Prevention & Treatment

Prevention & Treatment, Volume 2, Article 8, posted December 21, 1999

The Prevention of Depression and Anxiety

Martin E. P. Seligman, PhD, Peter Schulman, B.S., and Robert J. DeRubeis, PhD
Department of Psychology, University of Pennsylvania

Steven D. Hollon, PhD
Department of Psychology, Vanderbilt University

ABSTRACT

A brief and inexpensive cognitive–behavioral prevention program was given to university students at risk for depression. *At risk* was defined as being in the most pessimistic quarter of explanatory style. Two hundred thirty-one students were randomized into either an 8-week prevention workshop that met in groups of 10, once per week for 2 hr, or into an assessment-only control group. Participants were followed for 3 years and the authors report the preventive effects of the workshop on depression and anxiety. First, the workshop group had significantly fewer episodes of generalized anxiety disorder than the control group and showed a trend toward fewer major depressive episodes. The workshop group had significantly fewer moderate depressive episodes but no fewer severe depressive episodes. Second, the workshop group had significantly fewer depressive symptoms and anxiety symptoms than the control group, as measured by self-report but not by clinicians' ratings. Third, the workshop group had significantly greater improvements in explanatory style, hopelessness, and dysfunctional attitudes than the control group and these were significant mediators of depressive symptom prevention in the workshop group.

This research was supported by Grant MH19604 from the National Institute of Mental Health. Martin E. P. Seligman acknowledges the significant contribution of Mary Anne Layden to the authorship of the Attributional Style Questionnaire. The authors wish to thank the many research assistants, diagnostic interviewers, and trainers who contributed to this study:

Research Assistants: Amelia Balonek, Beth Berland, Bridget Blaney, Elizabeth Brannon, Steven Bromley, Muniya Choudhury, Deborah Clark, Kevin Colton, Robert Dauman, Linda Drummer, Amy Eisenberg, Tracy Epstein, Angela Estep, Michelle Fink, Jed Fishback, Amy Freundlich, Jennifer Friedman, Michael Friedman, Scott Glassman, Danielle Gray, Suzanne Johnson, Leslie Kaas, Beth Kaplan, Jeanine LaRouche, Nancy Lee, Andrea Levin, Patrick Lohmeyer, Irene Markman, Travis Marquette, Elizabeth McElwee, Jeanine McHugh, Janet Miller, Margaret Morris, Antonia Morocco, David Mraovitch, Stephanie Newman, Riaz Patel, Suzanne Pineles, Christopher Prokop, Veronica Rice, Daniel Richter, Andrew Rozmiarek, Yael Rubenstein, Jay

Salinger, Stacey Sarfatti, Jennifer Schecter, Sheena Sethi, Michael Singleton, Heather Smay, Audrey Smolkin, Todd Sweeney, Colleen Taylor, Caroline Tisot, Lisa Warren, Christopher Wasson, B.J. Zion, David Zlotchew, Julie Zures.

Workshop trainers: Norman Cotterell, Arthur Freeman, Robert DeRubeis, Jane Gillham, Melissa Hunt, Lisa Jaycox, Karen Reivich, Judy Saltzberg.

Diagnostic interviewers: John Abela, Carrie Bearden, Cynthia Brody, Gregory Buchanan, Esteban Cardemil, Tamar Chansky Stern, Timothy Davis, Martin Devine, Lisa Friedman Miller, Patricia Furlan, Carolyn Haskins, Nicholas Haslam, Derek Isaacowitz, Suzanne Johnson, Dora Klein, Karen Levinson, Lyn Mozley, Pagona Roussi, Jason Satterfield, Kelly Schmidt, Andrew Shatte, Karen Steinberg, Juliet Sternberg, Patricia White, Eric Zorrilla.

Correspondence concerning this article should be addressed to Martin E. P. Seligman, University of Pennsylvania, Department of Psychology, 3815 Walnut Street, Philadelphia, PA 19104.
E-mail: seligman@psych.upenn.edu

Depression affects almost 11 million individuals per year in the United States. Estimates of its monetary costs exceed $43 billion a year in treatment and lost productivity, a toll slightly larger than the costs of heart disease (Greenberg, Stiglin, Finkelstein, & Berndt, 1993). Cognitive therapy has proven roughly as effective in treating unipolar depression as antidepressant medication and produces marked relief in about 70% of patients (Beck, Hollon, Young, Bedrosian, & Budenz, 1985; Dobson, 1989; Hollon et al., 1992).

More striking is the finding that cognitive therapy prevents relapse after the termination of therapy and may have greater preventive effect than antidepressant drugs (Blackburn, Eunson, & Bishop, 1986; Evans et al., 1992; Hollon & Najavits, 1988; Shea et al., 1990; Simons, Murphy, Levine, & Wetzel, 1986). Unlike pharmacotherapy, cognitive therapy teaches a set of skills that can be applied long after the end of therapy. Because the majority of depressed individuals suffer multiple episodes, the capacity of an intervention to prevent future episodes is at least as important as its ability to treat the current episode.

Can the skills taught in cognitive therapy for depression and anxiety disorders also be taught preventively to individuals at risk but not currently suffering from these disorders? Despite the importance of prevention, there has been little research in this area. Jaycox, Reivich, Gillham, and Seligman (1994) found that a cognitive–behavioral workshop significantly prevented depressive symptoms compared with a control group at both postworkshop and 6-month follow-up among 10- to 13-year-olds identified as at risk for depression. Gillham, Reivich, Jaycox, and Seligman (1995) also found that this same sample still had significantly lowered depressive symptoms at 2-year follow-up (see Clarke, Hawkins, Murphy, Sheeber, Lewinsohn, & Seeley, 1995, for related cognitive interventions in adolescents).

The primary goal of this project was to explore a similar prevention workshop for college students. Our secondary goal was to look at mediators of any prevention effects. We looked at changes in explanatory style, hopelessness, self-esteem, and dysfunctional attitudes. Given the high comorbidity of depression and anxiety, our two main targets were the

prevention of new episodes of depressive disorder and its symptoms as well as anxiety disorders and its symptoms.

We used an 8-week cognitive–behavioral workshop designed to prevent depression and anxiety among individuals identified as at risk for depression. We now report the effects of the program for the first 3 years of follow-up, with frequent periodic diagnostic interviews and measurements.

Method

Participants

All 231 participants entered the study as first-year undergraduates at the University of Pennsylvania in 1991, 1992, and 1993. Forty eight percent of the participants were men and 52% were women. College students were used for two reasons: They are easy to recruit and track longitudinally, and they are at a formative age when the program could have long-lasting effects.

All participants were identified as being at risk for depression, with risk defined as scoring in the bottom quartile of the Attributional Style Questionnaire (ASQ: Seligman, Abramson, Semmel, & von Baeyer, 1979; Peterson et al., 1982). We have conducted six unpublished longitudinal studies with a total of 809 college students that showed that students who scored in the bottom quartile had between two and eight times more risk for later high levels of depressive symptoms (a Beck Depression Inventory [BDI] score of 16 or more).

We mailed ASQs to all new, incoming students in the summer before their first semester at the university. Students were eligible to participate if they scored in the most pessimistic quartile of the ASQ (using the full scale score, CPCN; see *Measures*) when they took it in the summer and if they met all of the following criteria at the preworkshop evaluation:

1. Not currently receiving psychotherapy or medication for psychological problems.
2. Still scoring in the most pessimistic quartile of the ASQ (CPCN $\leq$ 2.17) at the fall and spring preworkshop evaluation. This bottom quartile cutoff of 2.17 was determined by the distribution of ASQ scores for the university students tested in this research. We note that when we designed the project, we assumed that the ASQ score would remain stable between the time the ASQ was taken in the summer and the time of the preworkshop evaluation, a period of 2 to 6 months. In the first year of recruiting, therefore, we did not use the ASQ score to exclude participants during the preworkshop evaluation. At the end of the first year we discovered, to our surprise, that exactly half of the participants had ASQ scores above the bottom quartile at the preworkshop evaluation. We have excluded these participants from all analyses because they were not the intended target population for this research. We are not sure why the ASQ score improved for these students, but it is possible that the intervening major life change of leaving home and going to college was responsible. Regression to the mean is also a possible explanation for the rise in

CPCN scores.

3. Scoring 19 or less on the BDI, to exclude people likely to be in a major depressive episode before the program began.
4. Not meeting criteria for any of the following Axis I disorders: current major depression, past major depression with psychotic features (past major depression was not an exclusion criterion), past or current mania, current dysthymia, current cyclothymia, past or current psychosis, current suicide risk, past or current alcohol or substance dependence, current alcohol or substance abuse, current panic disorder, current panic disorder with agoraphobia, current agoraphobia without panic disorder, current obsessive–compulsive disorder, current somatization disorder, current hypochondriasis, current undifferentiated somatoform disorder, current anorexia or bulimia.
5. Reading and signing the voluntary consent form.

Measures: Diagnosis

The primary tool for screening at the preworkshop evaluation was the Structured Clinical Interview for the *DSM–III–R* (SCID; Spitzer & Williams, 1985). The SCID is a structured diagnostic interview designed to yield *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.; *DSM–III–R*) diagnoses across five axes. We only used Axis I for this study. At subsequent interviews, we used the Longitudinal Interval Follow Up Evaluation, modified to reflect *DSM–III–R* criteria (LIFE; Keller et al., 1987). All diagnostic interviews were audiotaped, unless the participant objected (less than 1% objected), and 10% of the tapes were randomly selected and checked by a second interviewer for diagnostic accuracy. The Research Diagnostic Criteria—Family History version, modified to reflect *DSM–III–R* criteria, was used to assess family history of psychopathology (Spitzer, Endicott, & Robins, 1978).

Measures: Symptoms

We used two additional measures of depression, the 21-item self-report BDI (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) and the 17-item clinician-rated Structured Interview Guide for the Hamilton Depression Rating Scale (SIGH-D; Hamilton, 1960; Williams, 1988). We also used two other measures of anxiety, the 21-item self-report Beck Anxiety Inventory (BAI: Beck, Epstein, Brown, & Steer, 1988) and the 14-item clinician-rated Structured Interview Guide for the Hamilton Anxiety Rating Scale (SIGH-A: Hamilton, 1959).

Measures: Mediators

We measured four possible mediators of prevention: explanatory style, dysfunctional attitudes, hopelessness, and self-esteem. The ASQ is a self-report instrument that yields scores for explanatory style for bad events and for good events using three causal dimensions: internal versus external, stable versus unstable, and global versus specific causes. The ASQ presents 12 hypothetical events, half good and half bad, and the test-taker

is asked to write down the one major cause of each event and then rate the cause along a 7-point continuum for each of the three causal dimensions. We report results for the ASQ variable that represents all of the questions on the instrument—the explanatory style for good events minus the explanatory style for bad events (CPCN). This is consistently the most valid correlate and predictor of depressive deficits (Peterson & Seligman, 1984).

The Dysfunctional Attitudes Scale (DAS; Weissman & Beck, 1978) is a 100-item self-report measure designed to assess the extent to which an individual endorses general attitudes and underlying assumptions hypothesized by cognitive theory to be associated with depression, a set of constructs similar to Albert Ellis's concept of irrational beliefs. Participants endorse statements on a 7-point Likert scale. The DAS has been shown to distinguish clinically depressed from nondepressed psychiatric patients and from normal controls (Hollon, Kendall, & Lumry, 1986).

The Hopelessness Scale (HS; Beck, Weissman, Lester, & Trexler, 1974) is a 20-item true–false self-report instrument that measures general expectations about the future and the degree of general pessimism. Research has found that the HS discriminates clinically depressed from nondepressed psychiatric patients (Brown & Beck, 1989).

The Self-Concept Test is a 25-item self-report instrument that assesses self-esteem by having test-takers compare themselves to others on a variety of physical, intellectual, and personality characteristics (Beck, Steer, Brown, & Epstein, 1990).

Diagnostic Interviewers

Diagnostic interviewers administered all diagnostic measures as well as the Hamilton Depression Rating Scale and the Hamilton Anxiety Rating Scale. All interviewers were graduate students at area doctoral clinical psychology programs and had experience and training in diagnosis and assessment prior to joining the study. Interviewers went through about 20 hr of in-class training, which consisted of lectures, discussion, and role-playing, as well as about 10 hr of homework. To be accepted as interviewers, trainees had to pass a written test of the diagnostic criteria for the SCID and have satisfactory reliability on several videotaped and audiotaped interviews. In addition, throughout the entire diagnostic interviewing phase, all interviewers met in a group with the supervisor every 2 or 3 weeks to discuss questions, and the supervisor did a reliability check using an audiotaped interview that was later discussed with the interviewer. The purpose of these meetings was to prevent the interviewers from drifting from the appropriate interview techniques.

The Prevention Workshop

The prevention training workshop consisted of 16 hr of meetings—1 2-hr meeting per week for 8 weeks—plus between-meeting homework. The workshop was given to 10 to 12 freshmen participants per group by a trainer and a cotrainer. In addition, one of the trainers met individually with each participant on six occasions—once in the beginning of the workshop, once in the middle of workshop, 1 month after the workshop, 3 months after the

workshop ended, once in the fall term of their sophomore year, and once in the spring term of their sophomore year. The purpose of these individualized meetings was to review the skills the participants learned in the workshop and discuss any questions they had about applying these skills to their lives.

All trainers were trained cognitive therapists who currently worked or previously had worked at Aaron Beck's Center for Cognitive Therapy, with experience ranging from 2 years to 30 years. All cotrainers were either these same cognitive therapists or doctoral students enrolled in the clinical psychology program at the University of Pennsylvania.

The workshop taught a range of cognitive–behavioral techniques based largely on Beck and colleagues' cognitive therapy for depression (Beck, 1964, 1967, 1976; Beck, Rush, Shaw, & Emery, 1979; Hollon & Beck, 1979; Seligman, 1991). We decided not to include any explicit discussion of explanatory style in the workshop because this would have undermined our ability to reliably measure this construct as a mediator.

The workshop included the following topics: (a) the cognitive theory of change (the relationship between thoughts, feelings, and behaviors); (b) identifying automatic negative thoughts and underlying beliefs; (c) marshaling evidence to question and dispute automatic negative thoughts and irrational beliefs (empirical hypothesis testing); (d) replacing automatic negative thoughts with more constructive interpretations, beliefs, and behaviors (generating alternatives, thought stopping, distraction techniques); (e) behavioral activation strategies (graded task breakdown, time management, antiprocrastination techniques, creative problem solving, assertiveness training), (f) interpersonal skills (active listening, taking each other’s perspectives, controlling emotions, passive vs. assertive vs. aggressive behaviors), (g) stress management (relaxation training), and (h) generalizing these coping skills to new and relevant situations.

We developed a highly detailed and scripted manual to standardize the delivery of the workshop (Gillham et al., 1991). The format of the workshop meetings consisted of rapport-building, some lecturing and audiovisual presentations, participant role-playing, games and activities, group discussion and homework reviews, and the use of a detailed participant's notebook with homework and written materials that reviewed the major points of the workshop. Attendance at the workshop averaged about 85%.

Procedure

We recruited a total of 231 participants during the summers of 1991, 1992, and 1993. We mailed ASQs in late May of each year to all incoming first-year undergraduates who lived in the United States. In the letter accompanying the ASQ, we told them that if they completed the ASQ, we might contact them later to see if they were interested in participating further in our research. We told them that if they chose to participate further, they could receive $400. All prospective participants were told: "The purpose of this study is to track the attitudes and expectations of students during their college experience." We sent a second ASQ mailing in early July to those who did not respond to the first mailing. See Table 1 for details.

Table 1
Participant Recruiting at Each Phase

	1991/1992	1992/1993	1993/1994
Total incoming class	2,316	2,221	2,453
First ASQ mailing (U.S. only)	2,052	2,007	2,248
Second ASQ mailing	831	903	—
Total response rate	66%[a]	63%[b]	49%[c]
Bottom quartile ASQ	331	337	294
Preworkshop evaluations	261	318	148
Participants	83	104	44

Note. ASQ = Attribution Style Questionnaire
[a]1,346/2,052. [b]1,261/2,007. [c]1,104/2,248

In August, we called the incoming students to see if they were interested in participating and discussed what participation would entail. At this point, all prospective participants were told: "The purpose of this research is to evaluate a seminar designed to teach first-year students how to deal with the challenges and take advantage of the opportunities that new students at Penn face." If they were interested, we scheduled them to come in for a preworkshop evaluation in the beginning of the fall or spring semester to determine their eligibility for the study. About 85% of those we reached were interested in participating, and we scheduled these individuals to come in for evaluations. This was a much higher percentage than we expected. We believe this high percentage (85%) was due to several reasons: (a) Participants were offered $400 for completing all phases of the research; (b) we sent participants an official-looking letter on University of Pennsylvania letterhead soon after they received their acceptance letter, at a time when they were likely to be receptive and compliant; and (c) participants were told that the results of this research would likely benefit many individuals.

At the preworkshop evaluation, research assistants first administered the ASQ and BDI to determine eligibility to participate further. If students still scored in the bottom quartile of the ASQ, had a BDI score of 19 or less, and signed the voluntary consent form, diagnostic interviewers then conducted a SCID to determine eligibility. If students did not meet criteria for any of the Axis I disorders listed above, they went on to fill out the questionnaire battery. They were then randomized into either the prevention workshop or a no-workshop control group. Participants were stratified on the basis of four variables to ensure that the groups were balanced—sex, ASQ score (above vs. below the median score of .83 for this bottom quartile on the ASQ group), BDI score (above vs. below the median score of 6), and whether they had a past history of depression. Seventeen of the 225 participants had past depression at the preworkshop diagnostic interview, based on the SCID. They were randomized into each of the two conditions.

At the end of the semester, about 2 months after the preworkshop evaluation, all participants returned for a postworkshop evaluation similar to the preworkshop evaluation. Diagnostic interviewers conducted the LIFE to collect information on Axis I disorders, after which participants took a questionnaire battery to measure possible mediators. We administered follow-up evaluations similar in format to the postworkshop evaluation once each semester for six semesters over the next 3-year period for a total of six follow-up evaluations. Prior to each diagnostic interview, research assistants asked participants not to tell the interviewer which group they were in. This was to ensure that interviewers were unaware of condition so that their interviews and diagnoses would be unbiased by our research goals.

In addition to the twice-a-year evaluations, we also mailed BDIs to all participants each month over the 3-year follow-up for a total of 36 measurements. Interviewers used these monthly BDIs in the follow-up evaluations to prompt their line of questioning and to prompt participants' memories about the period of time since the prior evaluation.

We report all follow-up data below for the first 3 years. The attrition rate was only 3.5% (8/231), which was a much lower rate than we expected. Three of the dropouts were in the workshop group and 5 were in the control group. All 8 dropouts said that their reason for discontinuing participation was that they were too busy. A review of the tables indicates that the missing data often exceeds 3.5%. At any given measurement, some participants were temporarily unreachable or unavailable, and this percentage ranged from 0% to 20%. The most common reasons given were that they were studying abroad for one or two semesters or that they were too busy to participate at that particular measurement.

We believe there are at least three reasons for the low attrition rate: (a) Participants received $400 for participating in all phases of the research; (b) all contacts with participants were carefully scripted, and research staff were trained to be professional and courteous at all times; and (c) participants were occasionally reminded that the results of this research would likely benefit many individuals.

Dependent Symptom Measures

We created two measures to represent all the different time measurements for each participant for each measure (postworkshop and follow-ups 1 through 6): *follow-up mean* and *follow-up maximum*. The mean variable was the mean of all available time measures from postworkshop through Follow-Up 6 for each participant. The maximum variable represented the most symptomatic level of a given variable for all available time measures from postworkshop through Follow-Up 6 for each participant.

Statistical Procedures

We conducted survival analyses to determine differences in the time to first episode between the intervention and control groups, covarying preworkshop symptoms (Greenhouse, Stangl, & Bromberg, 1989). Survival analyses not only take into account whether an episode has occurred but also reflect the time elapsed until the first episode.

Survival analyses compute estimates of the survival distribution for right-censored data. An observation is censored when a participant is no longer eligible to have an episode, such as when the participant withdraws prematurely from participation or completes the follow-up period without experiencing an episode. The Mantel–Cox test, also known as the log-rank test, was used to test for group differences (Mantel, 1966). We used SAS Proc Lifereg to perform this analysis (Cox & Oakes, 1984; SAS Institute Inc., 1985).

To determine the effect of time on the workshop effect, we modeled a growth curve using SAS Proc Mixed (Littel, Milliken, Stroup, & Wolfinger, 1996). We modeled the outcomes from postworkshop through the 3 years of follow-up, adjusting for preworkshop levels and for condition. Time was treated as a categorical response in which no specific pattern was assumed. We used the following model with the categorical outcome measures, the continuous outcome measures, and the cognitive measures:

$$\text{Outcome} = \text{Outcome}_{\text{pre}} \ \ \text{Time} \ \ \text{Condition} \ \ \text{Time*Condition}$$

We used the following technique to determine mediation of the workshop effects (Sobel, 1982). Path *a* is the treatment effect on cognition change. Path *b* is the cognition effect on symptom change. We computed the coefficient estimate and standard error for path *ab*.

The coefficient estimate for *ab* = coefficient of path *a* multiplied by the coefficient of path *b*.
Standard error (S) = square root of $(b^2 s_a^2 + a^2 s_b^2 + s_a^2 s_b^2)$
t statistic = coefficient *ab* divided by the standard error. The *t* statistic can be used to obtain the *p* value.

This procedure allows us to test the hypothesis that the workshop led to cognition change which in turn led to symptom change.

We report one-tailed *p* values when there is a clear, unidirectional hypothesis that the workshop group will do better than the control group. Two-tailed *p* values are noted when used. Also, in all analyses, symptoms at preworkshop were covaried to control for initial level of symptoms. In other words, for all the analyses of covariance (ANCOVAs), all postworkshop and follow-up measures are residualized using the preworkshop measure. The effect sizes in Tables 3 and 4 are equal to the difference between the means divided by the standard deviation of the control group.

Results

There were three main findings. First, the workshop group had significantly fewer episodes of generalized anxiety disorder than the control group and showed a trend toward fewer major depressive episodes. Second, the workshop group had significantly fewer depressive symptoms and anxiety symptoms than the control group, as measured by self-report but not by clinicians' ratings. Third, the workshop group had significantly greater improvements in explanatory style, hopelessness, and dysfunctional attitudes compared with the control group, and these were significant mediators of depressive symptom prevention in the

workshop group.

Baseline

There were no significant differences between the workshop and control groups at the preworkshop evaluation for any of the variables.

Blindness and Interrater Reliability

To determine if diagnostic interviewers were blind as to which condition participants were in, following each interview, we had them guess which condition the participant was in. At all the evaluations but one, a chi-square analysis revealed that diagnostic interviewers were unable to accurately guess which condition the participants were in. Diagnostic interviewers were accurate in 54% of their guesses at the postworkshop interview, $\chi^2(1, N = 183) = 1.2$, *ns,*, in 46% of their guesses at Follow-Up 1, $\chi^2(1, N = 188) = 1.3$, *ns,* in 52% of their guesses at Follow-Up 2, $\chi^2(1, N = 193) = 0.3$, *ns,* in 48% of their guesses at Follow-Up 3, $\chi^2(1, N = 190) = 0.3$, *ns,* in 59% of their guesses at Follow-Up 4, $\chi^2(1, N = 181) = 4.8$, $p < .03$, in 51% of their guesses at Follow-Up 5, $\chi^2(1, N = 185) = 0.0$, *ns,* and in 50% of their guesses at Follow-Up 6, $\chi^2(1, N = 191) = 0.0$, *ns*. Sample sizes for these analyses are less than the total sample size because interviewers sometimes forgot to guess which group the participant was in.

Interrater reliability was estimated from an intraclass correlation coefficient, using the random effects estimate, with two raters pooled (ICC[2,2]; Shrout & Fleiss, 1979.] For the 114 tapes assessed, we found intraclass correlations of .86 for the Hamilton Depression Rating Scale, .89 for the Hamilton Anxiety Rating Scale, and .84 for the LIFE MDD rating.

Episodes of Depression

We performed survival analyses to determine differences in the time to first episode between the intervention and control groups, covarying preworkshop symptoms. Survival analyses not only take into account whether an episode has occurred, but also reflect the time elapsed until the first episode. This analysis is described in *Statistical Procedures*. The LIFE rates major depression on a scale of 1 through 6. A rating of 3 is considered *moderate depression*, a rating of 4 is *marked*, 5 is *definite*, and 6 is *severe*. Few participants had a rating of 5 ($n = 23$) or 6 ($n = 4$) for major depression.

Defining a depressive episode as 3 or more (moderate and above), we found that 40% (42/106) of the workshop group had a depressive episode, as had 48% (57/119) of the control group, $\chi^2(1, N = 225) = 1.9$, $p < .08$. Participants in the control group were 1.4 times more likely to have a depressive episode than participants in the workshop group. We used an odds ratio calculation, which is the probability of an episode in the control group divided by the probability of an episode in the workshop group. See Table 2.

Table 2

Number of Participants With Depressive or Anxiety Episodes at Each Rating

	LIFE MDD rating				
	None(1-2)	Moderate(3)	Marked(4)	Definite(5)	Severe(6)
Workshop	64	15	13	12	2
Control	62	28	16	11	2

	LIFE GAD rating		
	None(0)	Moderate(1)	Definite(2)
Workshop	91	15	0
Control	94	24	1

Note. LIFE = Longitudinal Follow Evaluation; MDD = Major Depressive Disorder; GAD = General Anxiety Disorder

Most of the episodes were moderate rather than severe, so we analyzed each level of depression separately. When we defined an episode as equal to a rating of 3 (instead of greater than or equal to 3) and compared this with the number of participants who did not have a depressive episode (a rating of 1 or 2), we found that 19% (15/79) of the workshop group had a moderate episode of depression and 31% (28/90) of the control group did, $\chi^2(1, N = 169) = 3.4$, $p < .03$. (The denominator is the number of participants with a moderate episode plus the number of participants with a 1 or 2.) When a depressive episode was defined as equal to 4 (marked), 17% (13/77) of the workshop group were depressed, as were 21% (16/78) of the control group, $\chi^2(1, N = 155) = 0.4$, *ns*. When a depressive episode was defined as equal to 5 (definite), 16% (12/76) of the workshop group were depressed, as were 15% (11/73) of the control group, $\chi^2(1, N = 149) = 0.0$, *ns*. When a depressive episode was defined as equal to 6 (severe), 3% (2/66) of the workshop group were depressed, and 3% (2/64) of the control group were depressed, $\chi^2(1, N = 130) = 0.0$, *ns*.

Episodes of Anxiety

The LIFE assesses generalized anxiety disorder (GAD) on a 1 through 3 scale. A rating of 2 is considered moderate and a rating of 3 is definite GAD. For this analysis, we used a cutoff of 2 or more because only 1 participant had a rating of 3. For the diagnosis of generalized anxiety disorder, 15 out of 106 of the workshop group were at the moderate level or above (14%) versus 25 out of 119 of the controls (21%), $\chi^2(1, N = 225) = 3.0$, $p < .04$. Using the odds ratio calculation, participants in the control group were 1.6 times more likely to have an anxiety episode than participants in the workshop group. See Table 2.

Time Effect and Gender Effect on Episodes

The effect of the workshop on preventing depression or anxiety episodes neither increased

nor decreased over time. To determine this, we used the growth curve model discussed in *Statistical Procedures*. The Time × Condition interaction effect was not significant, indicating that the effect of the workshop did not change over time.

Gender seemed to be a moderator of the prevention effect on depressive episodes, but not on GAD episodes. In a survival analysis of women, defining a depressive episode as 3 or more (moderate and above), we found that 24 out of 52 of the workshop group had a depressive episode (46%), as had 39 out of 65 of the control group (60%), $\chi^2(1, N = 117) = 2.5$, $p < .06$. In a survival analysis of men, again defining a depressive episode as 3 or more (moderate and above), we found that 18 out of 54 of the workshop group had a depressive episode (33%), as did 18 out of 54 of the control group (33%), $\chi^2(1, N = 108) = 0.0$, *ns*.

In a survival analysis of women, defining a GAD episode as 2 or more (moderate and above), we found that 12 out of 52 of the workshop group had a GAD episode (23%), and 18 out of 65 of the control group had an episode (28%), $\chi^2(1, N = 117) = 0.8$, *ns*. In a survival analysis of men, defining a GAD episode as 2 or more, we found that 3 out of 54 of the workshop group had a GAD episode (6%), as had 7 out of 54 of the control group (13%), $\chi^2(1, N = 108) = 1.9$, $p < .08$. In other words, there was a trend for men to benefit more from the workshop than women, in terms of GAD episode prevention.

Compared with men, women had a significantly higher incidence of both depressive episodes and of GAD episodes (defining episodes as moderate and above). Disregarding which condition they were in, we found that women had a higher incidence of depressive episodes than men (54% vs. 33%), $\chi^2(1, N = 225) = 9.6$, $p < .002$, as well as a higher incidence of GAD (26% vs. 9%), $\chi^2(1, N = 225) = 10.3$, $p < .001$.

Symptom Levels of Depression and Anxiety

Using the mean of each participant's post and follow-up measures, ANCOVA analyses found that the workshop group had significantly fewer symptoms of depression than the control group, covarying depression symptoms at preworkshop, using the biannual BDI measure ($n = 225$), $F(1, 223) = 4.4$, $p < .02$, and the monthly BDI measure ($n = 225$), $F(1, 223) = 4.3$, $p < .02$. The Hamilton measure of depression was not significant. The workshop group also had significantly fewer symptoms of anxiety than the control group, covarying anxiety symptoms at preworkshop, using the BAI ($n = 225$), $F(1, 223) = 4.4$, $p < .02$. The Hamilton anxiety measure was not significant. See Table 2 for details.

Using the maximum (worst level) of each participant's post and follow-up measures, an ANCOVA analysis revealed that the workshop group had significantly fewer symptoms of depression than the control group, covarying depression symptoms at preworkshop, using the biannual BDI measure ($n = 225$), $F(1, 223) = 2.9$, $p < .04$, and the monthly BDI measure ($n = 225$), $F(1, 223) = 3.1$, $p < .04$. The Hamilton measure of depression was not significant. The workshop group also had significantly fewer symptoms of anxiety than the control group, covarying anxiety symptoms at preworkshop, using the BAI ($n = 225$), $F(1, 223) = 4.4$, $p < .02$. The Hamilton measure of anxiety was not significant. The prevention effect

sizes on all the symptom measures were small. See Table 3 for details.

Table 3

Adjusted Means and ANCOVAs for Symptom Measures

Variable	Workshop group			Control group			ANCOVA [a]			
	n	*M*	*SD*	*n*	*M*	*SD*	*F*	*df*	*p*	Effect size [b]
Hamilton Depression Rating Scale [c]										
Preworkshop [d]	106	3.2	2.8	119	3.0	2.9	—	—	—	—
Postworkshop	106	1.7	2.0	119	2.4	2.5	5.9	1, 223	.008	.28
Follow-Up 1	106	2.2	2.6	117	2.3	2.4	0.0	1, 219	*ns*	.04
Follow-Up 2	105	2.1	2.4	117	2.3	2.5	0.4	1, 221	*ns*	.08
Follow-Up 3	101	2.0	2.2	114	2.4	3.3	1.3	1, 220	*ns*	.12
Follow-Up 4	96	2.6	2.9	107	2.2	2.7	1.0	1, 213	*ns*	−.15
Follow-Up 5	96	2.6	3.8	109	2.2	2.4	0.6	1, 201	*ns*	−.17
Follow-Up 6	99	2.1	2.4	109	2.4	2.7	1.0	1, 206	*ns*	.11
Mean [e]	106	2.2	1.6	119	2.4	2.0	0.9	1, 223	*ns*	.10
Maximum [f]	106	5.4	3.8	119	5.3	3.4	0.1	1, 223	*ns*	−.03
Biannual BDI										
Preworkshop [d]	106	7.3	5.1	119	7.3	5.0	—	—	—	—
Postworkshop	106	3.2	3.0	119	4.3	3.4	7.3	1, 223	.004	.32
Follow-Up 1	106	2.9	3.9	117	3.3	3.3	0.8	1, 221	*ns*	.12
Follow-Up 2	103	2.2	3.3	116	3.1	3.8	3.9	1, 217	.03	.24
Follow-Up 3	97	2.0	2.8	110	3.4	3.9	7.8	1, 205	.003	.36
Follow-Up 4	84	2.7	3.4	99	2.4	3.8	0.2	1, 181	*ns*	−.08
Follow-Up 5	85	2.4	4.2	102	2.8	3.5	0.6	1, 185	*ns*	.11
Follow-Up 6	100	1.7	2.5	110	2.6	3.6	4.1	1, 208	.02	.25
Mean [e]	106	2.4	2.5	119	3.2	2.8	4.4	1, 223	.02	.29
Maximum [f]	106	5.6	4.7	119	6.6	4.4	2.9	1, 223	.04	.23
Monthly BDI [c, f]										
Mean [e]	106	2.6	2.6	119	3.4	3.2	4.3	1, 223	.02	.25
Maximum [f]	106	8.7	6.2	119	10.3	7.3	3.1	1, 223	.04	.22
Hamilton Anxiety Rating Scale [c]										
Preworkshop [d]	106	3.8	3.3	119	3.3	2.8	—	—	—	—
Postworkshop	106	2.1	2.2	119	3.1	2.6	9.6	1, 223	.001	.38

Follow-Up 1	106	2.4	2.8	115	2.6	2.7	0.2	1, 219	*ns*	.07
Follow-Up 2	105	2.6	2.5	117	3.0	2.8	0.8	1, 220	*ns*	.14
Follow-Up 3	101	2.2	2.5	114	2.8	3.6	1.7	1, 213	.09	.17
Follow-Up 4	96	2.5	2.9	108	2.4	2.7	0.2	1, 202	*ns*	–.04
Follow-Up 5	96	2.9	3.6	109	2.7	3.1	0.1	1, 203	*ns*	–.06
Follow-Up 6	100	2.4	2.9	109	2.5	2.6	0.1	1, 207	*ns*	.04
Mean [e]	106	2.5	1.9	119	2.7	2.0	1.2	1, 223	*ns*	.10
Maximum [f]	106	5.3	3.9	119	6.0	3.8	1.5	1, 223	*ns*	.18
Beck Anxiety Inventory [c]										
Preworkshop [d]	106	4.8	4.8	119	4.5	4.4	—	—	—	—
Postworkshop	106	2.6	2.9	119	2.8	2.3	0.6	1, 223	*ns*	.09
Follow-Up 1	106	2.6	3.6	117	2.8	3.1	0.1	1, 221	*ns*	.06
Follow-Up 2	103	2.5	2.8	116	3.2	3.8	2.6	1, 217	.05	.18
Follow-Up 3	97	1.5	2.2	110	3.0	3.9	10.8	1, 205	.001	.38
Follow-Up 4	84	2.6	2.9	99	2.7	4.3	0.1	1, 181	*ns*	.02
Follow-Up 5	85	2.0	2.8	102	2.5	3.0	1.5	1, 185	*ns*	.16
Follow-Up 6	100	1.7	2.7	110	2.3	2.9	1.9	1, 208	.09	.21
Mean [e]	106	2.2	1.9	119	2.8	2.5	4.4	1, 223	.02	.24
Maximum [f]	106	5.1	4.1	119	6.2	4.9	3.3	1, 223	.03	.22

Note. ANCOVA = analysis of covariance.

[a] Model: $\text{Symptom}_{\text{follow-up}} = \text{Symptom}_{\text{pre}} + \text{Condition}$

[b] Difference between the means divided by the standard deviation of the control group.

[c] Higher score is worse.

[d] All preworkshop variables are actual data rather than adjusted data.

[e] Mean of all postworkshop and follow-up measures for each participant.

[f] Maximum (worst level) of all postworkshop and follow-up measures for each participant.

[g] For 36 months after workshop.

To determine whether the prevention effect was greater for participants with more symptoms at intake, we used the following ANCOVA model with each of the symptom measures:

$$SX_{\text{Follow-Up}} = SX_{\text{Pre}} \ \ \text{Condition} \ \ SX_{\text{Pre}}*\text{Condition}$$

We found two baseline interaction effects. First, participants with more preworkshop depressive symptoms on the BDI had a greater prevention effect than those with fewer preworkshop depressive symptoms, using the mean monthly BDI measure ($n = 225$), $F(1, 221) = 4.4$, $p < .04$, two-tailed. Second, participants with more preworkshop anxiety symptoms on the Hamilton anxiety measure had a greater prevention effect than those with

measure (n = 225), $F(1, 221) = 4.1$, $p < .04$, two-tailed.

Time Effect and Gender Effect on Symptoms

We found no effect of time on the prevention effect, for any of the symptom measures. To determine this, we used the growth curve model discussed in *Statistical Procedures*. The Time × Condition interaction effect was not significant, indicating that the effect of the workshop did not change over time.

We also found that gender did not have a moderating effect for any of the symptom measures. In other words, the workshop effect was not moderated by gender for the symptom measures. The following ANCOVA model was used to determine this:

$SX_{Follow\text{-}Up} = SX_{Pre}$ Condition Gender Condition*Gender

Prediction of Grades

We collected cumulative grade point average (GPA) for the participants' 4 years at the University of Pennsylvania. A t test analysis revealed that the workshop group did not differ significantly from the control group in cumulative GPA, $t(196) = 1.1$, *ns,* n = 198, though there was a trend in the expected direction. Explanatory style at postworkshop, however, significantly predicted cumulative GPA, covarying explanatory style at preworkshop, $F(1, 196) = 3.9$, $p < .03$, n = 198. Explanatory style at preworkshop did not predict cumulative GPA, $F(1, 196) = 0.0$, *ns,* n = 198.

The University of Pennsylvania currently uses a measure called the predictive index (PI) that is used in admissions decisions and predicts how well a student will perform academically. The PI is a weighted average of the applicant's high school grades, high school rank, and Scholastic Aptitude Test (SAT) scores. The PI significantly predicted cumulative GPA, $F(1, 195) = 30.7$, $p < .0001$, n = 197. Explanatory style at postworkshop predicted cumulative GPA when the PI was covaried out, $F(1, 194) = 2.7$, $p < .05$, n = 197.

Cognitive Measures

At postworkshop and all follow-ups, an ANCOVA analysis revealed that the workshop group had significantly better levels of explanatory style and dysfunctional attitudes than the control group, covarying preworkshop levels of each of these measures. The hopelessness measure was significantly better for the workshop group than for the control group at postworkshop, Follow-Up 1, and Follow-Up 3, but there was no difference at Follow-Up 5. There were no significant differences between the groups on the self-esteem measure. The workshop effect sizes on all the cognitive measures were small. See Table 4 for details.

The effect of the workshop on the cognitive measures neither increased nor decreased over time. To determine this, we used the growth curve model discussed in *Statistical*

Procedures. The Time × Condition interaction effect was not significant, indicating that the effect of the workshop did not change over time.

Table 4

Adjusted Means and ANCOVAs for Possible Cognitive Mediators

Variable	Workshop group			Control group			ANCOVA			
	M	*SD*	*n*	*M*	*SD*	*n*	*F*	*df*	*p*	Effect size[a]
Attributional Style Questionnaire[b]										
Preworkshop	0.0	1.7	106	0.1	2.0	119	—	—	—	—
Postworkshop	1.4	2.2	106	0.8	1.9	119	4.6	1, 223	.02	.32
Follow-Up 1	1.5	2.5	106	1.0	2.0	117	3.4	1, 221	.03	.20
Follow-Up 3	1.8	2.5	96	0.8	2.7	110	7.5	1, 204	.003	.37
Follow-Up 5	2.0	2.7	85	1.0	2.5	102	6.7	1, 185	.005	.40
Hopelessness Scale[c]										
Preworkshop evaluation	3.9	3.0	106	3.7	3.5	119	—	—	—	—
Postworkshop evaluation	3.0	2.3	106	3.9	2.3	119	7.9	1, 223	.003	.39
Follow-Up 1	3.1	2.2	106	3.8	2.6	117	4.0	1, 221	.02	.27
Follow-Up 3	3.3	2.6	97	4.1	2.9	110	4.8	1, 218	.02	.28
Follow-Up 5	3.4	2.6	85	3.5	2.7	101	0.0	1, 214	*ns*	.04
Beck Self-Concept Scale[d]										
Preworkshop evaluation	79.7	6.1	106	80.5	7.2	119	—	—	—	—
Postworkshop evaluation	79.9	4.6	106	80.1	3.8	119	-0.2	1, 223	*ns*	–.05
Follow-Up 1	80.6	4.8	106	80.4	4.2	117	0.1	1, 221	*ns*	.05
Follow-Up 3	80.9	5.2	97	80.0	5.1	110	1.9	1, 205	.09	.18
Follow-Up 5	81.0	5.7	85	80.7	5.3	101	0.1	1, 184	*ns*	.06
Dysfunctional Attitudes Scale[c]										

Preworkshop evaluation	324.2	50.8	106	323.9	55.2	119	—	—	—	—
Postworkshop evaluation	302.9	35.8	106	320.3	31.7	119	14.9	1, 223	.0001	.55
Follow-Up 1	297.0	35.8	106	318.4	36.1	117	19.7	1, 221	.0001	.59
Follow-Up 3	299.1	42.9	97	316.0	46.6	110	7.3	1, 205	.004	.36
Follow-Up 5	291.4	42.3	85	301.5	41.8	102	2.7	1, 185	.05	.24

Note. ANCOVA (analysis of covariance) Model: $\text{Cognition}_{\text{follow up}} = \text{Cognition}_{\text{pre}} + \text{Condition}$. All preworkshop means are actual data rather than adjusted data.

[a] Difference between the means divided by the standard deviation of the control group.

[b] Measures explanatory style for good minus bad events; higher score is more optimistic.

[c] Higher score is worse.

[d] Measures self-esteem; higher score is better.

Mediation

We used the Sobel (1982) formulas to determine mediation (discussed in *Statistical Procedures*). We found that explanatory style, hopelessness, and dysfunctional attitudes were all significant mediators of depressive symptoms at postworkshop and follow-up. Only dysfunctional attitudes was a significant mediator of anxiety symptoms at postworkshop and follow-up. See Table 5 for statistics.

Table 5

Mediation Calculations Using Sobel Formulas

	Path *a*	Path *b*	Coefficient *ab/s*	*t*	*p*
		ASQ (n = 225)			
BDI *M*	.590	.222	.131/.081	1.6	.05
BAI *M*	.590	.126	.074/.058	1.3	.10
		HS (n = 225)			
BDI *M*	.863	.270	.233/.107	2.2	.02
BAI *M*	.863	.092	.079/.065	1.2	.10
		DAS (n = 225)			
BDI *M*	17.34	.011	.191/.108	1.8	.04
BAI *M*	17.34	.014	.243/.102	2.4	.01

Note. All symptom mean variables represent the mean of each participant's post and follow-up measures, residualized with the preworkshop measure. The cognition measure represents the measure at postworkshop, residualized with the preworkshop measure. ANCOVA = analysis of covariance. ASQ = Attributional Style Questionnaire; BDI = Beck Depression Inventory; BAI = Beck Anxiety Inventory; HS = Hopelessness Scale; DAS = Dysfunctional Attitudes Scale.

Discussion

There were three major findings. First, on episodes, the workshop group had significantly fewer episodes of generalized anxiety disorder than the control group and showed a trend toward fewer major depressive episodes. The workshop group had significantly fewer moderate depressive episodes but no fewer severe depressive episodes. Women benefited more from the workshop than men.

Second, on symptoms, the workshop group had significantly fewer depressive symptoms and anxiety symptoms than the control group, as measured by self-report but not by clinicians' ratings.

Third, on mediators, the workshop group had significantly greater improvements in the cognitive measures of explanatory style, hopelessness, and dysfunctional attitudes than the control group. The workshop effect sizes on the symptom measures were small. These improvements in explanatory style, hopelessness, and dysfunctional attitudes were all significant mediators of the prevention of depressive symptoms in the workshop group. Only dysfunctional attitudes was a significant mediator of the prevention of anxiety symptoms. There was no difference in self-esteem between the groups and self-esteem was therefore not a mediator. The workshop effect sizes on the cognitive mediators were small.

There are five issues we now discuss about these prevention findings: (a) the size of the effects, (b) the consistency of the effects across different measures, (c) whether prevention was found or was merely maintenance of gain, (d) why women benefited more than men, and (e) the lack of a placebo control.

Locus of Prevention Effects: Moderate Versus Severe Depression

The effects we found were modest in size. Let us consider the prevention of depressive episodes first. We found that there was a nonsignificant trend ($p < .08$) with the workshop group having fewer episodes than the control group when both moderate and more severe episodes were combined. Forty-eight percent of the control group had such an episode, and only 40% of the workshop group. When we analyzed moderate episodes of depression alone, the prevention effect was significant ($p < .03$).

The percentage of episodes, when moderate through severe are counted, is quite large absolutely, but it should be noted that these students were chosen because they were at high

risk for depression by virtue of pessimism scores on the ASQ in the worst quartile. We have found in extensive pilot work that such students have two to eight times the risk for developing moderate to severe symptoms of depression relative to the rest of the student population. The overall rate of more severe episodes was lower than we expected: 4/225 (1.8%) students had a severe episode and another 23/225 (10.1%) had a definite episode of depression, or a total of 12% at the most intense end of depressive episodes.

How might we account for the existence of only a trend with all depressive episodes, but a significant effect of prevention when only moderate episodes are counted? There are three salient possibilities.

1. We might have merely capitalized on chance by dividing up episodes this way. So it may be that there is no actual preventive effect on depression. We doubt this explanation because the pattern of results, including self-reported depressive symptoms, self-reported anxiety symptoms, mediators, and anxiety episodes is coherent, suggesting an underlying preventive effect.
2. Moderate depression is psychological and preventable, but severe depression is biological and not preventable by psychological interventions. This is an intriguing speculation and we have no data that bear directly on it. The speculation is consistent with the finding of preventive effects on the self-reported depression symptoms but not on clinicians' ratings. The clinicians' ratings may be more loaded on severe, "melancholic" features with self-report more loaded on cognitive and emotional features of depression.
3. Finally, it may be that the sample size of severe depression was too small for our design to have enough power to detect a difference. We lean toward this explanation. Our power estimates were based on an educated guess that 30% of our control group would display a definite or severe episode of depression over the 3-year follow-up, but we found only a 10.9% (13/119) rate in the control group. When an episode is defined to include moderate episodes or worse, thereby increasing the percentage of participants with an episode to more than 40%, a prevention effect seems to emerge. We do not, however, have a ready explanation of why only 11% of the at-risk control group showed a definite or severe episode of depression over the 3-year follow-up, rather than the 30% rate we had expected. But we speculate that students are more likely to survive college with moderate than with severe depression.

Consistency of Dependent Variables

The preventive effects were not uniform across all dependent variables. Preventive effects were found for the self-reported symptom measures but not for the clinicians' ratings of symptoms. There are three plausible accounts of this inconsistency:

1. We capitalized on chance and were able to come up with some measures that showed prevention, but there is no true underlying prevention effect. We doubt this explanation, again because the pattern of results, including self-reported depressive and anxiety symptoms, mediators, and clinician ratings of depressive and anxiety

episodes is coherent, suggesting an underlying preventive effect.

2. Cognitive interventions, such as our program, teach people to talk "nondepressively," but do not change underlying depression. They change the cognitive and emotional symptoms, which are the material of self-report but leave the somatic and behavioral symptoms untouched. This criticism is often launched against psychological therapies for depression generally, and against cognitive therapy in particular.

   We analyzed cognitive (cog), fatigue (fat), emotional (em), and somatic (som) symptoms on the BDI separately (White, 1997). We found high correlations among all four sets of symptoms. For all 225 participants in the study, we found the following correlations among the different sets of symptoms, averaging their BDIs across the 36 months: r(cog-som) = .54, r(cog-em) = .77, r(cog-fat) = .68, r(som-em) = .74, r(som-fat) = .73, r(em-fat) = .76, all ps < .0001.

   On the other hand, using a repeated measures analysis of variance across all 36 months and classifying BDI items into these four symptom clusters, we found a Group X Cluster interaction at the $p < .02$ level, in addition to the significant main effects of prevention on each cluster. The interaction appears to be based primarily on greater prevention effects for cognitive than for somatic symptoms (White, 1997). It should be noted, however, that somatic symptoms were quite rare in this study. We therefore doubt that the prevention workshop affected only the cognitive and emotional "talking" symptoms and not the behavioral–somatic symptoms, but it is possible that our prevention program prevents more cognitive symptoms than somatic symptoms.

3. Consistency of all variables is a function of effect size. We found modest preventive effects overall. When these are further disaggregated by looking at each dependent variable separately, some will be significant and others not, although the pattern will be coherent. If the power of our design had been greater, with a larger sample size, we speculate that we would have found preventive effects on every variable. This also seems consistent with the greater effect on cognitive and emotional variables, which are more prominent in less severe depressions, and which are the kind of depressive episodes that predominated in this study. In general, we found that the BDI effects were clearer than the Hamilton Depression Rating Scale effects. One possible explanation is that the Hamilton Depression Rating Scale is insensitive in the relatively nondepressed range, while the BDI remains sensitive.

Prevention or Maintenance

The pattern of the preventive effect over time was steady. Another way of saying this is that the direct effect of the workshop may have been relieving symptoms that existed at the time of the workshop. This relief is then maintained across the 3-year follow-up. Such a pattern might lead some to ask whether this was prevention at all or just therapy plus maintenance.

We believe this is a semantic confusion in need of some conceptual clarification, which we will now try to provide. There are three possible patterns of relief that a preventive intervention might bring about, where preventive intervention is defined as an intervention on a relatively asymptomatic, but at-risk, group with the result that the at-risk group then

stays relatively asymptomatic.

Steady. In this pattern, relief occurs during the intervention itself, and remains throughout the follow-up. If the effect immediately postintervention is covaried out, no additional preventive effect occurs during follow-up. This is the pattern we found. We believe that underlying this pattern is the fact that the workshop taught a set of cognitive and emotional skills that do not wane or wax over time. They are simply deployed steadily when problems arise over the course of life.

Waning. In this pattern, relief occurs during the intervention itself, and wanes across the follow-up. This is the almost universal effect that therapy has on preventing symptoms far into the future. What often underlies this pattern is the learning of ways of dealing with problems, in which the learning gets weaker over time or the problems change over time, making what is learned less applicable.

Waxing. This is the rarest pattern, the sleeper effect. In this pattern, little relief occurs during the intervention itself, but during follow-up, effects increase over time. More formally, if the effect immediately postintervention is covaried out, additional preventive effects will be found during follow-up. It is instructive that Gillham et al. (1995) found such a pattern using a similar workshop, but designed for 10- to 12-year-old children at risk for depression, who then went into puberty during 2 years of follow-up. We believe that in this case the children learned the same set of skills that the university students in our workshop did, but the 11-year-old children had less occasion than our 20-year-old students did to deploy these skills, until puberty hit. The numerous problems that puberty brought provided new and more frequent opportunities for the use of antidepressant cognitive and behavioral skills. The preventive effect therefore waxed in time for the children both because they faced more opportunities to do better than the control group over time and/or because the skills got stronger as they deployed them more successfully.

Women Versus Men

Why did the women benefit more from the workshop than the men, in terms of depressive episode prevention? One possibility is that women with depression benefit more in general from cognitively oriented programs like ours. We found the opposite effect in our programs with 10- to 12-year-old children, with boys benefiting more than girls (Jaycox et al. 1994; Gillham et al., 1995). Another possibility is that of a floor effect: Women had significantly higher rates of depression than men in our study, and so the preventive effects of the program would only become visible with a larger number of men over a longer time span.

No Placebo Control

We finally mention the major shortcoming of our experimental design—the absence of a placebo control. Factors such as social cohesion, expectation of gain, and attention from older authority figures might contribute to the prevention effect. Since this is one of the first attempts to prevent depression on a large and inexpensive scale, we deemed it cost effective

to first see if we could get any prevention at all. If so, future research would then attempt to dismantle the active from the inactive components of the workshop using placebo designs.

In conclusion, it is important to note that fewer than 20% of individuals with an affective disorder seek treatment (Shapiro et al., 1984). If episodes of depression and anxiety can be prevented at an early stage of life by interventions in schools and colleges, this could have long-lasting and beneficial effects on the mental health of the nation. Our ultimate goal is to provide middle schools, high schools, and colleges with easily implemented prevention programs in order to immunize vulnerable individuals against depression and anxiety. We hope to enable widespread delivery of this program by training educators, school counselors, and social workers to carry out cognitive–behavioral prevention of depression and anxiety.

References

Beck, A. T. (1964). Thinking and depression: II. Theory and therapy. *Archives of General Psychiatry, 10,* 561–571.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects.* New York: Harper and Row.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders.* New York: International Universities Press.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56,* 893–897.

Beck, A. T., Hollon, S. D., Young, J. E., Bedrosian, R. C., & Budenz, D. (1985). Treatment of depression with cognitive therapy and amitriptyline. *Archives of General Psychiatry, 42,* 142–148.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression: A treatment manual.* New York: Guilford Press.

Beck, A. T., Steer, R., Brown, G., & Epstein, N. (1990). The Beck Self-Concept Test. *Archives of General Psychiatry, 2,* 191–197.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. E., & Erbaugh, J. K. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571.

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.

Blackburn, I. M., Eunson, K. M., & Bishop, S. (1986). A two-year naturalistic follow-up of depressed patients treated with cognitive therapy, pharmacotherapy and a combination of both. *Journal of Affective Disorders, 10,* 67–75.

Brown, G., & Beck, A. T. (1989). The role of imperatives in psychopathology: A reply to Ellis. *Cognitive Therapy and Research, 13,* 315–322.

Clarke, G., 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 randomized trial of group cognitive intervention. *Journal of the American Academy of Child & Adolescent Psychiatry, 34,* 312–321.

Cox, D. R., & Oakes, D. (1984). *Analysis of survival data.* London: Chapman and Hall.

Dobson, K. S. (1989). A meta-analysis of the efficacy of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 57,* 414–419.
Evans, M. D., Hollon, S. D., DeRubeis, R. J., Piasecki, J. M., Grove, W. M., Garvey, M. J., & Tuason, V. B. (1992). Differential relapse following cognitive therapy and pharmacotherapy for depression. *Archives of General Psychiatry, 49,* 802–808.
Gillham, J., Reivich, K., Jaycox, L., & Seligman, M. E. P. (1995). Prevention of depressive symptoms in school children: Two year follow up. *Psychological Science, 6*, 343–351.
Gillham, J., Jaycox, L., Reivich, K., Hollon, S. D., Freeman, A., DeRubeis, R. J., & Seligman, M. E. P. (1991). *The Apex Project manual for group leaders.* Unpublished manuscript.
Greenberg, P. E., Stiglin, L. E., Finkelstein, S. N., & Berndt, E. R. (1993). The economic burden of depression in 1990. *Journal of Clinical Psychiatry, 54*, 405–426.
Greenhouse, J. B., Stangl, D., & Bromberg, J. (1989). An introduction to survival analysis: Statistical methods for analysis of clinical trial data. *Journal of Consulting and Clinical Psychology, 57,* 536–544.
Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology, 32,* 50–55.
Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry, 23,* 56–62.
Hollon, S. D. & Beck, A. T. (1979). Cognitive therapy of depression. In P. C. Kendall & S. D. Hollon (Eds.), *Cognitive–behavioral interventions: Theory, research, and procedures* (pp. 153–203). New York: Academic Press.

Hollon, S. D., DeRubeis, R. J., Evans, M. D., Wiemer, M. J., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1992). Cognitive therapy and pharmacotherapy for Depression: Singly and in combination. *Archives of General Psychiatry, 49,* 774–781.
Hollon, S. D., Kendall, P. E., & Lumry, A. (1986). Specificity of depressotypic cognitions in clinical depression. *Journal of Abnormal Psychology, 95,* 52–59.
Hollon, S. D., & Najavits, L. (1988). Review of empirical studies of cognitive therapy. In A. J. Frances & R. E. Hales (Eds.), *American Psychiatric Press Review of Psychiatry* (Vol. 7, pp. 643–666). Washington, DC: American Psychiatric Press.
Jaycox, L., Reivich, K., Gillham, J., & Seligman, M. E. P. (1994). Prevention of depressive symptoms in school children. *Behaviour Research and Therapy, 32,* 801–816.
Littel, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). *SAS system for mixed models.* Cary, NC: SAS Institute Inc.
Keller, M. B., Lavori, P. W., Friedman, B., Nielsen, E., Endicott, J., McDonald-Scott, P., & Andreasen, N. C. (1987). The longitudinal interval follow-up evaluation: A comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry, 44,* 540–548.
Mantel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports, 50,* 163–170.
Peterson, C., & Seligman, M. E. P. (1984). Causal explanations as a risk factor for depression: Theory and evidence. *Psychological Review, 91,* 347–374.
Peterson, C., Semmel, A., von Baeyer, C., Abramson, L. T., Metalsky, G. I., & Seligman, M. E. P. (1982). The Attributional Style Questionnaire. *Cognitive Therapy and Research, 6,*

SAS Institute Inc. (1985). *SAS User's Guide: Statistics, Version 5 Edition.* Cary, NC: SAS Institute Inc.
Seligman, M. E. P. (1991). *Learned optimism.* New York: Knopf.

Seligman, M. E. P., Abramson, L. Y., Semmel, A., & von Baeyer, C. (1979). Depressive attributional style. *Journal of Abnormal Psychology, 88,* 242–247.
Shapiro, S., Skinner, E. A., Kessler, L. G., Von Korff, M., German, P. S., Tischler, G. L., Leaf, P. J., Benham, L., Cottler, L., & Regier, D. A. (1984). Utilization of health and mental health services: Three epidemiological catchment area sites. *Archives of General Psychiatry, 41,* 971–978.
Shea, M. T., Elkin, I., Imber, S. D., Sotsky, S. M., Watkins, J., Collins, J. F., Pilkonis, P. A., Leber, W. R., Krupnick, J., Dolan, R. T., & Parloff, M. B. (1990). Course of depressive symptoms over follow-up: Findings from the National Institute of Mental Health treatment of depression collaborative research program. *Archives of General Psychiatry, 49,* 782–787.
Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428.
Simons, A. D., Murphy, G. E., Levine, J. E., & Wetzel, R. D. (1986). Cognitive therapy and pharmacotherapy for depression: Sustained improvement over one year. *Archives of General Psychiatry, 43,* 43–49.
Sobel, M. E.(1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological methodology* (pp. 290–312). Washington, DC : Jossey-Bass.
Spitzer, R. L., Endicott, J., & Robins, E. (1978). *Research diagnostic criteria.* New York: Biometrics Research, Evaluation Section, New York State Psychiatric Institute.
Spitzer, R. L., & Williams, J. B. W. (1985). *Structured clinical interview for DSM–III–R-Patient version (SCID-P).* New York: Biometrics Research Department, New York State Psychiatric Institute.
Weissman, A. N., & Beck, A. T. (1978). *Development and validation of the Dysfunctional Attitudes Scale: A preliminary investigation.* Paper presented at the meeting of the American Educational Research Association, Toronto, Canada.
White, P. T. (1997). *Cognitive symptoms: Do they predict depression?* Unpublished doctoral dissertation, University of Pennsylvania.
Williams, J. B. W. (1988). A structured interview guide for the Hamilton Depression Rating Scale. *Archives of General Psychiatry, 45,* 742–747.