

Efficacy of Self-Administered Treatments for Depression and Anxiety

Marisa Menchola and Hal S. Arkowitz
University of Arizona

Brian L. Burke
Fort Lewis College

Self-administered treatments (SATs) are widely used by the general public and mental health professionals. Previous reviews of the efficacy of SATs have included under this category interventions for nonclinical problems, group interventions, and interventions involving significant amounts of therapist contact. The efficacy of SATs for clinical levels of depression and anxiety with minimal therapeutic contact was examined by meta-analyzing 24 studies. The results show large effects for SATs when compared with no-treatment control groups ($d = 1.00$). However, unlike previous meta-analyses that found nonsignificant differences between SATs and therapist-administered treatments, in this sample SATs resulted in significantly poorer outcomes ($d = -0.31$). Some differences in effect size were observed between the clinical targets of depression and anxiety. However, there were high correlations between clinical target, methodological quality of the study, and amount of contact. This makes it impossible to determine whether the observed differences could be explained by the nature of the disorders, methodological quality, or the amount of contact with a member of the research team. The implications of the findings for the clinical use of SATs and for future research are discussed.

Keywords: self-administered treatments, self-help, depression, anxiety, meta-analysis

Supplemental data: <http://dx.doi.org/10.1037/0735-7028.38.4.421supp>

The use of self-administered treatments (SATs) as an adjunct to psychotherapy is common among practicing psychologists, with over 80% of psychotherapists recommending some form of self-help intervention to their patients (Norcross et al., 2000). However, many questions remain regarding the use of SATs in clinical practice and the ways in which their effects may be maximized. In particular, knowledge is limited regarding the usefulness of SATs

for clients with clinically significant psychological problems with minimal therapist contact.

SATs have been defined as “media-based treatment approaches (book, manual, audiotape, videotape, or some combination) that are used largely by an individual independent of a helping professional.” (Gould & Clum, 1993, p. 170). They have been used for decades in areas as diverse as obesity (e.g., Balch & Ross, 1974), sexual dysfunction (Van Lenkvel, 1998), alcohol abuse (e.g., Heather, Whitton, & Robertson, 1986), and memory skills (Scogin, Stordant, & Lott, 1985). The present article is a meta-analysis of outcome studies on the efficacy of SATs for depression and anxiety in clinical populations. We first review how SATs have been used clinically, and we next present a summary of previous reviews and explain how the present study differs from them.

Clinical Use of SATs

The self-help industry has proved to be enthusiastically embraced by the general public. In 2004, the self-improvement industry was estimated to be an \$8.56-billion business, with a market growth of 50% between 2000 and 2004 (Salerno, 2005). In 2003 alone, more than 3,500 new self-help books were published, accounting for more than \$650 million in sales (Salerno, 2005). But SATs are also popular among mental health professionals in their clinical practice. Psychologists themselves have long been involved in the development and promotion of SATs and self-help materials (Rosen, Glasgow, & Moore, 2003). The use of self-help methods is widespread as an adjunctive technique in psychotherapy (Norcross et al., 2000; Ogles, Lambert, & Craig, 1991), as it has been suggested that they can enhance treatment outcome (Scogin, 2003). SATs have also been integrated into the therapeutic process for the handling of secondary or less complex aspects of the client’s complaint or accelerating the process of his or her

MARISA MENCHOLA is a PhD candidate in the clinical psychology program at the University of Arizona, where she obtained her Master’s degree in psychology. She is currently a predoctoral intern at the Department of Psychiatry, University of Arizona, where her clinical activities focus on the psychological assessment of medical patients. Her main research interests are the processes of emotional change and the physiology of emotion regulation.

HAL S. ARKOWITZ received his PhD in clinical psychology from the University of Pennsylvania. He is an associate professor in the Department of Psychology at the University of Arizona. His areas of research include the development and evaluation of a model of resistance in psychotherapy as ambivalence about change, measuring resistance, the development and evaluation of interventions to work with resistance, and process and outcome research on motivational interviewing applied to anxiety disorders and depression.

BRIAN L. BURKE received his PhD in clinical psychology from the University of Arizona. He is an assistant professor of psychology at Fort Lewis College in Durango, Colorado. His areas of research include the teaching of psychology and motivational interviewing, and he remains active as a licensed clinical psychologist in the state of Colorado.

WE THANK Varda Shoham and Alfred Kaszniak for assistance with this article and Patrick McKnight for his generous consultation.

CORRESPONDENCE CONCERNING THIS ARTICLE should be addressed to Marisa Menchola, Department of Psychology, University of Arizona, Tucson, AZ 85721. E-mail: menchola@u.arizona.edu

getting acquainted with the therapeutic approach and techniques, thus saving in-session time for issues that may be better addressed with a therapist (Floyd, 2003). SATs also have the advantage of potentially enhancing the client's feelings of responsibility and self-efficacy for therapeutic change.

In addition to their use as an adjunct to psychotherapy, SATs have also been used as stand-alone treatments. Their low cost in terms of time and money, their wide availability, and their potential to reach a broad audience (Mains & Scogin, 2003) make SATs a potentially efficient way to help, if not everyone, at least clients with certain characteristics (Mahalik & Kivlighan, 1988; Scogin, Jamison, & Gochneaur, 1989). SATs are also a potentially useful alternative for specific populations that may not have access to regular psychotherapeutic treatment, such as the elderly, inmates, people in underserved areas, people with disabilities, and caregivers, or for people reluctant to engage in or who hold negative attitudes toward psychological treatment. It has also been suggested that SATs may be used as a cost-effective, nonintrusive, first-line treatment in a stepped-care model (Mains & Scogin, 2003; Scogin, Hanson, & Welsh, 2003).

The use of SATs as the main or only therapeutic intervention is not free of controversy. Rosen (1987; Rosen et al., 2003) has repeatedly pointed out that techniques applied successfully by a therapist are not always self-administered successfully, that ineffective SATs can actually lead to the worsening of a problem, and that well-intentioned changes to a self-administered treatment can lead to an ineffective intervention such that the effect of any change in instructional content must be assessed. Other authors (Marrs, 1995; Scogin, Bynum, Stephens, & Calhoun, 1990; Starker, 1988) have cautioned that the most widely researched books are not representative of those found at the self-help section of a bookstore or even of those most frequently prescribed by psychologists.

Previous Reviews of the Efficacy of SATs

A few published meta-analyses have examined the efficacy of SATs for a wide range of target problems. All have found SATs to be reliably more effective than no-treatment control groups (NTCs). Another consistent result has been that, when compared with standard therapist-administered treatments (TATs), effect sizes for SATs and TATs have not differed significantly. The effect sizes estimated for SATs have been variable but relatively large when compared with effect sizes estimated for psychotherapy such as those found in Wampold et al.'s (1997) comparison of bona fide psychotherapies with no treatment ($d = 0.82$) and placebo ($d = 0.48$). In their meta-analysis of 40 studies, Scogin et al. (1990) estimated an average effect size of $d = 0.96$ for SATs when they were compared with no treatment (the effect size for the subgroup composed of depression and anxiety studies was 0.73 and for the phobias group, 1.10). Gould and Clum (1993) found an effect size for self-help interventions of $d = 0.76$, with $d = 0.74$ for depression and $d = 1.11$ for fear reduction in 40 studies. In a meta-analysis of 70 samples, Marrs (1995) estimated an effect size for bibliotherapy studies of $d = 0.57$, with $d = 0.57$ for depression and $d = 0.91$ for anxiety. In a meta-analysis of six studies focusing on self-help methods for unipolar depression, Cuijpers (1997) found an effect size of $d = 0.82$. Finally, in a meta-analysis of 14 studies, den Boer, Wiersma, and Van Den Bosch (2004) estimated

an effect size for self-help methods of $d = 0.83$ when compared with control (placebo and waiting-list) conditions.

Why Another Meta-Analysis?

We were interested in conducting a new review to provide a cleaner test of SATs for anxiety and depression by controlling a number of confounds present in previous meta-analyses.

Nonclinical Samples

All existing meta-analyses except Cuijpers (1997) and den Boer et al. (2004) included studies done with college students and other nonclinical samples and studies lacking adequate criteria for inclusion (e.g., studies that did not use a diagnostic system or a cutoff score in a valid and reliable clinical instrument as inclusion criteria). This resulted in a combination of diagnosable problems such as depression, anxiety disorders, or substance dependence and target problems such as study behavior and parenting skills (Scogin et al., 1990), headache (Gould & Clum, 1993), ethnocentrism, and happiness (Marrs, 1995).

Amount of Therapist Contact

Studies included in previous meta-analyses have varied widely in the amount of therapist contact allowed under the SAT category. Marrs (1995) reported that "the participants in the bibliotherapy studies met with a therapist a mean of 36 minutes per week" (p. 852). Similarly, several studies included in den Boer et al.'s 2004 meta-analysis consisted of bibliotherapy in addition to contact with a therapist or researcher, including a study that involved six 60-min sessions with a clinician (Al-Kubaisy, Marks, Logsdail, & Marks, 1992). Our interest was in a purer form of SATs. Specifically, we were interested in the reduced influence of a therapist in an expert role guiding treatment, whether physically present or not, so we included only studies with very little or no therapist contact.

Contact With Other Participants

All previous meta-analyses except Cuijpers (1997) included self-help group interventions (e.g., support groups). Our interest was on treatments implemented by an individual alone, so studies of self-help group interventions were excluded.

Examining previous reviews did not allow us to determine the clinical significance of treatments administered by an individual alone, mainly because of the inclusion of too much contact with a therapist or researcher, the inclusion of group interventions, heterogeneous samples of different disorders, and the inclusion of subclinical diagnoses. Our review was concerned with the efficacy of relatively pure forms of SATs for clinical levels of depression and anxiety (*pure* within the limits imposed by pre- and posttreatment assessments and periodic contact to monitor progress).

Method

The first objective of this review was to provide a comprehensive account of studies conducted on SATs for clinical anxiety and depression that would allow us to identify some critical issues relevant to methodological quality, external validity, and limitations of the research. A second objective was to describe the

efficacy of SATs under the following conditions: (a) when compared with NTC groups and when compared with TATs; (b) at posttreatment (initial response) and at follow-up (sustained efficacy); and (c) as expressed through effect sizes, as well as by other indicators of efficacy such as percentage of participants who showed clinically significant improvement. Our third objective was to identify moderators that might account for variations in the efficacy of SATs such as the amount of contact with a therapist and/or researcher or the methodological quality of the study.

Study Selection

We conducted a database search in PsycInfo, Medline, and ProQuest using the keywords *self-administered treatment*, *self-help*, and *bibliotherapy*. No limitations were specified for cultural and linguistic characteristics or date of publication. We also reviewed the references of the prior reviews and outcome studies. Our criteria for inclusion in the meta-analysis were the following:

1. The design had to include at least a comparison between an SAT and a control group (no treatment or placebo) or a comparison between an SAT and a TAT.
2. Assignment to groups had to be randomized.
3. In order to be considered an SAT, the treatment had to be designed to be implemented by the client. This excluded treatments administered by therapists or that required regular group or individual therapeutic or instructional contact, for example, individual and group therapy, self-help groups, and psychoeducational courses.
4. The SAT had to be implemented as the primary treatment, not as an adjunct to another intervention (e.g., a TAT or treatment as usual in the community).
5. Weekly meetings or phone contacts should have had the sole purpose of monitoring progress and clarifying procedures (e.g., symptom assessment, answering questions about the book) and should not have exceeded 15 min per week.
6. The study had to use an adolescent or adult clinical sample. For our purposes, this meant all participants must have been screened to meet a clinical level of symptom severity, for example, with a Beck Depression Inventory (BDI; Beck, Ward, & Mendelson, 1961) cutoff for depressive symptoms, or some kind of diagnostic criteria such as the then-current edition of the *Diagnostic and Statistical Manual of Mental Disorders*.
7. The study had to report at least one measure of depression and/or anxiety at both pre- and posttreatment times.
8. Sufficient statistical information to permit estimation of an appropriate effect size had to be either reported in the study or provided at our request.
9. The study had to be either a published article or an unpublished doctoral dissertation.

Analyses

Relevant descriptive features. The first step of the analysis consisted of coding the relevant features of the selected studies, including but not limited to the following: sample size, relevant demographics, comparison treatment(s), target problem (depression or anxiety), diagnostic criteria used, therapeutic material used in the SAT, nature of the TAT, amount of therapist and/or researcher contact, measures used, length of treatment, percentage of

participants improved, percentage of participants seeking additional treatment at follow-up, attrition rates, and follow-up lengths and rates.

Individual effect sizes. To explore the efficacy of SATs at posttreatment (initial response), we estimated one effect size for each comparison in each study. For example, if one study had four groups (a cognitive SAT, a behavioral SAT, a TAT, and an NTC group), four effect sizes were estimated: SAT(cognitive) versus NTC, SAT(behavioral) versus NTC, SAT(cognitive) versus TAT, and SAT(behavioral) versus TAT. Because some studies reported more than one outcome measure for the same construct (e.g., mean scores on both the BDI and the Hamilton Rating Scale for Depression [HRSD; Hamilton, 1960]), we selected one instrument and estimated an effect size on the basis of that measure (Hedges & Olkin, 1985; Lipsey & Wilson, 2001). Our criteria for selecting an instrument were the following:

1. the one most closely related to the target problem (e.g., a scale measuring panic symptoms as opposed to a general anxiety scale in a study of panic disorder), or
2. the most psychometrically sound instrument, or
3. the instrument that had been used as diagnostic criterion in the study.

Because different studies operationalized the target problem differently (i.e., used different assessment instruments), the effect size estimated was the standardized mean difference. Following Hedges and Olkin (1985), we first computed the unit-free effect size g by subtracting the mean of the control group from the mean of the experimental group and dividing the difference by the pooled standard deviation. We subtracted the mean of the NTC or TAT group from the mean of the SAT mean, so a positive sign would indicate that the SAT group did better than the comparison group. Then d , an estimation that corrects for small sample bias on g , was obtained. Confidence intervals (CIs) were estimated for each individual effect size.

Combined effect sizes. Once individual effect sizes were estimated for all studies, a combined effect size for the efficacy of SAT versus NTC and a combined effect size for the efficacy of SAT versus TAT were estimated. First, in order to generate an independent sample of effect sizes, we used only one effect size per study to compute a combined effect size. In this way, each study could contribute only one effect size to the combined SAT versus NTC estimate and one effect size to the combined SAT versus TAT estimate. If the study had more than one SAT group, our criteria for selecting an individual effect size were the following:

1. if the SAT groups differed in amount of therapist contact, the purest SAT group was selected (e.g., a no-contact SAT instead of a minimal contact SAT); and
2. if the SAT groups differed in therapeutic orientation, the most common group was selected (e.g., cognitive SATs for depression).

Next, because effect sizes based on larger samples are more precise estimates of the corresponding population parameters, the effect size from each study was weighed by its inverse variance (Lipsey & Wilson, 2001). Finally, we estimated combined effect sizes using a macro for SPSS (Lipsey & Wilson, 2001).

Other measures of efficacy. Following Westen and Morrison (2001), we examined multiple indicators of clinical efficacy of the SATs that are significant for the evaluation of a treatment (Kazdin, 1999). These other indicators of initial response and sustained

efficacy included percentage of participants who improved at posttreatment, percentage who remained improved at follow-up, and percentage seeking additional treatment at follow-up.

Moderator analyses. Variables that were coded to be examined as potential moderators of effect size included the following:

1. Target problem. Previous meta-analyses have found the effect sizes for SATs to differ for depression and anxiety (Gould & Clum, 1993; Marrs, 1995), suggesting that SATs may be differentially effective, depending on the clinical problem (Mains & Scogin, 2003).

2. Methodological quality. The evidence relating methodological quality and effect size is mixed (e.g., Glass, McGaw, & Smith, 1981; Stanton & Shadish, 1997). We expanded the methodological quality scale developed by Miller et al. (1995). The original scale has 12 items and a maximum score of 17. New items related to treatment integrity, instruments used, and measures of process and efficacy were included, and the range of the scores for some items was increased (see Appendix A in supplemental data, available on the Internet). The new scale, with 19 items and a maximum score of 35, increased the variability in the quality scores of the studies.

3. Contact. Our inclusion criteria restricted our sample to low-contact studies. Still, there were the following three levels of contact in the studies included: (a) no contact at all during the treatment period, (b) contact halfway into the treatment period (for example, one phone call at Week 4 of an 8-week treatment), or (c) weekly contact.

4. Researcher allegiance. The importance of considering researcher allegiance in efficacy studies has been well documented (e.g., Luborsky et al., 1999). Although it was difficult to code for allegiance, some guidelines were whether the researcher was examining the efficacy of an SAT that he or she had authored and whether it had been hypothesized that the SAT would be more efficacious than alternative treatments.

5. Sample. Adolescents, adults, or older adults.

6. Availability of an alternative treatment. Following Rohen (2002), we coded whether there was a TAT group in the study because a participant's knowledge that he or she could have been assigned to a therapist-directed condition may have influenced his or her expectations or motivation.

We performed a homogeneity analysis on combined effect sizes using the Q statistic (Hedges & Olkin, 1985) to determine whether there was significant heterogeneity among the effect sizes that warranted further exploration (Hunter, Schmidt, & Jackson, 1982). Then we conducted weighted multiple regressions to examine the impact of each moderator on the combined effect sizes in an attempt to arrive at a predictive model through an exploratory approach of forward selection and backward elimination.

Results

Description of the Studies

Studies included. Twenty-four outcome studies were retrieved that fit our inclusion criteria: 11 for depression and 13 for anxiety. Of these, 20 compared SATs with NTCs (11 for depression, 9 for anxiety), and 9 compared SATs with TATs (2 for depression, 7 for anxiety). Additionally, 3 studies compared SATs with group treatments (2 for depression, 1 for anxiety), and 3 compared SATs with placebo treatments (1 for depression, 2 for anxiety). Within the

anxiety studies, there were 6 studies on panic disorder with and without agoraphobia, 2 on spider phobia, and 1 on each of the following: agoraphobia, panic attacks not necessarily meeting criteria for panic disorder, snake phobia, social anxiety, and all anxiety disorders. A description of the studies included and the comparison groups can be found in Appendix B (supplemental data, available on the Internet).

Diagnostic criteria. All the studies used either a diagnostic system or a cutoff in a valid measure indicating a clinical level of the symptoms. Six depression studies and 10 anxiety studies used *Diagnostic and Statistical Manual of Mental Disorders* criteria. All depression studies used a cutoff in one or more depression scales, mostly cutoffs of 10 on the HRSD or BDI.

Exclusion criteria. The great majority of studies excluded participants with comorbid conditions. All depression studies excluded participants believed to be at suicidal risk or presenting with other psychopathology (i.e., manic episodes, psychosis, substance abuse or dependence) or severe physical illness. Similarly, most anxiety studies ruled out comorbid psychological diagnoses (e.g., depression) and physical conditions (e.g., hypertension, seizure disorder, or respiratory problems). Seven depression and five panic studies did not exclude participants who were on stabilized medications.

Exclusion rates and attrition. Forty-two percent of the people inquiring about a given study actually started treatment. The main reasons why people who sought treatment were not included were the following: no further interest in the study once they learned more about it, failure to meet the inclusion criteria, or failure to start treatment after baseline assessment. Eighty-five percent of the intent-to-treat sample completed the treatment phase. The total number of completers for the 24 studies was 934 participants (mean sample size = 39). Of the completers, an average of 80% were available for assessment at follow-up.

Description of the sample. Seventy-nine percent of the sample were women, the mean age was 51 years, and the mean education for the adult participants was 13.7 years. Eighty-eight percent of participants were White Caucasian in the six studies that reported race. For the studies that did not exclude participants who were on stabilized medication, an average of 40% of the participants were on medications during the study.

Severity of symptoms at pretreatment. The mean pretreatment scores for depression were 21.9 on the BDI and 17.9 in the HRSD. For the panic studies, the mean number of full panic attacks per week at pretreatment was 2.6.

Nature of SATs. The depression studies were remarkably homogeneous. It is important to consider that 72% of the depression studies were conducted within a single research program (Ackerson, Scogin, McKendree-Smith, & Lyman, 1998; Bowman, Scogin, & Lyrene, 1995; Floyd, Scogin, McKendree-Smith, Floyd, & Rokke, 2004; Jamison & Scogin, 1995; McKendree-Smith, 1998; Rohen, 2002; Scogin, Hamblin, & Beutler, 1987; Scogin et al., 1989). The prototypical depression study compared using cognitive bibliotherapy, mostly Burns (1980), with a waiting list, during a treatment phase of 4 weeks, with weekly phone calls for monitoring purposes (Ackerson et al., 1998; Floyd et al., 2004; Jamison & Scogin, 1995; Landreville & Bissonnette, 1997; Scogin et al., 1987). Only Schmidt and Miller (1983) and Wollersheim and Wilson (1991) contacted their participants midway through treatment instead of weekly.

The SATs for anxiety disorders were more varied in terms of length (from 2 to 12 weeks) and contact (no contact at all or contact at midtreatment). The SAT for phobias was self-administered desensitization (Ghosh & Marks, 1987; Hellström & Öst, 1995; Öst, Salkovskis, & Hellström, 1991; Rosen, Glasgow, & Barrera, 1976). With the exception of Carlbring, Westling, Ljungstrand, Ekselius, and Anderson (2001), who evaluated the efficacy of an Internet-based cognitive-behavioral treatment, panic studies used bibliotherapy, either alone (Febbraro, Clum, Roodman, & Wright, 1999; Gould, Clum, & Shapiro, 1993; Hecker, Losee, Fritzler, & Fink, 1996; Lidren et al., 1994) or with other interventions such as monitoring (Febbraro et al., 1999) or video- and/or audiotapes (Gould & Clum, 1995; Parry & Killick, 1998; White, 1995).

Initial Response

For purposes of analysis, NTC refers to a waiting list, and TAT refers only to individual treatment with a therapist. Because only three studies included placebo controls and group interventions, these were analyzed separately. All *ns* refer to the number of studies included in the analysis.

Effect sizes for the comparison of SATs with NTCs. Table 1 (supplemental data, available on the Internet) lists the effect sizes for those studies that compared an SAT with an NTC, as well as the outcome measures used to estimate effect sizes. The overall effect size of SATs when compared with no-treatment controls was $d = 1.00$ ($p < .01$; $n = 20$; 95% CI = 0.81, 1.18), with individual effect sizes ranging from -0.14 to 2.48. The effect size for depression was $d = 1.28$ ($p < .01$; $n = 11$; 95% CI = 1.03, 1.53), and the effect size for anxiety was $d = 0.67$ ($p < .01$; $n = 9$; 95% CI = 0.40, 0.94). The combined effect size for the subgroup of anxiety studies aimed at panic was $d = 0.45$ ($p < .01$; $n = 6$; 95% CI = 0.13, 0.77), and for the rest of anxiety studies, it was $d = 1.14$ ($p < .01$; $n = 3$; 95% CI = 0.67, 1.61). However, it is interesting to note that when we estimated the effect size for panic using the frequency of full-blown panic attacks as an outcome measure (instead of the Panic Attack Symptom Questionnaire; Clum, Broyles, Borden, Watkins, & Hayes, 1990), the resulting effect size was $d = 0.70$. Figure 1 illustrates the combined effect sizes for different clinical targets.

Effect sizes for the comparison of SATs with TATs. Table 2 (supplemental data, available on the Internet) lists the effect sizes

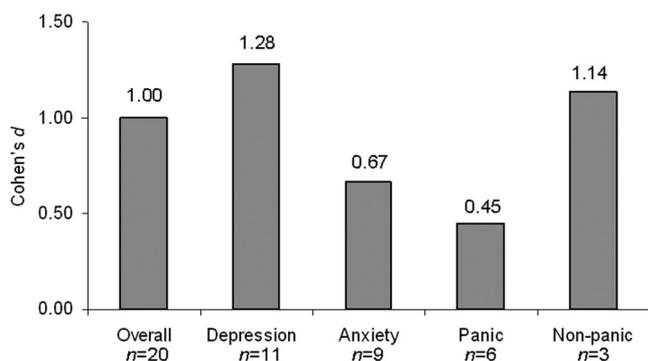


Figure 1. Effect sizes for self-administered treatments versus no-treatment controls.

for the studies that compared an SAT with a TAT. The overall effect size was $d = -0.31$ ($p < .05$; $n = 9$; 95% CI = -0.59 , -0.03), with individual effect sizes ranging from -1.51 to 0.66. This indicates that SATs yielded poorer outcomes than did individual interventions directed by a therapist. The effect size for the depression studies was $d = -0.44$ ($p = .15$; $n = 2$; 95% CI = -1.04 , 0.16), and for anxiety $d = -0.27$ ($p = .09$; $n = 7$; 95% CI = -0.59 , 0.04). Only two panic studies included a comparison of SAT with TAT, with a nonsignificant $d = 0.14$ ($p = .70$; 95% CI = -0.56 , 0.83). The remaining five anxiety studies yielded a significant effect size of $d = -0.38$ ($p < .05$; 95% CI = -0.73 , -0.03).

Effect sizes for other comparisons. Three studies compared SATs with group interventions (Lidren et al., 1994; Schmidt & Miller, 1983; Wollersheim & Wilson, 1991). The overall effect size for this comparison was a nonsignificant $d = 0.21$ ($p = .40$; $n = 3$; 95% CI = -0.28 , 0.71). Three studies compared SATs with placebo treatments (Rosen et al., 1976; Schelver & Gutsch, 1983; Scogin et al., 1987), with an overall $d = 0.66$ ($p < .05$; $n = 3$; 95% CI = 0.13, 1.19). Thus, in this small number of studies, SATs were better than placebo interventions but not significantly better than group treatments.

Although exploring the differential efficacy of SATs was not an objective of this study, we estimated effect sizes for those studies that had more than one SAT group (e.g., a cognitive vs. a behavioral intervention). No effect size in a study comparing alternative SATs was significant.

Clinical significance. Different studies defined clinically significant improvement differently. Most studies followed the criteria proposed by Jacobson and Truax (1991). These criteria require (a) a posttreatment score that is closer to the mean of a normal population than to the mean of the depressed sample at pretreatment and (b) evidence of reliable change as indicated by an index estimated on the basis of reliability of the instrument. Some panic studies defined clinically significant improvement as a decrease of more than 50% on the number of panic attacks experienced, whereas others required being panic free at posttreatment. Two out of the three phobia studies used a criterion of being able to complete the exposure hierarchy, which entailed, for example, having a spider crawl on the individual's hands for 20 s. Across studies, an average of 55% of depressed patients, 42% of panic patients, and 7% of phobic patients achieved clinically significant levels of improvement with SATs at posttreatment. Only three depression studies explicitly reported the percentage of participants experiencing deterioration during the treatment phase, with an average of 9%. Also on the basis of three studies, we determined that the average of patients not meeting diagnostic criteria for depression at posttreatment was 67%. The mean posttreatment scores were 9.1 for the HRSD and 9.3 for the BDI.

Sustained Efficacy

Effect sizes for follow-ups were not estimated because in the great majority of studies the control groups received treatment after posttreatment assessment. Because of the limitations in interpreting within-groups effect sizes, we report data on clinical significance only at follow-up.

Whereas studies varied in the length of follow-ups (ranging from 2 weeks to 3 years posttreatment), they have consistently

found that the effects of SATs are maintained at follow-up, with some studies reporting further improvement from posttreatment to follow-up in some measures of depression (e.g., Ackerson et al., 1998; Floyd et al., 2004; McKendree-Smith, 1998; Scogin et al., 1990), phobia (Rosen et al., 1976), and agoraphobia (Ghosh & Marks, 1987). The mean HRSD score was 7.2 for follow-ups conducted up to 6 months posttreatment (mean BDI score was 9.6). The mean HRSD score for follow-ups conducted 1 year or more posttreatment was 7.8. The percentage of participants showing clinically significant improvement from posttreatment to follow-up was 27% for depression and 12% for panic. Thirty-one percent of depressed patients showed no clinically significant change from posttreatment to follow-up, whereas 7% (on the basis of three studies) showed clinically significant deterioration. Unfortunately, it was impossible to estimate the percentage of those showing clinically significant improvement at posttreatment who remained improved at follow-up on the basis of available reports. The percentage of participants not meeting criteria for depression at follow-up was the same as at posttreatment, 67%. Twenty-seven percent of participants had sought further treatment, and 52% reported rereading the assigned book at least partially after treatment.

Moderator Analyses

Moderator analyses were performed only on the SAT versus NTC subgroup, given the small number of studies that compared SATs with TATs. The homogeneity analysis performed in the 20 studies comparing SATs with NTCs showed that it was a heterogeneous group ($Q = 49.61$; $p < .01$). Both the subgroup of depression studies ($Q = 22.61$; $p < .05$) and the subgroup of anxiety studies ($Q = 16.34$; $p < .05$) were heterogeneous as well. Given this heterogeneity, we explored the effect of different moderators on the variance observed. Because of the small number of studies, regression analyses were performed on the 20 studies comparing SATs with NTCs, and not separately for depression and anxiety. Of the moderators tested, three yielded significant regression coefficients: clinical target ($B = -0.61$; $p < .01$), quality ($B = 0.07$; $p < .01$), and contact ($B = 0.37$; $p < .01$). Researcher allegiance, sample, and availability of an alternative treatment were not significant predictors of effect size.

The inclusion of more than one predictor in the regression equation did not show any increment in the predictive power of the individual predictors (R -squares for clinical target, methodological quality, and amount of therapist and/or researcher contact were .22, .29, and .24, respectively). This is explained by the finding of high correlations, ranging in magnitude from 0.71 to 0.87, among these three moderators. In other words, the depression studies included were high-quality studies ($M = 22.5$, $SD = 2.8$) with higher levels of contact, whereas the anxiety studies had lower quality scores ($M = 16.0$, $SD = 4.0$) and involved significantly less contact. This is an important observation for one to keep in mind when interpreting the results of the moderator analyses. The fact that these three highly correlated moderators basically accounted for the same portion of the variance in the effect sizes was evident when we attempted a hierarchical regression: Whichever moderator was inserted first into the equation took all the predictive power, rendering the other two predictors nonsignificant.

Discussion

The purpose of the present review was to explore the efficacy of SATs when applied to a selective sample of patients with clinical levels of depression and anxiety symptoms. The results indicated that although SATs were more effective than NTCs, they produced significantly lower levels of improvement than did therapist-directed individual interventions. There was also some preliminary evidence that SATs may be better than placebo interventions but not significantly better than therapist-administered group interventions. There was no evidence in our sample of differential efficacy when we compared SATs of different modalities or orientations.

Previous meta-analyses had found nonsignificant differences between the outcome of TATs and SATs (den Boer et al., 2004; Gould & Clum, 1993; Marrs, 1995; Scogin et al., 1990). If this were valid, one implication would be that SATs could replace TATs, at least under certain conditions. However, when we examined anxiety and depression outcomes in clinical samples and with minimal therapist and/or researcher contact, this was not the case. Although the subset of studies comparing SATs with TATs had limitations (i.e., small number of studies, limited variety of SATs, inclusion of one-session TATs), our results show that TATs were significantly more effective than were SATs. Thus, whereas for milder disorders other reviews have suggested that SATs may be helpful, for more serious disorders SATs alone may be insufficient without therapist contact.

Limitations and Caveats

In addition to the limitation imposed by a small sample size of 24 studies, certain characteristics of the studies reviewed limit the generalization of the results. First, although it may be tempting to conclude that SATs are effective for clinical depression and anxiety disorders other than panic disorder but less effective for panic disorder, there were significant differences among the studies reviewed. Depression studies were more homogeneous than were anxiety studies in terms of measures used, length and nature of the SAT, methodological quality, and amount of therapist contact. The prototypical depression study rated high in methodological quality and involved weekly contact with participants for assessment purposes. The prototypical anxiety study had less contact (either no contact at all or contact once or twice during the treatment phase) and rated lower in methodological quality. The high level of overlap among clinical target, methodological quality of the study, and amount of contact makes it impossible to discern which one of these potential predictors was responsible for the variation in effect sizes.

A second issue for one to bear in mind when evaluating the efficacy of SATs is the importance of considering the clinical impact of these treatments beyond the magnitude of the effect sizes estimated. Especially in the case of anxiety, although the effect sizes suggest overall improvement in the treatment groups, other indicators of efficacy revealed that the effects at the clinical level may have been small. Some measures not used for the estimation of effect sizes (e.g., behavioral measures for phobia or measures of frequency or duration of panic attacks for panic disorder) failed altogether to differentiate between treatment and control groups at posttreatment (e.g., Carlbring et al., 2001; Febraro et al., 1999; Gould & Clum, 1995; Parry & Killick, 1998; Rosen et al., 1976).

Another important point is the extent to which these samples were self-selected. Although some kind of self-selection is present to some degree in most intervention studies, an additional bias may have been present in our sample because of the participants' knowledge that treatment would specifically involve a self-administered intervention, most commonly reading a book. It could be argued that people with certain characteristics (e.g., low levels of self-efficacy, external locus of control, low expectations of success, or lack of enjoyment in reading) might not agree to an SAT. One cannot predict what effect the prescription of an SAT would have had on these people.

Fourth, one should keep in mind that very little is known as yet about differential responses to SATs. Beutler et al. (1991) reported that depressed patients with an internalizing coping style and those high in defensiveness and/or resistance did better in a self-directed intervention than did those in group cognitive-behavioral or experiential therapies. However, in our sample, no participant characteristics were consistently related to likelihood of completing treatment or to better outcome. Results were mixed regarding the relationship between variables such as age, education, and severity of symptoms and the likelihood of successfully completing and responding to SATs.

Fifth, very little is known about the therapeutic processes responsible for change in SATs, as studies have failed to identify reliable mediators of outcome. Some studies found significant changes on cognitive variables in cognitive SAT groups when they were compared with controls (Ackerson et al., 1998; Bowman et al., 1995; Gould & Clum, 1995; Jamison & Scogin, 1995; Parry & Killick, 1998), but these changes may not have been limited to the cognitive groups (e.g., Bowman et al., 1995). Moreover, tests of the mediational effects of cognitive or behavioral variables on depression failed to confirm such effect (Floyd et al., 2004; McKendree-Smith, 1998; Rohen, 2002). Results were also mixed regarding the relation between adherence to the SAT and outcome. In bibliotherapy for depression, the self-reported average number of pages read was 84% of the assigned reading material, but the range was from 0 to 100%. Rosen et al. (1976) reported that only 47% of participants in self-administered desensitization groups completed at least half of their hierarchy. Although some studies found no correlation between outcome and completion, comprehension, or practice of the self-help techniques (Ackerson et al., 1998; Floyd et al., 2004; Lidren et al., 1994; Öst et al., 1991), Rosen et al. (1976) and Rohen (2002) did find associations between adherence to the SAT and outcome.

Finally, one needs to keep in mind the limitations of these estimates of the efficacy of SATs in terms of external validity. As Rosen et al. (2003) pointed out, grouping a limited number of studies into a meta-analysis provides no empirical basis for evaluating the vast majority of untested programs. Not only are *Feeling Good* (Burns, 1980), *Coping With Panic* (Clum, 1990), and the self-exposure manuals created by researchers not necessarily representative of materials that patients use in their everyday lives, but one knows nothing about the efficacy of these same materials used outside of a research context. Even in the studies with no contact during the treatment phase, participants had gone through a pre-treatment assessment, knew that they would be evaluated at post-treatment, and in some cases had to mail in forms related to their mood and compliance with the program. The use of SATs by individuals outside of treatment or research contexts typically does

not involve such feedback, pre- or postassessments, or compliance checks. The contribution of those procedures to outcome is yet to be determined.

Implications for the Clinical Use of SATs

In spite of these caveats, SATs are valuable clinical tools that clearly have a significant effect on anxiety and depression. However, caution is needed when interpreting these results for two main reasons: (a) SATs do not appear to be as effective as TATs in populations with clinically significant symptoms, and (b) we were not able to determine to what extent the amount of contact with a therapist is responsible for positive outcomes in the reviewed studies. The SATs whose efficacy was supported by the research were accompanied by some degree of contact for monitoring adherence and comprehension of the material as well as for assessing symptoms. In light of these results and limitations, it seems reasonable to conclude that TATs should be favored in cases in which symptoms are clinically significant. However, SATs remain both a useful therapeutic component as an adjunct to TATs and a valuable treatment option when no other interventions are available.

There are good reasons to incorporate SATs as part of a therapist-administered intervention, as self-help methods may not only facilitate the therapeutic work but also contribute some unique elements to TATs. For example, it has been argued that SATs may facilitate improvement beyond the treatment phase, perhaps by developing a sense of self-efficacy. Floyd et al. (2004) found that although cognitive therapy for depression had better results than did cognitive bibliotherapy at posttreatment, the difference did not remain at 3-month follow-up because the SAT group improved and the TAT group maintained its gains. Rosen et al. (1976) found that although a group of phobics who received a TAT for desensitization showed no further improvement at 2-month follow-up, the SAT desensitization group showed improvement in some measures of phobic anxiety.

On the basis of the available evidence, we propose a few guidelines related to the use of SATs as adjuncts to psychotherapy. First, therapists should be thoroughly familiar with a particular self-help material before recommending it to a client. Second, therapists should favor those SATs that have research support and avoid interventions whose principles are inconsistent with known and researched psychological principles. Third, the material should be discussed in session. This applies even when the material recommended addresses a secondary aspect of the client's complaint that is not a focus of psychotherapy. Often a clinician may recommend a book that may be potentially helpful for a client but not discuss the material in any detail in session. If this is the case, the client may not get as much benefit from SATs as he or she would if the therapist monitored the patient's progress with the material, evaluated his or her comprehension of the principles and techniques, and monitored the relevant symptoms. If the therapist spends a few minutes each session exploring these issues, the SAT may become a much more powerful ingredient of treatment.

What about SATs as stand-alone treatments? For individuals living in remote areas or having insurance difficulties or other limitations (e.g., inmates, medical inpatients, or caregivers) for whom a regular TAT is unavailable, SATs are a valuable treatment option, particularly if accompanied by some degree of contact with

a mental health professional. For someone who cannot attend or afford regular sessions, an initial meeting can be scheduled to assess symptoms and introduce the material, followed by scheduled follow-ups for discussion of the material prescribed and further monitoring of symptoms. Follow-ups can even consist of brief phone sessions, an affordable option for the client and a cost-effective way for a therapist to aid multiple individuals who would otherwise not receive appropriate help.

In addition to its practical implications, the efficacy of stand-alone SATs has significance for the understanding of therapeutic change processes. Bohart and Tallman (1999), in their description of their "client as an active self-healer" approach to psychotherapy, viewed evidence for the effectiveness of self-help methods as suggesting that "the expert therapist's application of techniques is not always necessary" (p. 35). Further, the fact that the client is the primary agent of change is illustrated by significant improvement that can take place in the absence of a therapeutic relationship. The efficacy of stand-alone SATs also has significance for psychotherapy research. Rosen et al. (1976) suggested that empirically validated SATs can serve as a standard against which more costly forms of treatment might be compared. Given the difficulty of designing a true placebo psychotherapy for outcome research, Schmidt and Miller (1983) have pointed out that "minimal interventions delivered with a high degree of positive expectancy" (p. 329), as is the case with SATs, would be appropriate as comparison therapies. That is, a new treatment should be better for a client than reading a book or watching a tape on his or her own.

Future research in this domain should explore further which individuals benefit most from which self-administered treatments and with what level of therapeutic contact before SATs can be prescribed as the sole therapeutic intervention with full confidence. Until then, however, one can be reasonably confident that SATs, when prescribed and used appropriately, can be of significant benefit to clients experiencing symptoms of depression and anxiety.

References

- References marked with an asterisk indicate studies included in the meta-analysis.
- *Ackerson, J., Scogin, F., McKendree-Smith, N., & Lyman, R. (1998). Cognitive bibliotherapy for mild and moderate adolescent depressive symptomatology. *Journal of Consulting and Clinical Psychology, 66*, 685–690.
- Al-Kubaisy, T., Marks, I. M., Logsdail, S., & Marks, M. P. (1992). Role of exposure homework in phobia reduction: A controlled study. *Behavior Therapy, 23*, 599–621.
- Balch, P., & Ross, A. W. (1974). A behaviorally oriented didactic group treatment of obesity: An exploratory study. *Journal of Behavior Therapy and Experimental Psychiatry, 5*, 239–243.
- Beck, A. T., Ward, C. H., & Mendelson, M. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 861–871.
- Beutler, L. E., Engle, D., Mohr, D., Daldrup, R. J., Bergan, J., Meredith, K., & Merry, W. (1991). Predictors of differential response to cognitive, experiential, and self-directed psychotherapeutic procedures. *Journal of Consulting and Clinical Psychology, 59*, 333–340.
- Bohart, A. C., & Tallman, K. (1999). *How clients make therapy work: The process of active self-healing*. Washington, DC: American Psychological Association.
- *Bowman, D., Scogin, F., & Lyrene, B. (1995). The efficacy of self-examination therapy and cognitive bibliotherapy in the treatment of mild to moderate depression. *Psychotherapy Research, 5*, 131–140.
- Burns, D. (1980). *Feeling good: The new mood therapy*. New York: Guilford Press.
- *Carlbring, P., Westling, B., Ljungstrand, P., Ekselius, L., & Anderson, G. (2001). Treatment of panic disorder via the Internet: A randomized trial of a self-help program. *Behavior Therapy, 32*, 751–764.
- Clum, G. (1990). *Coping with panic*. Pacific Grove, CA: Brooks/Cole.
- Clum, G., Broyles, S., Borden, J., Watkins, P., & Hayes, J. (1990). Validity and reliability of the Panic Attack Symptoms and Cognitions Questionnaires. *Journal of Psychopathology and Behavioral Assessment, 12*, 233–245.
- Cuijpers, P. (1997). Bibliotherapy in unipolar depression: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 28*, 139–147.
- den Boer, P. C. A. M., Wiersma, D., Van Den Bosch, R. J. (2004). Why is self-help neglected in the treatment of emotional disorders? A meta-analysis. *Psychological Medicine, 34*, 959–971.
- *Febbraro, G., Clum, G., Roodman, A., & Wright, J. (1999). The limits of bibliotherapy: A study of the differential effectiveness of self-administered interventions in individuals with panic attacks. *Behavior Therapy, 30*, 209–222.
- Floyd, M. (2003). Bibliotherapy as an adjunct to psychotherapy for depression in older adults. *Journal of Clinical Psychology/In Session: Psychotherapy in Practice, 59*, 187–195.
- *Floyd, M., Scogin, F., McKendree-Smith, N., Floyd, D. L., & Rokke, P. D. (2004). Cognitive therapy for depression. A comparison of individual psychotherapy and bibliotherapy for depressed older adults. *Behavior Modification, 28*, 297–318.
- *Ghosh, A., & Marks, I. (1987). Self-treatment of agoraphobia by exposure. *Behavior Therapy, 18*, 3–16.
- Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.
- Gould, R. A., & Clum, G. A. (1993). A meta-analysis of self-help treatment approaches. *Clinical Psychology Review, 13*, 169–186.
- *Gould, R. A., & Clum, G. A., (1995). Self-help plus minimal therapist contact in the treatment of panic disorder: A replication and extension. *Behavior Therapy, 26*, 533–546.
- Greenberger, D., & Padesky, C. (1995). *Mind over mood: Change how you feel by changing the way you think*. New York: Guilford Press.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry, 23*, 56–62.
- Heather, N., Whitton, B., & Robertson, I. (1986). Evaluation of self-help manual for media-recruited problem drinkers: Six-month follow-up results. *British Journal of Clinical Psychology, 25*, 19–34.
- *Hecker, J., Losee, M., Fritzler, B., & Fink, C. (1996). Self-directed vs. therapist-directed cognitive behavioral treatment for panic disorder. *Journal of Anxiety Disorders, 10*, 253–265.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- *Hellström, K., & Öst, L. (1995). One-session therapist-directed exposure vs. two forms of manual-directed self-exposure in the treatment of spider phobia. *Behaviour Research and Therapy, 33*, 959–965.
- Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies*. Beverly Hills, CA: Sage.
- Jacobson, N., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19.
- *Jamison, C., & Scogin, F. (1995). The outcome of cognitive bibliotherapy with depressed adults. *Journal of Consulting and Clinical Psychology, 63*, 644–650.
- Kazdin, A. E. (1999). The meanings and measurement of clinical significance. *Journal of Consulting and Clinical Psychology, 67*, 332–339.

- *Landreville, P. (1998). Cognitive bibliotherapy for depression in older adults with a disability. *Clinical Gerontologist, 19*(3), 69–74.
- *Landreville, P., & Bissonnette, L. (1997). Effects of cognitive bibliotherapy for depressed older adults with a disability. *Clinical Gerontologist, 17*(4), 35–55.
- *Lidren, D., Watkins, P., Gould, R., Clum, G., Asterino, M., & Tulloch, H. (1994). A comparison of bibliotherapy and group therapy in the treatment of panic disorder. *Journal of Consulting and Clinical Psychology, 62*, 865–869.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Applied social research methods series (Vol. 49). Thousand Oaks, CA: Sage.
- Luborsky, L., Diguier, L., Seligman, D., Rosenthal, R., Krause, E., Johnson, et al. (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice, 6*, 95–106.
- Mahalik, J. R., & Kivlighan, D. M. (1988). Self-help treatment for depression: Who succeeds? *Journal of Counseling Psychology, 35*, 237–242.
- Mains, J., & Scogin, F. (2003). The effectiveness of self-administered treatments: A practice-friendly review of the research. *Journal of Clinical Psychology/In Session: Psychotherapy in Practice, 59*, 237–246.
- Marrs, R. W. (1995). A meta-analysis of bibliotherapy studies. *American Journal of Community Psychology, 23*, 843–870.
- *McKendree-Smith, N. (1998). *Cognitive and behavioral bibliotherapy for depression: An examination of efficacy and mediators and moderators of change*. Unpublished doctoral dissertation, University of Alabama, Tuscaloosa.
- Miller, W. R., Brown, J. M., Simpson, R. L., Handmaker, N. S., Bien, T. H., Luckie, L. F., et al. (1995). What works? A methodological analysis of the alcohol treatment outcome literature. In R. K. Hester & W. R. Miller (Eds.), *Handbook of alcoholism treatment approaches: Effective alternatives* (2nd ed., pp. 12–44). Boston: Allyn & Bacon.
- *Norcross, J. (Ed.). (2003). Integrating self-help into psychotherapy. [Special issue]. *Journal of Clinical Psychology/In Session: Psychotherapy in Practice, 59*(2).
- Norcross, J., Santrock, J., Campbell, L., Smith, T., Sommer, R., & Zucjerman, E. (2000). *Authoritative guide to self-help resources in mental health*. New York: Guilford Press.
- Ogles, B. M., Lambert, M. J., & Craig, D. E. (1991). Comparison of self-help books for coping with loss: Expectations and attributions. *Journal of Counseling Psychology, 38*, 387–393.
- *Öst, L., Salkovskis, P., & Hellström, K. (1991). One-session therapist-directed exposure vs. self-exposure in the treatment of spider phobia. *Behavior Therapy, 22*, 407–422.
- *Parry, R., & Killick, S. (1998). An evaluation of the impact of an individually administered videotape for people with panic disorder. *Behavioural and Cognitive Psychotherapy, 26*, 153–161.
- *Rohen, N. (2002). *Analysis of efficacy and mediators of outcome in minimal-contact cognitive bibliotherapy used in the treatment of depressive symptoms*. Unpublished doctoral dissertation, University of Alabama, Tuscaloosa.
- Rosen, G. (1987). Self-help treatment books and the commercialization of psychotherapy. *American Psychologist, 42*, 46–51.
- *Rosen, G., Glasgow, R., & Barrera, M. (1976). A controlled study to assess the clinical efficacy of totally self-administered systematic desensitization. *Journal of Consulting and Clinical Psychology, 44*, 208–217.
- Rosen, G., Glasgow, R., & Moore, T. (2003). Self-help therapy: The science and business of giving psychology away. In S. Lilienfeld, S. Lynn, & J. Lohr (Eds.), *Science and pseudoscience in clinical psychology* (pp. 399–424). New York: Guilford Press.
- Salerno, S. (2005). SHAM: How the self-help movement made America helpless. New York: Crown.
- *Schelver, S., & Gutsch, K. (1983). The effects of self-administered cognitive therapy on social-evaluative anxiety. *Journal of Clinical Psychology, 39*, 658–666.
- *Schmidt, M. M., & Miller, W. R. (1983). Amount of therapist contact and outcome in a multidimensional depression treatment program. *Acta Psychiatrica Scandinava, 67*, 319–332.
- Scogin, F. (2003). Introduction: The status of self-administered treatments. *Journal of Clinical Psychology, 59*, 247–249.
- Scogin, F., Bynum, J., Stephens, G., & Calhoun, S. (1990). Efficacy of self-administered treatment programs: Meta-analytic review. *Professional Psychology: Research and Practice, 21*, 42–47.
- *Scogin, F., Hamblin, D., & Beutler, L. (1987). Bibliotherapy for depressed older adults: A self-help alternative. *The Gerontologist, 27*, 382–387.
- Scogin, F., Hanson, A., & Welsh, D. (2003). Self-administered treatment in stepped-care models of depression treatment. *Journal of Clinical Psychology, 59*, 341–349.
- *Scogin, F., Jamison, C., & Davis, N. (1990). Two-year follow-up of bibliotherapy for depression in older adults. *Journal of Consulting and Clinical Psychology, 58*, 665–667.
- *Scogin, F., Jamison, C., & Gochneaur, K. (1989). Comparative efficacy of cognitive and behavioral bibliotherapy for mildly and moderately depressed older adults. *Journal of Consulting and Clinical Psychology, 57*, 403–407.
- Scogin, F., Storandt, M., & Lott, L. (1985). Memory-skills training, memory complaints, and depression in older adults. *Journal of Gerontology, 40*, 562–568.
- *Smith, N., Floyd, M., Scogin, F., & Jamison, C. (1997). Three-year follow-up of bibliotherapy for depression. *Journal of Consulting and Clinical Psychology, 65*, 324–327.
- Stanton, M. D., & Shadish, W. R. (1997). Outcome, attrition, and family-couples treatment for drug abuse: A meta-analysis and review of the controlled, comparative studies. *Psychological Bulletin, 122*, 170–191.
- Starker, S. (1988). Psychologists and self-help books: Attitudes and prescriptive practices of clinicians. *American Journal of Psychotherapy, 42*, 448–455.
- *Van Lenkvel, J. (1998). Bibliotherapy in the treatment of sexual dysfunctions: A meta-analysis. *Journal of Consulting and Clinical Psychology, 66*, 702–708.
- Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "All must have prizes." *Psychological Bulletin, 122*, 203–215.
- Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology, 69*, 875–899.
- *White, J. (1995). Stresspac: A controlled trial of a self-help package for the anxiety disorders. *Behavioural and Cognitive Psychotherapy, 23*, 89–107.
- *Wollersheim, J. P., & Wilson, G. L. (1991). Group treatment of unipolar depression: A comparison of coping, supportive, bibliotherapy, and delayed treatment groups. *Professional Psychology: Research and Practice, 22*, 496–502.

Received March 14, 2006

Revision received October 2, 2006

Accepted October 9, 2006 ■