a school-based universal program designed to prevent depression in adolescence.

Studies have repeatedly demonstrated the high prevalence and harmful effects of depression during adolescence (Fleming & Offord, 1990; Keller, Lavori, Beardslee, Wunder, & Ryan, 1991; Peterson et al., 1993). Although there are variations across studies because of differences in methods employed, the point of prevalence of clinical depression for adolescents ranges between 0.4% to 8.4% (Anderson & McGee, 1994; Fleming & Offord, 1990). Lifetime prevalence rate for major depressive disorder in adolescents has been assessed to vary between 15% and 20% (Birmaher et al., 1996). Thus, a high priority for mental health research is to develop cost-effective prevention programs for depression in adolescents.

An important consideration in the area of prevention of adolescent depression is whether to implement universal programs that involve whole populations, selective programs that involve adolescents selected on the basis of a known population risk factor, or indicated programs that involve individuals showing mild to moderate signs or symptoms of a disorder (Mrazek & Haggerty, 1994). Each method has advantages and disadvantages. Universal prevention programs have been criticized for being expensive to mount and for involv-

ing implementation difficulties; however, they overcome other problems generally associated with indicated and selective prevention programs. The latter raise ethical problems of labeling and stigmatizing and the practical problem of low participation and retention rates. However, some innovative indicated and selective programs have recently been trialed. Indicated programs have typically targeted adolescents with elevated or subclinical levels of depressive symptoms (Clarke et al., 1995; Jaycox, Reivich, Gillham, & Seligman, 1994).

One indicated trial, the Pennsylvania Depression Program for adolescents ages 10 to 13 years (Jaycox et al., 1994), included three separate programs focusing on teaching (a) cognitive skills, (b) social problem-solving skills, and (c) a combination of cognitive and social problem-solving skills. Training in assertiveness, negotiation, and coping skills was also included. After finding no significant difference among the three intervention modalities, the groups were combined, resulting in a treatment sample of 69 participants and a wait-list control group of 74 participants. Significant improvements in depressive symptoms were obtained for the intervention group compared to controls at posttest, 6-month follow-up, and 2-year follow-up (Gillham, Reivich, Jaycox, & Seligman, 1995). This study indicates that prevention efforts to build resistance to depression seem promising during early adolescence. A limitation of the study was the possible

Requests for reprints should be sent to Ian Shochet, School of Applied Psychology, Griffith University, Queensland, Australia 4111. Email: i.shochet@mailbox.gu.edu.au

biasing effect of a self-selected sample in conjunction with the low initial recruitment rate (between 13% and 19%) and high attrition rate (30%).

In a second indicated prevention study using an adaptation of the treatment approach developed by Lewinsohn, Clarke, Hops, and Andrews (1990), Clarke et al. (1995) reported significant improvement in depression for an indicated intervention group compared to a wait-list group of 14- to 15-year-old adolescents. The program was more successful than Jaycox et al. (1994) at recruiting adolescents; however, it still only succeeded in engaging less than 50% of the adolescents identified as being at risk for depression. There was also a reasonably high attrition rate, particularly in the intervention group (21 out of 76). In another indicated trial, Hains and Ellmann (1994) reported positive results for their program, which consisted of problem solving, cognitive restructuring, and anxiety management, reducing depression scores in volunteer adolescents who had been classified as having high arousal levels. These authors also reported difficulty with possible self-selection bias.

These studies provide evidence for the usefulness of indicated prevention programs. They also highlight difficulties associated with recruitment and retention of adolescents. To the adolescent, such programs could be seen to put them at risk of being singled out from their peer group at an age when peer group acceptance is especially important. This problem might be substantially reduced if intervention programs for adolescent depression could be implemented routinely as part of the school curriculum, as either an alternative or complement to indicated programs. The question arises as to whether such a universal intervention is both feasible and effective given the “real-world” restrictions imposed by the school context.

The Resourceful Adolescent Program (RAP; Shochet, Holland, & Whitefield, 1997) was developed to meet this need. An 11-session program targeting Year 9 adolescents—Resourceful Adolescent Program—Adolescents (RAP–A)—was designed and piloted for implementation in the classroom as part of the school curriculum. The program content (outlined in the following) utilizes many of the cognitive-behavioral interventions of the successful indicated prevention trials discussed previously, while also addressing interpersonal and family risk and protective factors for depression. A three-session parallel program for parents was designed. The main goal of the parent program was to address the risk factor of severe family conflict and the protective factors of a responsive and warm parent-adolescent relationship (for research, reviews, and discussion on family factors and depression, see Kaslow, Deering, & Racusin, 1994; Lewinsohn, Clarke, Seeley, & Rohde, 1994; Shochet & Dadds, 1997).

The purpose of this study was to investigate the efficacy of a universal depression prevention program

within the constraints of the school environment. We needed to determine, in the first instance, whether it was feasible to effectively recruit and keep adolescents in a universal school-based program and whether the intervention would achieve adequate program fidelity in the classroom context. Next we needed to assess (a) whether such a program would be effective in reducing and preventing depressive symptoms and (b) whether the adolescents would approve of the program and consider it beneficial.

The study compared three prevention conditions: (a) RAP–A, in which only adolescents participated in the program; (b) Resourceful Adolescent Program–Family (RAP–F), in which adolescents participated in the same RAP program and their parents were invited to attend the RAP program for parents; and (c) Adolescent Watch (AW) comparison group. It was expected that both intervention groups (RAP–A and RAP–F) would be associated with fewer depressive symptoms at postintervention and 10-month follow-up compared to the nonintervention comparison group. It was also further expected that intervention effects would be greater for RAP–F than for RAP–A.

Method

Participants

The study was conducted at a large secondary school in Brisbane, Australia, an urban city of approximately 1.3 million people. Participating students were from families of low- to middle-class socioeconomic status, the majority of whom were Anglo-Saxon and Christian in origin. All students in Year 9 ($N = 295$) from two separate cohorts (1996 and 1997) were invited to participate in the study. After written parental consent was obtained, a total of 260 adolescents ages 12 to 15 years old (88.1% of the total Year 9 population; M age = 13.49 years, $SD = 0.54$) and comprising 121 boys (46.54%) and 139 girls (53.46%) remained. Class teachers identified students who were “integrated” (those previously noted by educational professionals as having developmental delays or learning disabilities and not intellectually able to understand the RAP session content). Although integrated students continued to participate in the study, they ($n = 4$) were not included in the total number of participants and no data were collected for purposes of this study. Participant numbers after referrals for individual counseling (discussed in the following), absentees, and attrition at postintervention and follow-up (broken down by cohort and intervention groups) are presented in Table 1.

Given that this was a universal program being conducted in one school, it was considered important to control for contamination from the active treatment to the comparison group. This was achieved by conduct-

Table 1. Participants in RAP–A or RAP–F and AW at the Various Program Stages

	1996 Cohort		1997 Cohort		Combined Total
	AW	RAP–A	RAP–F		
Total Sample Invited	144		151		295
Consented	126		134		260
Total Absent at Preintervention Assessment	1		4		5
Total Assessed at Preintervention	125		130		255
Excluded for Reasons of Developmental Disability	2		2		4
Excluded Because of Clinical Diagnosis at Preintervention and Referred for Individual Counseling	5		4		9
Total Usable Assessments at Preintervention	118		124		242
Total Assigned Into Intervention Groups	118	68	56		242
Total Assessed at Postintervention	107 ^a	65	56		228
Total Assessed at Follow-Up	90	53	51		194

Note: RAP–A = Resourceful Adolescent Program–Adolescent; RAP–F = Resourceful Adolescent Program–Families; AW = Adolescent Watch.

^aIncludes relocations and absentees.

ing the program with two different cohorts in the same school (1996, 1997) with the 1996 Year 9 students serving as the control group, and the 1997 students serving as the active intervention groups (RAP–A or RAP–F).

Measures

Three depression measures were administered to provide convergent data on adolescents' self-report of depressive symptomatology: the Child Depression Inventory (CDI; Kovacs, 1992), the Reynolds Adolescent Depression Scale (RADS; Reynolds, 1987), and the Beck Hopelessness Scale (BHS; Beck, Weissman, Lester, & Trexler, 1974). The CDI is a commonly used measure of childhood depression in which 27 items each provide three response options. In this study, and in accordance with the agreement established with the school, one item relating to suicidal ideation was omitted, leaving 26 items and a total possible score of 52. Kovacs described the CDI as having good internal consistency and moderate test–retest reliability in identifying depressive states.

The RADS (Reynolds, 1987) consists of 30 items on a 4-point scale, ranging from 1 (*almost never*) to 4 (*most of the time*), yielding a total possible score of 120. Again, in accordance with the agreement with the school, one item ("I feel like hurting myself") was omitted, leaving a total score possible of 116. The RADS is described as having high internal consistency and extensive evidence of content validity (Reynolds, 1987). The BHS (Beck et al., 1974) consists of 20 true/false items that measure the extent of an adolescent's negative attitudes about the future. Total scores range from 0 to 20 with higher scores indicating greater levels of hopelessness (Beck & Steer, 1988).

For ethical reasons, all students were actively monitored and those considered in need of clinical intervention were referred for additional assessment and

possible treatment. Attempts were made to balance participants' rights to privacy with a duty of care to intervene when students were considered at risk (Shochet & O'Gorman, 1995). To this end, clinicians scrutinized the questionnaires of adolescents who scored in the clinical range on the RADS (above 74). When there was also consistent evidence of depression on the basis of scores above the clinical range in either the CDI (20) and the BHS (9), the adolescents concerned were individually assessed. As a result, 15 adolescents were individually interviewed, and 9 were subsequently referred for additional assistance (we have no detailed information on the course of this assistance). These 9 participants continued in the universal program but were excluded from data analysis. (Table 1 reflects these exclusions).

Procedure

All adolescents were assessed on three occasions, preintervention, postintervention, and follow-up. Preintervention assessment occurred 2 weeks before start of the program. Postintervention was completed within 3 weeks of the end of the adolescent intervention program. The follow-up assessment occurred 10 months after postintervention. Both pre- and postintervention assessment comprised the same measures and procedures, and there were no differences in assessment conditions for the intervention groups versus controls. Participants in the comparison cohort (1996) were assessed at the same intervals, planned to coincide with those on the intervention cohort. There was a 5-month time span between the first week of preintervention and the end of the postintervention (i.e., 2 weeks from pretesting to start of the program, 11 weeks of the program including the midsemester vacation, and 3 weeks before completion of the postintervention). For practical reasons related to

school time-tabling, the AW pre- and postassessments occurred in the 5-month period during the latter half of the school year, as opposed to the first half of the school year for the RAP group.

The assignment to intervention versus control groups was done purely on the basis of the cohort. Given the universal nature of the program, there was a strong likelihood that the teenagers would share information about sessions with those peers who were not assigned to the program. In addition, school authorities were unhappy about the possibility of running the program with only half the cohort. Thus, the 1996 Year 9 cohort was assigned to the AW comparison group, which involved the normal school curriculum and three assessments, and the 1997 students served as the active intervention groups (RAP-A or RAP-F). Although there were no differences in cohorts on variables such as socioeconomic status and academic ability or gender composition, we were aware that this design could create a potential cohort effect confound. Probability of a confound from cohort effects was considered less likely, however, than the potential of a treatment contamination effect if each cohort was divided in half and then randomly allocated to intervention and comparison groups.

The RAP-A program was designed to be run in groups of approximately 8 to 12 participants per group. Preexisting classes of approximately 23 to 30 students each were divided into 2 or 3 groups and then randomly assigned one facilitator for each of 13 groups. There were no known differences in educational attainment levels or emotional problems among facilitation groups. Within the intervention cohort there was random assignment by facilitation group of participants into RAP-A (in which participants received the 11-week adolescent program only) and RAP-F (in which participants received the 11-week adolescent program and parents of participants were invited to the parent program). As a result of this process, 56 adolescents (that is, approximately half) were assigned to RAP-F, and their parents were invited to attend three parent evenings, which occurred within Weeks 4 and 9 of the adolescent program.

RAP-A program. The RAP program is a fully manualized group treatment available from the first author. Eleven group sessions were conducted weekly for between 40 and 50 min during school class time. Specifically, sessions were as follows: Session 1, establishing rapport; Session 2, affirmation of existing strengths; Sessions 3 and 4, promoting self-management and self-calming skills in the face of stress; Sessions 5 and 6, cognitive restructuring; Session 7, problem solving; Session 8, building and accessing psychological support networks; Sessions 9 and 10, interpersonal components designed to promote family

harmony and avoid escalation of conflict; Session 11, summary and termination. The content and process of each session of the program is specified in a Group Leader's Manual (Shochet et al., 1997). Complimentary participant workbooks (Shochet et al., 1997) were provided to each student in the program.

As is evident, Sessions 1 through 7 follow cognitive-behavioral therapy (CBT) approaches (e.g., Clarke et al., 1995). Sessions 8 to 10 address interpersonal risk and protective factors in adolescent development (Mufson, Moreau, Weissman, & Klerman, 1993). Interpersonal conflict is a well-established correlate and predictor of future depression (Hammen & Peters, 1978; Harter & Jackson, 1993; Lewinsohn et al., 1994). In the interpersonal components of the RAP program, groups discuss role transitions during adolescence (such as moving towards greater independence while maintaining positive relations with parents) and encourage use of skills to promote harmony, avoid escalation of conflict, and find ways of repairing conflictual situations. Interpersonal components also cover the skills of perspective taking, that is, the ability to understand another person's perspective, and skills to broaden social support.

RAP-F program. Three parent group sessions were conducted at 3-week intervals over the 11-week duration of the RAP-A program. Each session lasted 3 hr and was conducted at night to maximize parental attendance. Sessions were conducted with two group facilitators. The content of the sessions was as follows: Session 1, identification of existing parental strengths and identification and management of stress to enhance calm and effective parenting; Session 2, provision of information on normal adolescent development and strategies for promoting adolescent self-esteem and balancing independence and attachment issues; and Session 3, provision of strategies to promote family harmony and manage conflict. A number of strategies were used to maximize involvement of parents in RAP-F, including scheduling night meetings, individual invitations, reminder letters, and the provision of supper.

Facilitators. Group leaders for the adolescent program were psychologists with varied training and experience, ranging from experienced clinicians to graduate psychology students on practicum. Thirteen facilitators of the adolescent program were randomly assigned to the RAP-F or RAP-A conditions. The parent program was facilitated by trained psychologists that did not facilitate any of the adolescent programs. All facilitators attended 15 training sessions involving 25 hr of training. Training involved three elements: (a) two orientation workshops prior to commencement of

the RAP program; (b) 11 supervision and session-planning meetings of approximately 1½ hr each, held after every RAP session; and (c) one final debriefing meeting at conclusion of the RAP program. Training and supervision meetings were conducted by three experienced clinical psychologists involved in the program. Each facilitator was asked to present any group process and practical problems that arose in the session, and these were all discussed in turn (for example, how to encourage nonparticipating students, how to manage disruptive behavior).

To ensure the program was delivered within the parameters specified in the Group Leader's Manual, group facilitators self-monitored their implementation or were monitored by independent observers who were clinical supervisors from the RAP project, or both. The observers rotated across the various facilitation groups. Integrity checks required the observer to indicate ("yes-no" response) whether the facilitator had covered the content areas of the particular session as outlined on a session outline checklist. The number of content areas per session ranged from 5 to 11. Integrity checks by independent observers were carried out on 12.5% of RAP sessions, and self-ratings were completed on 57% of sessions.

Process evaluation. Process evaluation of RAP-A was completed by adolescents at the completion of their participation in the program. A 5-point Likert scale ranging from 1 (*no value*) to 5 (*great value*) was provided for participants to respond to eight questions. Questions probed perceived value of the program for adolescents in their everyday life, that is, whether it assisted their interactions with family and friends, improved their confidence levels, and improved their coping skills. A further six items asked participants to rate the usefulness of each of the RAP topics out of 10, and the overall usefulness of the program on a 5-point scale of 1 (*not at all useful*) to 5 (*very useful*).

Results

Statistical analyses revealed no significant differences among the intervention groups on variables of gender and any of the preintervention depression measures. As expected, there was a small age difference resulting from the fact that assessment of the AW group began in the second semester as opposed to the first semester for the RAP group. The AW group had a mean age of 13.7 years ($SD = 0.47$), whereas the RAP groups had a mean age of 13.2 years ($SD = 0.45$). However, the pattern of results did not change when age was or was not included as a covariate. Consequently, the results presented here do not involve the use of age as a covariate.

As can be seen in Table 1, there appeared to be a higher attrition rate in the AW groups compared to the RAP groups. This difference was not statistically significant, $\chi^2(2, N = 242) = 5.44, p = .067$. We nevertheless established that there were no significant differences in preintervention depression levels for participants who remained in the study compared to those who left the program or were absent at follow-up.

Program Integrity and Attendance

Mean percentage of content areas correctly covered by each facilitator for Sessions 2 to 10 in RAP-A was 90.5, 92.5, 93.5, 85.0, 85.6, 88.6, 88.9, 91.1, and 87.9, respectively. There was no difference between mean percentage of correct implementation according to self versus other ratings, $F(1, 141) = 1.27, ns$. Thus, facilitators achieved high levels of integrity in terms of program content. Given time-tabling constraints of the school, this adherence to program content was considered satisfactory. Attendance rates were accurately kept on 75% of participants in the program. All participating adolescents attended at least 9 of 11 sessions. Attrition rate for the program was low, as only 2% of students did not complete the program, and the majority of these students left the school during the term of the program. Attendance rate for the parents in RAP-F was very low. Only 36% of adolescents in RAP-F had a parent, or parents, who attended at least one of the parent sessions, and only 10% had a parent, or parents, who attended all three of the parent sessions.

Analyses of Intervention Effects

We conducted a repeated measures multivariate analysis of variance (MANOVA) with the between-participants factors of intervention group (RAP-A, RAP-F, and AW) and gender, and within-participants factor of time (preintervention, postintervention, and follow-up), on the depression measures (RADS, CDI, and BHS) considered together. The MANOVA showed main effects for intervention group, $F(6, 342) = 2.22, p < .05$; time, $F(6, 167) = 8.20, p < .001$; and gender, $F(3, 170) = 5.33, p < .01$, as well as time by intervention group interaction effects, $F(12, 336) = 1.81, p < .05$. There were no Gender \times Time or Gender \times Time \times Intervention Group effects. Univariate results of the MANOVA showed that the Intervention Group \times Time interaction effect was found on the CDI, $F(4, 344) = 3.15, p < .05$, and the BHS, $F(4, 344) = 3.37, p < .01$, but not on the RADS.

To examine the source of the Intervention Group \times Time interaction on the CDI and the BHS, three separate univariate repeated measures analyses of variance were conducted for the two measures, comparing each

intervention group against the other. For RAP–A versus AW, there was a significant Time \times Intervention Group interaction for the CDI, $F(2, 130) = 3.19, p < .05$, and the BHS, $F(2, 130) = 5.19, p < .01$. Similarly, for RAP–F versus AW, there was a significant Time \times Intervention Group interaction for both the CDI, $F(2, 130) = 4.85, p < .01$, and the BHS, $F(2, 129) = 5.18, p < .01$. For RAP–A versus RAP–F there were time effects only for CDI, $F(2, 99) = 8.76, p < .001$, and BHS, $F(2, 96) = 7.13, p < .001$, but no Time \times Intervention Group effects. Thus the two intervention groups differed from the comparison group but not from one another.

Paired (two-tailed) t tests were conducted to examine the time effect within the Time \times Intervention Group interaction effects on the CDI and BHS. From preintervention to postintervention, there were significant reductions for the RAP–F group on the BHS ($t = -2.09, p < .05$) and CDI ($t = -2.67, p < .01$) and for the RAP–A group on the CDI ($t = -2.35, p < .05$). For the AW group there was a significant increase on both the BHS ($t = 2.8, p < .01$) and on the CDI ($t = 2.07, p < .05$). From preintervention to follow-up, there were significant reductions for the RAP–F group on both the BHS ($t = -2.30, p < .05$) and CDI ($t = -2.15, p < .01$) as well as for the RAP–A group on the BHS ($t = -2.75, p < .01$) and CDI ($t = -2.59, p < .05$). For the AW group there were no significant differences in depression scores from preintervention to follow-up. From postintervention to follow-up there were no significant shifts with the exception of a significant decrease in CDI scores for the AW group ($t = -2.37, p < .05$). Thus these analyses indicate that active intervention groups had significant reductions in CDI and BHS levels largely through changes from pre- to postintervention and maintenance of these changes at 10-month follow-up (resulting in significant shifts from preintervention to follow-up). For the comparison group, there was an initial increase at postintervention. There was also a slight decrease at follow-up, resulting in no significant shifts from preintervention to follow-up (see Table 2).

In summary, results show statistically significant intervention effects for both the RAP intervention groups compared to the comparison group. They also show that the two RAP intervention groups did not differ from one another. Analyses presented in the following examine whether these effects are clinically significant ones.

Clinical and Health Promotion Significance

Examining clinical significance is considered standard practice in treatment outcome studies (Jacobson & Truax, 1991), but there is no predefined approach to examining this issue with respect to prevention and health promotion. Extrapolating from treatment outcome

studies, it is important to determine intervention effects on the number of people falling into clinical categories of depression, but also for those falling into the subclinical and healthy categories of depression at pretest, posttest, and follow-up. Thus we examined three questions: (a) How many adolescents fell into the clinical, subclinical, and healthy categories of depression for the RAP and AW groups at preintervention, postintervention, and follow-up? (b) How many adolescents who began with subclinical symptoms fell into the healthy, clinical, or subclinical ranges at posttest and follow-up? (c) How many adolescents who were healthy to begin with moved either into subclinical levels or clinical levels at postintervention and follow-up? Because of the problems of defining these categories based on one self-report score only, we aimed to reduce the possibility of false positives by defining the categories of healthy, subclinical, and clinical based on a convergence of the two measures (CDI and BHS). These measures were chosen as they measured different aspects of depression symptomology and were both sensitive to intervention effects as shown in the previous analyses. (The inclusion or exclusion of the RADS from the analyses presented here did not alter the pattern of results.) It was also important to establish criteria based on preexisting empirical data. Thus, criteria for judging the clinical status of each adolescent were as follows:

Clinical: On or above the clinical range on both the CDI and the BHS. We used the clinical cutoff of 20 and above as suggested in the test manual of the CDI (Kovacs, 1992). A score of 9 and above was used as the clinical cutoff on the BHS, which is considered indicative of moderate to severe hopelessness (Beck et al., 1974).

Subclinical: Above average symptoms, as defined in the test manuals, on both the CDI (12 and above) and the BHS (4 and above), but not in the clinical range on both measures.

Healthy: Not in the clinical range on either of the measures, and in the normal range on one or both of the measures.

For reasons of small sizes in the clinical subgrouping analyses and given that there were no differences in intervention effects for the two intervention groups, the RAP–A and RAP–F groups were combined for clinical significance analyses. Table 3 shows the preintervention, postintervention, and follow-up clinical status of the adolescents. At preintervention, there were no significant differences between the intervention groups with respect to the clinical and health status of participants, $\chi^2(2, N = 240) = 1.99, p > .37$. At postintervention, there were significant differences between the RAP and AW groups, $\chi^2(2, N = 220) = 11.89, p < .005$, with the

Table 2. Means and Standard Deviations for Depression Measures for the RAP-A, RAP-F, and AW Groups at Preintervention, Postintervention and Follow-Up

	RAP-A						RAP-F						AW					
	Pre		Post		Follow-Up		Pre		Post		Follow-Up		Pre		Post		Follow-Up	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RADS	56.09	9.62	51.75	12.48	50.57	11.42	55.06	10.12	49.57	11.19	49.59	11.16	55.42	12.09	51.46	13.59	49.80	14.11
CDI	7.25	4.96	5.82	4.80	5.74	4.80	7.92	5.45	5.84	4.42	6.37	4.95	8.50	6.81	8.90	7.87	7.82	7.14
BHS	3.54	3.17	2.94	2.52	2.62	2.33	3.64	2.98	2.56	2.78	2.73	2.58	3.28	3.05	3.71	3.65	3.68	3.87

Note: RAP-A = Resourceful Adolescent Program-Adolescent; RAP-F = Resourceful Adolescent Program-Families; AW = Adolescent Watch; RADS = Reynolds Adolescent Depression Scale; CDI = Child Depression Inventory; BHS = Beck Hopelessness Scale.

RAP group associated with lower rates of clinical and subclinical depression and higher rates in the healthy range compared to their comparison counterparts. At follow-up there were similar significant differences between the RAP and AW groups, $\chi^2(2, N = 192) = 10.63, p < .01$, with RAP groups showing lower rates of clinical and subclinical depression and higher rates in the healthy range. Next, we tracked movements into the three subcategories for the AW and RAP groups at postintervention and follow-up, based on their preintervention clinical status. Table 4 sets out the numbers and percentages in this regard.

First we examined differences between RAP and AW in the percentages of those who moved into subclinical and clinical categories for children who were healthy at preintervention. There was a nonsignificant trend toward fewer clinical cases in the RAP group at postintervention, $\chi^2(2, N = 177) = 5.09, p = .08$, and this difference was significant at follow-up, $\chi^2(2, N = 150) = 9.38, p < .01$. At follow-up, 10.1% of the “healthy” adolescents in the AW group moved into the subclinical or clinical categories compared to 1.2% in the RAP group.

Next, we analyzed data for adolescents who were in the subclinical category at preintervention. There were significant differences between the RAP and AW groups in clinical status at postintervention, $\chi^2(2, N = 40) = 7.25, p < .05$, and follow-up, $\chi^2(2, N = 37) = 6.04, p < .05$. None of the subclinical adolescents in the RAP program moved into the clinical range at postintervention or follow-up, compared to 10.5% at postintervention and 17.6% at follow-up for the AW group. In contrast, 71.4% of the subclinical adolescents in the RAP group moved into the healthy category at postintervention and 75% at follow-up. For subclinical children in the AW group, only 31.6% moved into the healthy category at postintervention and 41% at follow-up.

Table 4 also shows that the numbers of adolescents who were in the clinical range at preintervention were too small to compute statistical tests or make definitive conclusions.

Process Evaluation

Overall, participants rated the usefulness of the program as moderate to high. The overall rating for the program’s usefulness on a 5-point scale was $M = 3.89, SD = .91$. Means for the eight items measuring positive outcomes on a 5-point scale ranged from a low of 2.96 (“How much value were the sessions to the way you relate to family members?”) to 3.58 (“Would you recommend the program to your friends?”). There were no differences in ratings of outcomes between boys and girls. With reference to the six program content areas, the ranked mean ratings out of 10 were: problem solving, $M = 6.13, SD = 2.34$; self-talk, $M = 6.06, SD = 2.40$; resolving interpersonal problems, $M = 6.04, SD = 2.56$; recognizing personal strengths, $M = 5.95, SD = 2.25$; understanding body cues, $M = 5.75, SD = 2.32$; and support networks, $M = 5.69, SD = 2.35$. There were no gender differences in ratings of the content areas and no interaction between gender and content area. However, a main effect for content area, $F(5, 114) = 9.14, p < .001$, indicated that participants discriminated between usefulness of different content areas. A series of paired *t* tests comparing combinations of content areas revealed that “support networks” and “understanding body cues and stress management” were rated as less useful than other components, although it should be noted from the means that effect sizes were not substantial.

Discussion

The aim of this study was to investigate the implementation and efficacy of a universal school-based program to reduce depressive symptoms in adolescents, given problems of recruitment, retention, and possible stigmatization with indicated prevention programs. The first question was whether universal programs would result in more encouraging recruitment and retention rates for adolescents. The rate of recruitment of adolescents into this project was high. Eighty-eight

Table 3. Numbers and Percentage of Adolescents Falling Into the Healthy, Subclinical or Clinical Ranges at Preintervention, Postintervention and Follow-up for the RAP and AW Group

	Healthy		Subclinical		Clinical		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
RAP								
Preintervention	98	80.3	23	18.9	1	0.8	122	100
Postintervention	101	87.8	14	12.2	0	0.0	115	100
Follow-Up	95	93.1	6	5.9	1	1.0	102	100
AW								
Preintervention	91	77.1	23	19.5	4	3.4	118	100
Postintervention	76	72.4	22	21.0	7	6.6	105	100
Follow-Up	69	76.7	16	17.8	5	5.6	90	100

Note: RAP = Resourceful Adolescent Program; AW = Adolescent Watch.

Table 4. *Clinical Status of Adolescents at Postintervention and Follow-up, Subgrouped on Clinical Status at Preintervention in the RAP and AW Groups*

	Postintervention Clinical Status						Follow-Up Clinical Status						
	Healthy			Clinical			Healthy			Clinical			
	<i>n</i>	%		<i>n</i>	%		<i>n</i>	%		<i>n</i>	%		
Healthy													
RAP	86	92.5	7	7.5	0	0.0	80	98.8	0	0.0	1	1.2	
AW	70	83.3	11	13.1	3	3.6	62	89.9	7	10.1	0	0.0	
Subclinical													
RAP	15	71.4	6	28.6	0	0.0	15	75.0	5	25.0	0	0.0	
AW	6	31.6	11	57.9	2	10.5	7	41.2	7	41.2	3	17.6	
Clinical													
RAP	0	0.0	1	100.0	0	0.0	0	0.0	1	100.0	0	0.0	
AW	0	0.0	0	0.0	2	100.0	0	0.0	2	50.0	2	50.0	
Total	178		36		7		164		22		6		

Note: RAP = Resourceful Adolescent Program; AW = Adolescent Watch.

percent of those who had been invited participated in the program, with a 5.8% attrition rate from pre- to postintervention, and a 19.8% attrition rate from preintervention to follow-up, due largely to absences from school on the day of testing and relocations. The recruitment rate was higher than for studies utilizing an indicated approach to the prevention of adolescent depression (e.g., Clarke et al., 1995; Jaycox et al., 1994). The recruitment rate supported a major impetus of this study (i.e., to attempt a school-based universal depression prevention approach). One note of caution is that it is possible (although unlikely) that there was an overrepresentation of at-risk children in the 12% that declined to participate in this study.

The next question was whether the program would be implemented with sound program fidelity given general classroom and school constraints, such as competing scheduling priorities, unexpected interruptions, discipline problems, space constraints, and so on. The RAP-A program was implemented with good fidelity to the program protocol with an average of 89.3% session accuracy rate. (Time constraints appeared to account for any missed program content.)

With regard to the prevention effect, it was expected that both intervention groups (RAP-A and RAP-F) would be associated with fewer depressive symptoms at postintervention and follow-up compared to the nonintervention group. It was further expected that the predicted effects would be greater for RAP-F than for RAP-A. With regard to the former hypothesis, results can be summarized as follows: There were significant Intervention \times Time effects on depressive symptoms, confirming that both RAP intervention conditions (RAP-A and RAP-F) showed a significantly greater decrease in depressive symptoms as measured by the CDI and the BHS at postintervention and at follow-up approximately 10 months later when compared to the comparison group. There were, however, no significant differences between the RAP-A and RAP-F groups, which appears to disconfirm our expectation that the effects would be stronger for RAP-F than for RAP-A. Given the very low rate of parental participation, it is probable that the hypothesis of the increased effect for RAP-F could not be adequately tested. Nevertheless, from the outcome data and the low parental attendance rates, it would be reasonable to conclude the intervention effects in this study appear to be accounted for by the intervention targeted at adolescents.

One aspect of these findings that is difficult to explain is the fact that the intervention effect occurred on the CDI and the BHS only and not the RADS. The RADS had very significant time effects for all intervention groups, in which scores decreased across all three groups from pre- to post- to follow-up testing. The RADS differs from the CDI and the BHS in that it is a Likert scale rather than the endorsement of a particular proposition. The scale range and standard deviation is

much larger than the other two scales. It may be possible that this different form of questionnaire accounts for the difference between the RADS and the other two measures.

Nevertheless, the intervention programs seem to produce effects that have clinical and health promotion significance, when taking into account the scores on a convergence of depressive and hopelessness symptoms as measured by the CDI and BHS. Because the effects of the two intervention programs were similar, the two intervention groups were combined for statistical power purposes for analyses of clinical significance. Adolescents in the RAP program at postintervention and follow-up were found to have significantly lower rates of clinical and subclinical levels of depressive and hopelessness symptoms as defined by a convergence of clinical indicators on the CDI and BHS. It was very important to achieve positive results at follow-up and not just at postintervention. This would suggest that findings were more robust and more strongly indicative of a prevention effect, rather than some transient or time-specific effect.

Major beneficiaries of the program were those subclinical adolescents who began with moderately elevated depressive and hopelessness symptoms. Those in the RAP group were more likely to fall in the healthy range and less likely to fall in the clinical range compared to the comparison group. None of the subclinical adolescents in the RAP program moved into the clinical range at postintervention (5 months later) or at 10-month follow-up. For the AW subclinical adolescents, 17.6% fell in the clinical range at follow-up. From the health-promotion perspective, 75% of the subclinical adolescents in the RAP group fell into the healthy category at follow-up, compared to 41.2% of the AW group.

The universal nature of the program also appeared to be of benefit to adolescents who were initially considered healthy, and who would not normally be recruited into indicated or selective programs. That is, there was a significant difference in the clinical and health status between the RAP and AW groups for these initially healthy participants at follow-up. At follow-up, 10.1% of the healthy adolescents in the AW group moved into the subclinical category compared to none in the RAP group. Thus, although RAP was helpful for subclinical adolescents, the healthy group benefitted as well. The intervention effect coupled with the recruitment rate would appear to justify a universal intervention approach. However, two additional questions were whether the program was acceptable to the participants and whether the effects justify the costs.

The process evaluation provided an important source of validation from participants that the program was considered worthwhile. Overall, the program was positively endorsed, and all components of the program were ranked more positively than negatively. It was not

possible to determine which aspects of the program contributed to change in depression scores. Although there were differences in the process evaluations of various program components, differences were small and provided no clear evidence regarding the mechanisms of change. One of the key messages from the process evaluations, in spite of being positive, is the degree of variance in individual adolescents' reactions to different aspects of the program. Given the anonymous nature of the process evaluations, links between program preferences and outcome could not be examined. Future research should aim to link process evaluations with program outcomes and employ specific pre- and postmeasures of well-being such as attributional style, coping mechanisms, and interpersonal problem solving to help understand the mechanisms of change. Besides the issue of the mechanisms of change, as part of a comprehensive evaluation of a program, it is important to be able to demonstrate changes in the skills and domains that the program is targeting. Only then would we be able to confidently conclude that the program *per se* is bringing about the prevention effect.

It is important to do a cost-benefit analysis of the intervention approach. The direct costs of providing the intervention to approximately 125 adolescents consisted of the following: 13 intervention groups \times 1 hr sessions \times 1 group leader \times 11 sessions = 143 therapist hours. Given our failure to show additional benefits associated with the family component, it is not included in these calculations. From Table 3, the benefits can be summarized as the provision of RAP being associated with a reduction from 19.7% to 6.9% of children being in subclinical or clinical categories of depression at follow-up compared with stable rates in the comparison condition. Thus, approximately 13% or 13 participants per 100 can be argued to have either moved out of depression or to have been prevented from developing subclinical or clinical depression after receiving the intervention. This equates to 11 hr of therapist time per successful participant. The cost for training and supervision were not taken into account as these are implied costs in the delivery of all treatment.

Clearly, these costs are in a range roughly equivalent to the most efficient of treatments for depression. That is, most manualized treatments for depression in clinically referred adolescents are in the order of 12 to 20 sessions. Prevention and early intervention programs generally aim to be relatively less expensive than tertiary treatments and, in that sense, these results could be seen as less than optimal. However, this would miss the point of universal interventions for adolescents. The scarcity with which these young people refer themselves, or are referred by others, to clinical facilities and then receive adequate treatment has been documented (Keller et al., 1991). Most adolescents with depression problems simply do not receive treatment. Although it is certainly possible that there might have been at-risk

adolescents within the 12% that chose not to participate, this approach appears to be able to cast a wider net than indicated approaches.

Another important point is that approximately 10% of previously healthy adolescents in the AW group moved into clinical or subclinical ranges at follow-up. These are adolescents who might not normally be detected in an indicated model. Thus, given the available alternatives, the cost of 11 therapist hours to each successful participant seems realistic and perhaps the best available method of getting treatments out to adolescents at risk. It is difficult to estimate costs to the community of depression remaining untreated (or being left to develop); however, it is likely the costs are significantly higher than these direct therapist costs. Also, although the adolescents' perceived benefits from the program were not entered into a cost-benefit calculation, it should not be excluded from public health decisions on how best to prevent depression in young people. The question remains, however, as to whether the universal approach is the most cost-effective strategy to prevent depression. It is possible that a screening and early detection model might be more cost-beneficial and also overcome the problem of adolescents not self-referring.

This study possesses a methodological limitation that would suggest that we should interpret these encouraging results with some caution. We were able to control for school and socioeconomic status effects by conducting the research in one school, and the administration conditions of the two groups were identical, almost to the same classroom location. Nevertheless, running the program in one school created the methodological constraints of the study, namely that we had a potential cohort and time confound. The RAP versus AW was allocated on the basis of cohort and assessed at different times (although the time span was the same across the two groups). As mentioned, we had to choose between dividing each cohort in half and then randomly allocating to RAP and AW conditions in each cohort or allocating one cohort to the intervention condition and one to the comparison. We were concerned about treatment contamination problems, given that the program was run on a universal basis. If we divided a cohort in half, those in the program were almost certainly going to talk to their excluded peers about their program experiences. School authorities were also not in favor of an approach that provided different benefits to students within a particular cohort, as this would inevitably lead to complaints. Although the results mentioned previously may reflect some error associated with a cohort effect, effects at both postintervention and follow-up reduce (but do not exclude) an interpretation based on cohort or seasonal fluctuation.

A disappointing aspect of this study was the difficulty in recruiting parents of participating adolescents in the RAP-F component of the program, which made

it impossible to draw any definitive conclusions about the value of incorporating parents in prevention interventions for adolescent depression. As families play such an important role in adolescent depression, as both risk and protective factors, it would seem appropriate to put greater effort in future research into improving the technology for engagement of parents. One possible approach would be to develop a flexible learning model that does not rely on stressed parents having to devote precious time to attend after-hours workshops. One possibility could be the use of videotapes with telephone facilitation, although there is no assurance that this would increase parental recruitment.

Another limitation of the study is its reliance on self-report measures of depression. We used multiple measures of depression to overcome error associated with one self-report measure and ran the analyses to reflect convergence (e.g., MANOVA, the clinical significance data). Reynolds (1994) argued persuasively that self-report measures are among the most valid and reliable tools for assessing depression in adolescents. Nevertheless, this study would have been strengthened if the school allocated the time and we had the resources to conduct independent clinical assessments. Future research should include measures other than self-report.

Notwithstanding these concerns, results of this study are encouraging in terms of the value of running a cost-effective universal school-based program to prevent depression in adolescents. These findings suggest that future research in the area may prove beneficial, especially if it could be conducted across a number of schools and cohorts. One possible design is to have an intervention school and a matched control school in a particular cohort and then reverse intervention conditions in a subsequent cohort of the same schools. This would allow one to test or control for school and cohort effects.

It seems possible within the constraints of the school environment to design and implement classroom-based depression prevention programs, which appear effective both in the short term and at follow-up. Such an approach has many advantages over indicated or selective programs in terms of ease of recruitment, low attrition, and avoidance of stigma, and in the potential public health benefits. School-based universal approaches to preventing depression in adolescents could effectively cast a wider and important net, without problems of stigmatization and recruitment.

References

- Anderson, C. A., & McGee, R. (1994). Comorbidity of depression in children and adolescents. In W. M. Reynolds & H. F. Johnson (Eds.), *Handbook of depression in children and adolescents* (pp. 581–601). New York: Plenum.
- Beck, A. T., & Steer, R. A. (1988). *Beck Hopelessness Scale*. New York: Psychological Corporation.
- Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The Hopelessness Scale. *Journal of Consulting and Clinical Psychology, 42*, 861–865.
- Birmaher, B., Ryan, N. D., Williamson, D. E., Brent, D. A., Kaufman, J., Dahl, R. E., Perel, J., & Nelson, B. (1996). Childhood and adolescent depression: A review of the past 10 years. Part 1. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 1427–1439.
- Clarke, G. N., Hawkins, W., Murphy, M., Sheeber, L. B., Lewinsohn, P. M., & Seeley, J. R. (1995). Targeted prevention of unipolar depressive disorder in an at-risk sample of high school adolescents: A randomised trial of a group cognitive intervention. *Journal of the American Academy of Child Adolescent Psychiatry, 34*, 312–321.
- Fleming, J. E., & Offord, D. R. (1990). Epidemiology of childhood depressive disorders: A critical review. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*, 571–580.
- Gillham, J. E., Reivich, K. J., Jaycox, L. H., & Seligman, M. E. (1995). Prevention of depressive symptoms in school children: Two year follow-up. *Psychological Science, 6*, 343–350.
- Hains, A. A., & Ellman, S. W. (1994). Stress inoculation training as a preventative intervention for high school youths. *Journal of Cognitive Psychotherapy, 8*, 219–232.
- Hammen, C. L., & Peters, S. D. (1978). Interpersonal consequences of depression: Responses to men and women enacting a depressed role. *Journal of Abnormal Psychology, 87*, 322–332.
- Harter, S., & Jackson, B. K. (1993). Young adolescents' perceptions of the link between low self-worth and depressed affect. *Journal of Early Adolescence, 13*, 383–407.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 61*, 85–93.
- Jaycox, L. H., Reivich, K. J., Gillham, J., & Seligman, M. E. P. (1994). Prevention of depressive symptoms in school children. *Behaviour Research and Therapy, 32*, 801–816.
- Kaslow, N. J., Deering, C. G., & Racusin, G. R. (1994). Depressed children and their families. *Clinical Psychology Review, 14*, 39–59.
- Keller, M., Lavori, P. W., Beardslee, W. R., Wunder, J., & Ryan, N. (1991). Depression in children and adolescents: New data on “undertreatment” and a literature review on the efficacy of available treatments. *Journal of Affective Disorders, 21*, 163–171.
- Kovacs, M. (1992). *The Children's Depression Inventory manual*. North Tonawanda, NY: Multi-Health Systems.
- Lewinsohn, P. M., Clarke, G. N., Hops, H., & Andrews, J. (1990). Cognitive-behavioural treatment for depressed adolescents. *Behavior Therapy, 21*, 385–401.
- Lewinsohn, P. M., Clarke, G. N., Seeley, J. R., & Rohde, P. (1994). Major depression in community adolescents: Age at onset, episode duration, and time to recurrence. *Journal of the American Academy of Child and Adolescent Psychiatry, 33*, 714–722.
- Mrazek, P. J., & Haggerty, R. J. (Eds.). (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academy Press.
- Mufson, L., Moreau, D., Weissman, M. M., & Klerman, G. L. (1993). Interpersonal psychotherapy for adolescent depression. In G. L. Klerman & M. M. Weissman (Eds.), *New applications of interpersonal psychotherapy* (pp. 130–166). Washington, DC: American Psychiatric Association Press.
- Peterson, A. C., Compas, B. E., Brooks-Gunn, J., Stemmler, M., Ey, S., & Grant, K. E. (1993). Depression in adolescence. *American Psychologist, 48*, 155–168.
- Reynolds, W. M. (1987). *Reynolds Adolescent Depression Scale*. Odessa, FL: Psychological Assessment Resources.

PREVENTION OF ADOLESCENT DEPRESSION

- Reynolds, W. M. (1994). Assessment of depression in children and adolescents by self-report questionnaires. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 209–233). New York: Plenum.
- Shochet, I., & Dadds, M. (1997). Adolescent depression and the family: A paradox. *Clinical Child Psychology and Psychiatry*, 2, 307–312.
- Shochet, I., Holland, D., & Whitefield, K. (1997). *Resourceful adolescent program: Group leader's manual*. Brisbane, Australia: Griffith University.
- Shochet, I. M., & O'Gorman, J. (1995). Ethical issues in research on adolescent depression and suicidal behaviour. *Australian Psychologist*, 30, 183–187.
- Shochet, I., Whitefield, K., & Holland, D. (1997). *Resourceful adolescent program: Group leader's manual*. Brisbane, Australia: Griffith University.

Manuscript received March 21, 2000

Final revision received August 29, 2000

Copyright of Journal of Clinical Child Psychology is the property of Lawrence Erlbaum Associates and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.