This review synthesized findings from 17 studies since 1998 regarding evaluation of outpatient treatments for adolescent substance abuse. These studies represented systematic design advances in adolescent clinical trial science. The research examined 46 different intervention conditions with a total sample of 2,307 adolescents. The sample included 7 individual cognitive behavior therapy (CBT) replications \((n = 367)\), 13 group CBT replications \((n = 771)\), 17 family therapy replications \((n = 850)\) and 9 minimal treatment control conditions \((n = 319)\). The total sample was composed of approximately 75% males, and the ethnic/racial distribution was approximately 45% White, 25% Hispanic, 25% African American, and 5% other groups. Meta-analysis was used to evaluate within-group effect sizes as well as differences between active treatment conditions and the minimal treatment control conditions. Methodological rigor of studies was classified using Nathan and Gorman (2002) criteria, and treatments were classified using criteria for well-established and probably efficacious interventions based on Chambless et al. (1996). Three treatment approaches, multidimensional family therapy, functional family therapy, and group CBT emerged as well-established models for substance abuse treatment. However, a number of other models are probably efficacious, and none of the treatment approaches appeared to be clearly superior to any others in terms of treatment effectiveness for adolescent substance abuse.
conditions, careful measurement of substance abuse, manual-guided interventions, and relatively long-term follow-up of treatment effects.

The purpose of this review is to examine the findings of randomized clinical trials evaluating psychosocial treatments for adolescent substance abuse since 1998. We use the term substance abuse to refer to pathologic use of drugs and/or alcohol, including the American Psychiatric Association’s (APA’s; 1980, 1987, 1994) Diagnostic and Statistical Manual of Mental Disorders categories of substance abuse disorders and substance dependence disorders. According to DSM-IV, a substance abuse disorder is composed of three or more of a severe array of symptoms, including withdrawal, tolerance, high effort to obtain substance, giving up other activities, and unsuccessful efforts to reduce use (APA, 1994). The intent of our relatively narrow focus on randomized trials of substance abuse treatments was to base our conclusions about the efficacy of particular approaches on the most rigorous studies conducted, in which the interpretation of findings would be clearest (Chambless & Hollon, 1998). Although many broader reviews have been conducted (Vaughn & Howard, 2004; Williams & Chang, 2000), these reviews have focused more on the overall effectiveness of adolescent substance abuse treatment without evaluating the evidence for specific practices. Taken together, the randomized trials evaluating adolescent substance abuse treatments since 1998 have shown relatively similar findings, despite differences in design and methodology. The clarification of methods and design reflected in the studies reviewed, relative to earlier investigations, allows more confidence in the findings and provides a foundation of empirical support for specific treatment approaches to adolescent substance abuse and dependence.

**EARLY RESEARCH SUPPORT**

The majority of evidence-based practices for adolescent substance abuse were developed and initially evaluated prior to 1998, and the review presented here builds on the early treatment outcome research with this population. Initial treatment evaluation efforts focused on a number of family therapy approaches based on conceptual models derived from family systems perspectives, as well as evidenced-based practices developed using behavioral and CBT approaches (see reviews of Davidge & Forman, 1988; Deas & Thomas, 2001; Liddle & Dakof, 1995; Muck et al., 2001; Ozechowski & Liddle, 2000; Ozechowski, Turner, & Waldron, 2003; Stanton & Shaddish, 1997; Waldron, 1997; Waldron & Kaminer, 2004). Although many of the early studies of adolescent substance abuse treatment were characterized by significant methodological limitations, including small sample sizes, lack of appropriate control or comparison conditions, nonrandomized assignment to treatment, poorly specified treatments, and inadequate measurement of substance use (e.g., exclusive reliance on self-report of drug use on summary measures; lack of urine drug screening). Still, some investigations involved group designs and random assignment to treatment conditions, providing a basis for the formal clinical trials conducted in recent years.

**Family Therapy Approaches**

Some of the first randomized trials for adolescent substance abuse were conducted by Szapocznik and colleagues to evaluate an ecologically based structural and strategic family systems therapy model, Brief Strategic Family Therapy (BSFT; Szapocznik, Kurtines, Foote, Perez-Vida, & Hervis, 1983, 1986). BSFT is based on the conceptualization that adolescent substance abuse and other problem behaviors stem from maladaptive family interactions. BSFT focuses on inappropriate family alliances, overly rigid or permeable family boundaries, and the identification of individual family members (typically the adolescent) as the source of dysfunction. It uses strategies such as “joining” the family to enhance engagement, identifying repeated maladaptive interaction patterns, and “restructuring” to develop new, more adaptive patterns (Szapocznik & Kurtines, 1989).

In a pair of studies, families of substance abusing adolescents were randomly assigned to either BSFT involving the whole family or to BSFT implemented with only one family member to evaluate whether family therapy outcome for adolescent substance abuse varies on the basis of who participates in the sessions. Participants in the two studies were 37 and 35 Hispanic families, respectively, all with a substance abusing youth ranging in age from 12 to 20 years. The replication (Szapocznik et al., 1986) was virtually identical to the original study (Szapocznik et al., 1983), but with 15 sessions instead of 12. A self-report of frequency of substance use was obtained from the adolescents using the substance abuse score from the Psychiatric Status Schedule (PSS; Spitzer, Endicott, Fleiss, & Cohen, 1970). In both studies, the two intervention conditions showed significant reductions in drug abuse from pre- to post-treatment and follow-up assessments, with no between-group differences between the two active treatment approaches.

These findings were consistent with an earlier unpublished randomized clinical trial evaluating an ecological BSFT intervention for 33 Cuban American families (Scopetta, King, Szapocznik, & Tillman, 1979). In this study, significant pre- to posttreatment reductions in substance abuse and other problem behaviors were
found for both conditions, with number of target clients and other substance-abusing family members reporting a drug problem of any kind declining from 44 to 19 after an average of 12 sessions. In related work, Joanning, Thomas, Quinn, and Mullen (1992) found support for structural-strategic family therapy in a study of 134 families of substance abusing youth between the ages of 11 and 20 years. Compared to a family drug education condition and adolescent group therapy, family therapy was associated with significantly higher abstinence rates. Beyond treatment outcome, BSFT has also been shown effective in enhancing treatment engagement for youth and their families (Santisteban et al., 1996; Szapocznik et al., 1988). These studies provided additional support for BSFT based on links between treatment engagement and outcome.

BSFT research provided some of the first systematic efforts to establish an empirical basis of support for adolescent substance abuse treatment. As in other early research, however, the studies were limited by small sample sizes, high attrition, inadequate sample characterization with respect to substance use, and the absence of a distinctive comparison condition. In addition, the studies relied on a relatively insensitive treatment outcome measures. A common problem with many early studies was the reliance on measures based on diagnostic status indicating only the presence or absence of substance abuse or on instruments providing only summary measures of use (e.g., highest category of use—three or more times per week), which limited detection of clinically meaningful pre- to posttreatment substance use reductions (Leccese & Waldron, 1994; Liddle & Dakof, 1995; Miller, Westerberg, & Waldron, 2003).

The Functional Family Therapy model (FFT; Alexander & Parsons, 1982; Barton & Alexander, 1981) was evaluated by Friedman (1989) in a randomized trial involving 135 families of youth between the ages of 14 and 21 years presenting with heavy alcohol and drug use (e.g., daily cannabis use). They compared FFT to a parenting skills group intervention. FFT is an integrative ecological model that combines a family systems view of family functioning with behavioral techniques and a multisystemic emphasis. Early FFT sessions focus on enhancing treatment engagement and increasing the family’s motivation for change, using techniques such as relabeling and emphasizing family member interrelatedness (Alexander & Parsons, 1982). Communication, problem solving, and other behavioral intervention strategies are then implemented to establish new patterns of family interaction, and therapists begin to work directly with extrafamilial systems such as juvenile justice and school systems to facilitate generalization of change to the home and community environments. Substance use was assessed using a diagnostic measure developed by the investigator.

Both the FFT and parent training groups showed significant reductions in substance use of more than 50% at follow-up, with improvements in other areas of functioning as well. Although no differences were found between the two treatment groups for treated families, engagement rates differed dramatically (93% in FFT vs. 67% in the parenting condition). In a reanalysis of the entire “intent-to-treat” sample (i.e., including treatment dropouts as failures), Stanton and Shadish (1997) found significantly greater substance use reductions for FFT than the comparison condition.

Tentative support for FFT and structural-strategic therapy was also found in a study conducted by Lewis, Piercy, Sprenkle, and Trepper (1990) with 136 youth described as “regular” substance users. The integrated treatment and a parenting skills intervention both showed significant pre- to posttreatment reductions in drug use, with a greater percentage of youth in family therapy decreasing their use.

Early research also included the evaluation of Multisystemic Therapy (MST). Henggeler and his colleagues conducted two studies, comparing MST to individual therapy (Borduin et al., 1995) and then to a treatment-as-usual juvenile probation supervision intervention (Henggeler, Pickrel & Brondino, 1999). MST has much in common with the structural family therapy approaches studied by Lewis et al. (1990) and Szapocznik et al. (1988) but is also similar to FFT in the ecological focus of the intervention. According to Henggeler and his colleagues (1991), MST views individuals as “nested within a complex of interconnected systems that encompass individual, family, and extra-familial (peer, school, neighborhood) factors” (p. 43). The treatment focus is on changing dysfunction processes that occur in these other systems. Consistent with the effects of MST on delinquent youth, both MST studies (Borduin et al., 1995; Henggeler et al., 1999) were associated with significantly greater reductions in substance use on some measures, relative to the comparison conditions (i.e., individual therapy, treatment-as-usual juvenile probation supervision intervention). Unfortunately, the substance use measure was limited. Self-reported use was based on six questions, three each for marijuana and cocaine. One question focused on use in the last 6 months (dichotomous: yes/no) and two questions on days of use in the past 30 days (categorical: 0, 1–3, 4–9, 10 or more), providing a summary measure unlikely to be sensitive enough to detect pre-to posttreatment change. The 4-year follow-up study, however, revealed a significantly greater reduction in marijuana use for MST compared to the treatment-as-usual condition, on the marijuana urine screen measure. The findings provided initial support for further investigation of MST for adolescent substance abuse.
Behavioral Therapy and CBT Approaches

These approaches have relied primarily on classical, operant, and social learning theories to understand behavior. Within this framework, substance use is viewed as behavior that is learned in the context of social interactions (e.g., observing parents, siblings, peers, or other models in the media) and established as a result of the contingencies in the environment (Akers, Kohn, Lanza-Kaduce, & Radojevic, 1979). The majority of CBT-I and CBT-G models have been multi-component approaches. Cognitive strategies (e.g., identifying distorted thinking patterns) are typically combined with behavioral strategies (e.g., coping with cravings to use; communication; problem-solving; and substance refusal skills training, in which strategies for avoiding high-risk drug use situations are learned). In some models, a motivational enhancement component is offered in the initial sessions (Miller & Rollnick, 1991). Behavioral family therapy models extend the behavioral framework to examining individual behavior in the social context of the family. Within such family models, substance use and abuse are assumed to be determined, at least in part, by the familial antecedents to or consequences of the substance use behaviors (Bry, 1988).

A number of early studies evaluated individual and group behavior therapy to treat adolescent substance abuse (e.g., Davidge & Forman, 1988). The behavioral treatments varied widely, including aversive conditioning (Duehn, 1978); social skills training, problem-solving, self-monitoring, and other behavioral modification approaches in inpatient, residential, and outpatient settings (DeJong & Henrich, 1980; T. E. Smith, 1983); and contingency management and other programs for parents and adolescents (Cook & Petersen, 1985; Iverson, Jurs, Johnson, & Rohen, 1978; Iverson & Roberts, 1980). Although the findings from this research showed promise, many of the studies involved quasi-experimental designs, relied on single-subject designs or had very small samples, lacked control or comparison conditions, and/or included limited measures of substance use such as diagnostic status (i.e., present vs. absent), “summary” measures that aggregate reports of use into categories (e.g., days of use per month: 1–3, 5–7, 8–12, 13–19, 20 or more) or urine drug screening (i.e., present vs. absent). Such measures can be useful in epidemiological research, but are relatively insensitive measures for detecting change from pre- to posttreatment.

Still, two family-based behavioral treatment studies are noteworthy. Azrin and his colleagues compared family therapy to a process-oriented, nondirective (supportive) adolescent group therapy intervention (Azrin, Donohue, Besalel, Kogan, & Acierno, 1994; Azrin, McMahon, & Donohue, 1994), and Krinsley and Bry (1995) compared family therapy plus school-based intervention to the school-based intervention only. Each of these studies showed that behavioral family therapy was superior to the comparison conditions, providing consistent support for the family-based treatments. Although the sample sizes were small, each of these studies involved random assignment of youth to treatments and clearly specified interventions. In addition, substance use was assessed using the timeline followback procedure (TLFB), now a standard in addictions treatment research (Sobell & Sobell, 1992). TLFB is a structured interview technique that samples a specific period, using a monthly calendar and memory anchor points to reconstruct daily consumption during the period of interest. These methodological strengths enhanced the quality of the studies, lending confidence to the empirical findings. In recent years, several more carefully designed randomized clinical trials have been conducted evaluating behavioral and CBT interventions (Waldron & Kaminer, 2004). These studies are included in the current review.

PARAMETERS OF THE CURRENT REVIEW

An important facet of substance abuse treatment research designs is that placebo, no-treatment, and waitlist control groups have rarely been used. One reason for this, according to Stanton and Shaddish (1997), may be ethical concerns associated with withholding promising or beneficial interventions from severely impaired, treatment seeking individuals such as those with substance use disorders. Although this issue has been raised in the broader arena of clinical research (Arean & Alvidrez, 2002; Basham, 1986; Baucom, Hahlweg, & Kuschel, 2003; Kazdin, 1986, 2003; Kendall, Holmbeck, & Verduin, 2004; La Vaque & Rossiter, 2001; O’Leary & Borkovec, 1978), the absence of no-treatment and placebo control conditions in addictions treatment research is an accepted standard (Carroll, Rounsaville, & Gawin, 1991; DeLeon, Inciardi & Martin, 1995; Stanton & Shadish, 1997; Szapocznik et al., 1986). According to Kazdin (2003), no-treatment or waitlist control groups are not always necessary if evidence consistently shows that the absence of treatment has no effect. This is the basis for the standard followed in adolescent substance abuse research “in light of previous research showing lack of improvement in untreated or minimally treated samples” (Dennis et al., 2004, p. 198). However, approximately half of the studies examined and reviewed in this article have used random assignment to a “service as usual” or minimal treatment control conditions to evaluate treatment outcomes. Such studies, comparing active treatments with minimal treatment conditions, are considered “highly prized” by
Chambless and Hollon (1998, p. 8) because they control for the effects of nonspecific factors.

The absence of no-treatment conditions in a significant percentage of adolescent treatment outcome research has implications for the application of the review criteria (Chambless et al., 1996; Nathan & Gorman, 2002) by which the treatment approaches included in this article are evaluated. According to Chambless et al. (1996), interventions evaluated using group research designs may meet one of two sets of criteria: well-established or probably efficacious. Both sets of criteria require manual-guided treatment and clearly specified sample characteristics. Well-established psychosocial interventions for child and adolescent disorders specifically require at least two well-conducted group-design studies, conducted by different investigative teams, showing the treatment to be either superior to a placebo or alternative treatment or showing the treatment to be equivalent to an already-established treatment. Probably efficacious interventions require two studies showing the intervention to be more effective than a no-treatment control group or two group-design studies that would meet criteria for well-established interventions but conducted by the same investigator.

The Nathan and Gorman (2002) criteria classify treatment outcome studies into six types. Because of our restricted focus on randomized clinical trials, we have reviewed only Type 1 and Type 2 studies. Type 1 studies are clinical trials that include random assignment to comparison groups, blinded assessments, clear exclusion and inclusion criteria, strong measurement, sample sizes sufficient for statistical power, sufficient length of follow-up to adequately evaluate outcomes, and clearly described statistical methods. Type 2 studies are clinical trials in which some aspects of Type 1 studies are missing but the study is not fatally flawed. Type 3 studies and beyond, not included in this review, are those studies that have clear methodological limitations or otherwise have more limited utility for evaluating treatments that work (e.g., secondary data analysis, review papers).

Without the systematic inclusion of a no-treatment, placebo, or waitlist control groups, otherwise rigorous adolescent substance abuse treatment outcome studies comparing two or more active treatments with replications across multiple investigative sites cannot meet criteria for a well-established treatment because of the limited power to detect between group differences when both active treatments show benefit. In our view, this approach to categorization for the adolescent substance abuse treatment field is limited and could have negative consequences. The purpose of establishing evidence-based practices rests on the assumption that clients receive better care with clinicians who implement empirically supported treatments (Sackett, Richardson, Rosenberg, & Haynes, 2000; Woody, Weisz, & McLean, 2005). The application of criteria set forth by Chambless and colleagues (Chambless, 1996; Chambless et al., 1998; Chambless & Hollon, 1998; Chambless et al., 1996) and the consequent acceptance that no well-established treatments exist for adolescent substance abuse could lead clinical consumers to dismiss the existing support for the treatments or to be discouraged or dissuaded from adopting substance abuse interventions that collective evidence has shown to be efficacious. Thus, in this article, we present the empirical evidence for the treatment models reviewed, applying the Nathan and Gorman criteria to classify study type, and we discuss the application of the Chambless et al. (1998) and Chambless et al. (1996) criteria to the interventions in light of the concerns noted previously.

In addition, we examine the evidence for adolescent substance abuse treatments as a whole, relying on meta-analysis. Our meta-analysis approach is generally consistent with the guidelines described in Morris and DeShon (2002) and Lipsey and Wilson (2001) for synthesizing between study effects in repeated measures designs. In this approach, we first calculate a pre- to posttreatment effect size within each of 46 treatment condition replications. This strategy results in a standardized effect size measure for within-condition change that controls for initial differences between the distinct replications (Lipsey & Wilson, 2001; Morris & DeShon, 2002). That is, the standardization controls for initial between-replication characteristics such as measurement procedures, unique study site attributes, and treatment population case mix variations. We classified the 46 replications into four treatment types: minimal treatment control conditions (9 replications, \( n = 319 \)), CBT-I (9 replications, \( n = 367 \)), CBT-G (13 replications, \( n = 771 \)), and family therapy (17 replications, \( n = 850 \)). Next, we used random effects regression procedures within a hierarchical linear modeling (HLM) approach to compare the effect sizes from the four treatment types (Raudenbush & Bryk, 2002).

We recognize the important limitations for this form of meta-analysis as compared to one in which each individual study has random assignment between a specific treatment and a well-designed control condition. In our approach, the comparisons among treatment types are confounded with possible differences between the separate studies. However, the standardization procedure using pretreatment means and standard deviations helps to control for these between study characteristics. Still, the findings from the current approach must be interpreted more cautiously than would occur for meta-analyses composed of randomized trials that have placebo controls with double-blind assessments.
Our review focuses on research evaluating adolescent substance abuse treatment approaches published in English since 1998 or conducted since 1998, if unpublished. Unpublished studies were included to control for possible biases against the publication of nonsignificant findings that can lead to a distorted representation of the actual research conducted in a research domain (Begg, 1994), with consequent overrepresentation by statistically significant rather than nonsignificant findings. To identify unpublished studies, we contacted investigators in the adolescent substance abuse field to determine whether they had or knew of any unpublished studies. The selection of outcome studies for this review was based on a literature search designed to identify all studies evaluating specific treatment models for adolescent substance abuse and involving two or more treatment comparison groups since 1998. Databases systematically searched included PsycINFO, Medline, and Psychological Abstracts. Manual searches of the reference sections searched included PsycINFO, Medline, and Psycholog-
ically Abstracts. Manual searches of the reference sections of related studies and previous reviews were also conducted. The online searches used a variety of keywords and combinations, including adolescent, substance use, substance abuse, drug use, drug abuse, addiction, intervention, and treatment. Specific drugs of abuse (e.g., alcohol, marijuana, cocaine) were also combined with intervention and with treatment in separate searches.

Studies were included if the research met the criteria cited by Nathan and Gorman (2002) as being either a Type 1 or Type 2 investigation. Restricting the review to studies with group designs involving comparisons of two or more conditions ensured that interpretations about efficacious treatments were being drawn from the most rigorous of adolescent substance abuse treatment studies and maximized our ability to make comparisons across studies and intervention approaches. Of the 17 studies included in this review, 14 met nearly all of the Type 1 criteria whereas 3 others met Type 2 criteria. For these latter investigations, Kaminer, Burleson, Blitz, Sussman, and Rounsaville (1998) was limited by a small sample size and significant attrition; Battjes et al. (2004) compared treatments that were not clearly distinguished (and the clinical procedures were not well specified), and Winters et al. (2000) did not include a random assignment to conditions. All of the studies included in this review, except for Dennis et al. (2004) and Winters et al. (2000), were efficacy trials. Investigations were included if (a) study participants were adolescents between 12 and 19 years of age, (b) one or more specific treatment models was evaluated, (c) interventions included outpatient treatments for adolescent substance abuse involving alcohol or other illicit drugs, and (d) a clearly specified substance use measure was used as a primary outcome variable. In general, substance abusing adolescents represent a diverse, heterogeneous population. Although the majority of substance abusing youth are referred for and receive outpatient treatment (Muck et al., 2001), many youth who have engaged in only limited substance use are identified to receive a substance use intervention in schools or other community agencies. Others have engaged in such serious concomitant behaviors that they have been placed in a restrictive environment. Our focus on outpatient treatments for adolescent substance abuse was designed to maximize the homogeneity of study populations to facilitate interpretation of the findings.

Excluded from the review were smoking cessation and tobacco-focused interventions, which represent a specialized and extensive treatment domain (cf. Myers, 2001). Comparisons of findings across tobacco and other substance use interventions were beyond the scope of this review. As previously noted, community-based interventions such as those offered in schools (e.g., preventive interventions), emergency rooms (e.g., motivational interviewing), or juvenile justice settings (e.g., diversion programs) were not included because these interventions typically target a much broader population of substance-using adolescents than adolescents formally diagnosed with substance abuse or dependence disorders. Participants in community-based intervention programs, as well as those in residential or inpatient treatment programs, may differ in systematic ways from substance abusing youth referred to outpatient treatment, yielding a mixture of outcome findings that would be more difficult to interpret.

Using the aforementioned criteria, 17 studies were identified, some of which were included in one or more previous reviews (Deas & Thomas, 2001; Muck et al., 2001; Ozechowski & Liddle, 2000; Waldron & Kaminer, 2004; Williams & Chang, 2000). Table 1 provides a list of the studies reviewed and a summary of the major study characteristics. The Treatment and Contrast Group column in Table 1 shows that, across the studies, 46 distinct treatment conditions were evaluated. Eight family-based interventions were represented: an enhanced ecologically focused variant of BSFT, FFT, Multidimensional Family Therapy (MDFT), MST, Behavioral Family Therapy (BFT), and Integrative Behavioral Family Therapy, Transitional Family Therapy, and Strength Oriented Family Therapy (SOFT). The SOFT model is similar to other strength-focused family models. The treatment employs family-focused, motivational sessions in combination with solution focused family therapy, multifamily skills training groups supplemented with targeted case management (Smith, Hall, Williams, An, & Gotman, 2006).

Several variations of group and individual CBT interventions were also represented, including group CBT (CBT-G), 5- or 12-session versions of CBT-G combined with a Motivational Enhancement Therapy (MET; Miller & Rollnick, 1981) component (MET/CBT5; MET/CBT12), individual CBT, individual MET/CBT,
<table>
<thead>
<tr>
<th>Study Type</th>
<th>Nature of Substance Use</th>
<th>N</th>
<th>Patient Age, Sex, Race, Ethnicity</th>
<th>Treatment and Contrast Group(s)</th>
<th>Therapists</th>
<th>Tx Manual</th>
<th>Fidelity Ratings</th>
</tr>
</thead>
</table>
| Azrin et al., 2001 | Dually diagnosed: SUD + CD/ODD | 56 | 12–17 years, 82% male, 79% White, 16% Hispanic | a. CBT-I: ~20 hr  
M = 13.5, SD = 2.4  
b. BFT: ~20 hr  
M = 13.7, SD = 2.3 | N = 19; therapists nested within condition; clinical psychology doctoral students | Yes | Yes |
| Battjes et al., 2004 | Low to moderate use | 194 | 14–18 years, 84.5% male, 71% Caucasian, 22% African American | a. MI + CBT-G: 19 sessions  
2) Counseling Overview + CBT-G: 19 sessions  
M = 12.9, SD = 11.73 | No information | Unclear | No |
| Dennis et al., 2004 | Low to moderate use | 600 | 13–18 years, 83% male, 61% White, 30% African American | Trial 1  
a. MET = CBT5  
b. MET = CBT12  
c. MET = CBT, FS  
Trial 2  
a. MET/CBT5  
b. ACRA c. MDFT | N not specified; Therapists at four sites: 20% doctoral, 50% master's, 30% bachelor level | Yes | Yes, Ratings not reported |
| Henggeler et al., 1999; Henggeler et al., 2002 | SUD diagnosis | 118 | 12–17 years, 79% male, 50% African American, 47% White, non-Hispanic | a. MST: 12–187 hr;  
M = 40, SD = 28  
b. Usual care: 78% received 0 hr, 22% received ≤10 hr | MST: N = 3; 2 master's, 1 bachelor level | Yes | Yes |
| Kaminer et al., 1998; Kaminer & Burleson, 1999 | Dual diagnosis | 32 | 13–18 years, 61% male 89% White | a. CBT-G: 12 sessions  
b. Interactional group: 12 sessions | N = 4; 1 doctor, 1 master's level in each condition | Yes | Yes, Ratings not reported |
| Kaminer et al., 2002 | Dual diagnosis | 88 | 13–18 years 62 males 79 White | a. CBT-G  
b. Psycho-ed group | N = 4; Doctoral and master's levels | Yes | Yes, Ratings not reported n/a |
| Liddle et al., 2001 | SUD diagnosis; Moderate-heavy use | 182 | 13–18 years, 51% White, non-Hispanic 18% African American 15% Hispanic 6% Asian 10% Other | a. MDFT: 16 sessions  
b. CBT-G: 16 sessions  
c. Family psycho-ed group: 16 sessions | | | |
| Liddle et al., 2004 | SUD diagnosis | 224 | 13–18 years 81% male 72% African American | a. MDFT; b. CBT-I: | | | |
| Liddle et al., 2004 | Not specified | 80 | 11–15 years, 73% males 42% Hispanic 38% African American 11% Hait/Jamaican | a. MDFT: ~24 hr  
b. CBT-G: ~24 hr | N = specified | Yes | Ratings not reported |
<table>
<thead>
<tr>
<th>Study</th>
<th>Years</th>
<th>Gender</th>
<th>Ethnicity</th>
<th>Treatments</th>
<th>Contact</th>
<th>Outcome</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robbins et al. (2007)</td>
<td>13–17</td>
<td>82% male</td>
<td>100% Hispanic</td>
<td>SET: 12–16 family, 12 eco-systemic, FAM: 12–16 family sessions, SAU: 3 MA, 1 PhD-level Hispanic Therapists across 2 sites</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Santisteban et al., 2003</td>
<td>12–18</td>
<td>75% Male</td>
<td>100% Hispanic</td>
<td>BSFT: 4–20 sessions; M = 11.2, SD = 3.8, Group therapy</td>
<td>BSFT only</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Smith et al., 2006</td>
<td>12–18</td>
<td>71% male</td>
<td>8% African American 76% White, non-Hispanic 5% Hispanic</td>
<td>SOFT: 10 multifamily, 5 conjoint 2-hr sessions, 7C: 10 group sessions (2hr) and 5 individual sessions (1hr)</td>
<td>3 MA-level therapists in SOFT 2 MA and 2 BA-level therapists in 7C</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Stanton et al. (2007)</td>
<td>13–17</td>
<td>75% male</td>
<td>13% African American 83% White, non-Hispanic 2% Hispanic</td>
<td>TFT: 12 hr, AGT: 12 90-min sessions + 2 follow-up sessions as needed</td>
<td>Experienced community based therapists for both models</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Waldron, Slesnick, et al., 2001</td>
<td>13–17</td>
<td>79% male</td>
<td>45% White, non-Hispanic 45% Hispanic</td>
<td>FFT: 12 hr, FFT + CBT-I: 24 hr, CBT-G: 12 hr, CBT-I: 12 hr</td>
<td>5 therapists; 2 doctoral level</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Waldron et al., 2005b</td>
<td>13–19</td>
<td>79% male</td>
<td>45% White, non-Hispanic 45% Hispanic</td>
<td>FFT: 14 hr, IBFT: 14 hr, CBT-G: 14 hrd. CBT-I: 14 hr</td>
<td>8 Therapists; All master’s level</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Hops et al., 2007; Waldron et al., 2007</td>
<td>13–19</td>
<td>79% male</td>
<td>45% White, non-Hispanic 45% Hispanic</td>
<td>IBFT: 14 hr, CBT-I: 14 hr</td>
<td>13 therapists across 2 sites</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Winters et al., 2000</td>
<td>12–18</td>
<td></td>
<td></td>
<td>12-step, Minnesota Model, Treatment milieu: psychologists, counselors &amp; others</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

Note: tx = treatment; SUD = substance use disorder; CD/ODD = conduct disorder/oppositional defiant disorder; CBT-I = cognitive behavioral therapy-individual; BFT = behavioral family therapy; MI + CBT = motivational interviewing plus CBT; CBT-G = CBT-group; MET/CBT12 = motivation enhancement therapy/CBT-12 sessions; MET/CBT5 = MET/CBT-5 sessions; ACRA = adolescent community reinforcement approach; MDFT = multidimensional family therapy; MST = multisystemic therapy; SET = structural ecosystems therapy; FAM = family therapy; SAU = service as usual; BSFT = behavioral strategic family therapy; SOFT = strengths oriented family therapy; 7C = seven challenges; TFT = transition family therapy; AGT = adolescent group therapy; FFT = functional family therapy; IBFT = integrated behavioral and family therapy.
the Adolescent Community Reinforcement Approach (ACRA), and the Seven Challenges (7C). The 7C model combines individual and group based interventions (Schwebel, 2004; Smith et al., 2006). A combined group and individual therapy condition was examined in one study. The Minnesota Model 12-step intervention approach was also evaluated (Winters et al., 2000).

ADOLESCENT SUBSTANCE ABUSE TREATMENT STUDIES

Summary of Outcome Findings

Table 2 summarizes the treatment outcome findings for the studies reviewed. As expected for Type 1 and Type 2 studies, the studies had many strengths that represent significant improvements over previous studies, including the following: (a) comparing well-implemented active treatments to one another, (b) evaluating interventions with a clearly defined target population that mirrors the population found in regular practice settings, (c) using treatment manuals as well as therapist training and monitoring procedures, and (d) assessing subsequent treatment and symptom status during follow-up. As previously noted, the ethical standard in adolescent substance abuse treatment research often limits the use of comparison conditions in which treatment is withheld, and this practice poses a challenge for establishing the empirical support for interventions. Still, confidence in the within-group effect sizes is increased by the meta-analysis findings we present below and by virtue of other research in the field. For example, the notion that improvements associated with adolescent substance abuse treatment may represent regression to the mean is inconsistent with findings from the Services Research Outcome Study (SAMHSA, 1998), which showed, overall, that adolescents had an increase in substance use following treatment. Moreover, the differential efficacy of treatments across multiple studies (e.g., Henggeler et al., 1999; Liddle et al., 2001; Waldron, Slesnick, Brody, Turner, & Peterson, 2001) provides evidence that the reductions in substance use were a direct function of the treatments clients received, rather than an artifact of the passage of time or involvement in a clinical trial.

In classifying intervention approaches as well-established or probably efficacious, we have deviated from the criteria established by Chambless et al. (1996) by relying on the findings of the meta-analysis, presented next, rather than looking at individual comparisons within each of the studies. Thus, in determining whether an intervention is superior to placebo, no-treatment control, or alternative treatment, as required for well-established intervention, we are comparing individual interventions to the pooled minimal treatment control condition previously above.

Given the wide variation in samples, measures, treatment modality, intervention approaches, study procedures, research sites, and analytic methods, the outcomes are remarkably consistent. Overall, the studies reviewed provide additional support for a number of distinct treatment approaches. The remainder of this section focuses on the empirical support for model-specific interventions within these broad treatment types.

The studies evaluating BSFT in this review, taken together with the early studies (Szapocznik et al., 1983, 1986), provide collective evidence that BSFT is probably efficacious for adolescent substance use disorders. BSFT has been associated with significant and clinically meaningful reductions in substance use and has been replicated by the developers using rigorous methods. The sample characteristics are clearly specified. Moreover, the large clinical literature devoted to BSFT, along with a treatment manual in the public domain, provides excellent model specification. In addition, the focus on Hispanic families in these studies makes a unique and vital contribution to the field. Models similar to BSFT also have been replicated by independent research teams in randomized trials (e.g., Joanning et al., 1992), with careful implementation and adequate outcome measures. Although the inclusion of such similar models in terms of empirical support is questioned by Chambless and Hollon (1998), these variations are probably no greater than across those interventions in the universally labeled CBT. Still, despite the many strengths in studies of BSFT, we classified BSFT as probably efficacious rather than well-established because of relatively small effect size (0.19) in the current evaluation (Santisteban et al., 2003), the small sample sizes across studies, and the limited (i.e., comparison of two family-based conditions only) comparison conditions. However, BSFT is currently being evaluated (Feaster, Robbins, Horigian, & Szapocznik, 2004; Robbins, Bachrach, & Szapocznik, 2002; Szapocznik & Robbins, 2002) as part of the National Drug Abuse Clinical Trials Network (2006; NIDA Director, 2002) and the results of this evaluation will address limitations in previous studies. If BSFT is shown to be superior to the usual care comparison, the model could shift to the well-established category.

BSFT can also be considered a probably efficacious family therapy approach. Azrin et al. (2001) replicated an earlier randomized trial (Azrin, Donohue, et al., 1994; Azrin, McMahon, & Donohue, 1994), which demonstrated that BFT is associated with significant reductions in substance use as well as improvements in psychiatric functioning and school performance. Although the sample sizes of these studies were relatively small, the quality of the substance use measures, the manual-guided treatments, and the adherence measures are all strengths of these studies. Additional replication
<table>
<thead>
<tr>
<th>Study</th>
<th>Treatment Conditions</th>
<th>Outcome Variable &amp; Measure</th>
<th>ES (Treatment)</th>
<th>ES (Pre-FU)</th>
<th>Key Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Azrin et al., 2001</td>
<td>a. CBT-I</td>
<td>Urine drug screen</td>
<td>a. 0.49</td>
<td>a. 0.59</td>
<td>Both groups showed significant reductions in substance use and related problem behaviors; improvements maintained at 6-month follow-up</td>
</tr>
<tr>
<td></td>
<td>b. FBT</td>
<td>Timeline Followback</td>
<td>b. 0.47</td>
<td>b. 0.50</td>
<td></td>
</tr>
<tr>
<td>Battjes et al., 2004</td>
<td>a. MET + CBT-G</td>
<td>Global Appraisal of Individual Needs</td>
<td>na</td>
<td>na</td>
<td>Significant reductions in marijuana use in both group treatment conditions, largely sustained at 12-month follow-up</td>
</tr>
<tr>
<td></td>
<td>b. Counseling</td>
<td></td>
<td>na</td>
<td>na</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Overview + CBT-G</td>
<td></td>
<td>na</td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>Dennis et al., 2004</td>
<td>Trial 1</td>
<td>Urine drug screen</td>
<td>a. 0.40</td>
<td>a. 0.40</td>
<td>All treatment conditions in both trials showed significant pre-post reductions in substance use, with some relapse at 12-month follow-up; similar outcomes across sites; economic analysis favored ACRA, followed by MET/CBT5</td>
</tr>
<tr>
<td></td>
<td>a. MET/CBT5</td>
<td>Global Appraisal of Individual Needs</td>
<td>b. 0.20</td>
<td>b. 0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b. MET/CBT12</td>
<td>Collateral reports</td>
<td>c. 0.47</td>
<td>c. 0.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c. FSN</td>
<td></td>
<td>d. 0.49</td>
<td>d. 0.65</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trial 2</td>
<td></td>
<td>e. 0.46</td>
<td>e. 0.46</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d. MET/CBT</td>
<td></td>
<td>f. 0.45</td>
<td>f. 0.90</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e. ACRA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>f. MDFT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Henggeler et al., 1999;</td>
<td>a. MST</td>
<td>Urine drug screen</td>
<td>a. 0.38</td>
<td>a. 0.34</td>
<td>Youths in MST reported significantly greater reduction in drug use compared to control</td>
</tr>
<tr>
<td>Henggeler et al., 2002</td>
<td>b. Usual care</td>
<td>Self report</td>
<td>b. 0.04</td>
<td>b. 0.04</td>
<td></td>
</tr>
<tr>
<td>Kaminer et al., 1998; Henggeler et al., 2002</td>
<td>a. CBT-G</td>
<td>Self Report</td>
<td>a. 1.08</td>
<td>na</td>
<td>CBT-G demonstrated greater reduction in substance use compared to control</td>
</tr>
<tr>
<td>Kaminer &amp; Burleson, 1999</td>
<td>b. OtherGroup</td>
<td></td>
<td>b. 0.36</td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>Kaminer et al., 2002</td>
<td>a. CBT-G</td>
<td>Self Report</td>
<td>a. 1.20</td>
<td>a. 1.20</td>
<td>CBT-G demonstrated greater reduction in substance use compared to control</td>
</tr>
<tr>
<td></td>
<td>b. Psycho-ed group</td>
<td></td>
<td>b. 0.54</td>
<td>b. 0.91</td>
<td></td>
</tr>
<tr>
<td>Liddle et al., 2001</td>
<td>a. MDFT</td>
<td>Urine drug screen</td>
<td>a. 1.34</td>
<td>a. 1.28</td>
<td>Significant improvement in all conditions with MDFT associated with most improvement</td>
</tr>
<tr>
<td></td>
<td>b. AGT</td>
<td>Categorical self-report</td>
<td>b. 0.54</td>
<td>b. 0.95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c. MEI</td>
<td>Collateral reports;</td>
<td>c. 0.80</td>
<td>c. 0.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>composite measure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liddle et al., 2003</td>
<td>a. MDFT</td>
<td>Urine drug screen</td>
<td>a. 0.46</td>
<td>a. 0.74</td>
<td>MDFT demonstrated greater improvement in drug use compared to CBT-I</td>
</tr>
<tr>
<td></td>
<td>b. CBT-I</td>
<td>Categorical self-report</td>
<td>b. 0.16</td>
<td>b. 0.44</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Collateral reports;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>composite measure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liddle et al., 2004</td>
<td>a. MDFT</td>
<td>Urine drug screen</td>
<td>a. 0.50</td>
<td>a. 0.50</td>
<td>MDFT and CBT-G both showed significant reductions in substance use at post and follow-up; small between group effect size</td>
</tr>
<tr>
<td></td>
<td>b. CBT-G</td>
<td>Categorical self-report</td>
<td>b. 0.36</td>
<td>b. 0.44</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Collateral reports;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>composite measure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robbins et al., 2007</td>
<td>a. SET</td>
<td>TLFB, ADAD, urine screens</td>
<td>a. 0.60</td>
<td>a. 0.47</td>
<td>SET showed significantly greater reduction in drug use than FAM or control</td>
</tr>
<tr>
<td></td>
<td>b. FAM</td>
<td></td>
<td>b. 0.18</td>
<td>b. –0.28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c. Control</td>
<td></td>
<td>c. 0.02</td>
<td>c. 0.18</td>
<td></td>
</tr>
<tr>
<td>Santisteban et al., 2003</td>
<td>a. BSFT</td>
<td>Urine drug screen</td>
<td>a. 0.22</td>
<td>na</td>
<td>BSFT superior to group in reducing marijuana use and related problems</td>
</tr>
<tr>
<td></td>
<td>b. Group</td>
<td>Addiction Severity Index</td>
<td>b. –0.03</td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>Smith et al., 2006</td>
<td>a. 7C</td>
<td>Global Appraisal of</td>
<td>a. 0.46</td>
<td>a. 0.45</td>
<td>Both treatment conditions showed significant pre-post reductions in marijuana use</td>
</tr>
<tr>
<td></td>
<td>b. SOFT</td>
<td>Individual Needs</td>
<td>b. 0.48</td>
<td>b. 0.50</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
TABLE 2

Continued

<table>
<thead>
<tr>
<th>Study</th>
<th>Treatment Conditions</th>
<th>Outcome Variable &amp; Measure</th>
<th>ES (Treatment)</th>
<th>ES (Pre-FU)</th>
<th>Key Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stanton et al. 2007</td>
<td>a. TFT</td>
<td>Urine drug screen</td>
<td>a. 0.48</td>
<td>a. 0.50</td>
<td>Both group and family treatments associated with reduced marijuana use. TFT had greater reduction than AGT at long-term follow-up</td>
</tr>
<tr>
<td></td>
<td>b. AGT</td>
<td>TLFB; % days use</td>
<td>b. 0.46</td>
<td>b. 0.45</td>
<td></td>
</tr>
<tr>
<td>Waldron, Slesnick, et al., 2001</td>
<td>a. FFT</td>
<td>Urine drug screen</td>
<td>a. 0.79</td>
<td>a. 0.34</td>
<td>FFT, CBT-G, and IBFT all showed significant reductions in substance use at post and/or follow-up; FFT and IBFT superior to CBT-I</td>
</tr>
<tr>
<td></td>
<td>b. CBT-I</td>
<td>TLFB; % days use</td>
<td>b. 0.00</td>
<td>b. 0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c. CBT-G</td>
<td></td>
<td>c. 0.29</td>
<td>c. 0.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d. IBFT</td>
<td></td>
<td>d. 0.43</td>
<td>d. 0.46</td>
<td></td>
</tr>
<tr>
<td>Waldron et al., 2005</td>
<td>a. FFT</td>
<td>Urine drug screen</td>
<td>a. 0.73</td>
<td>a. 0.41</td>
<td>FFT, CBT-G, and IBFT all showed significant reductions in substance use at posttreatment and follow-up</td>
</tr>
<tr>
<td></td>
<td>b. CBT-I</td>
<td>TLFB; % days use</td>
<td>b. 0.61</td>
<td>b. 0.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c. CBT-G</td>
<td></td>
<td>c. 0.25</td>
<td>c. 0.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d. IBFT</td>
<td></td>
<td>d. 0.33</td>
<td>d. –0.04</td>
<td></td>
</tr>
<tr>
<td>Hops et al., 2007;</td>
<td>Hispanic</td>
<td>Urine drug screen</td>
<td>a. 0.20</td>
<td>a. 0.04</td>
<td>Improvement in marijuana use greater for Hispanics in IBFT versus CBT-I Improvement in marijuana use for Anglos in both treatments</td>
</tr>
<tr>
<td>Waldron et al., 2007</td>
<td>a. CBT-I</td>
<td>TLFB; % days use</td>
<td>b. 0.38</td>
<td>b. 0.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b. IBFT Anglo</td>
<td></td>
<td>c. 0.34</td>
<td>c. 0.66</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c. CBT-I</td>
<td></td>
<td>d. 0.38</td>
<td>d. 0.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d. IBFT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winters et al., 2000</td>
<td>a. MM 12 Step</td>
<td>Urinalysis Self-report</td>
<td>a. 0.55</td>
<td>na</td>
<td>Treatment completion associated with improved outcomes in MM 12 Step</td>
</tr>
<tr>
<td></td>
<td>b. Waitlist (8 weeks)</td>
<td>drug use PEI</td>
<td>b. 0.17</td>
<td>na</td>
<td></td>
</tr>
</tbody>
</table>

Note: We chose not to include some of the data from the Robbins et al. (in press). The authors raised concerns about the validity of these findings due to the high differential (across condition) attrition rate for the African American participants. ES = effect sizes; CBT-I = cognitive behavioral therapy-individual; MET/CBT12 = motivation enhancement therapy/CBT-12 sessions; MET/CBT5 = MET/CBT-5 sessions; CBT-G = CBT-group; na = not applicable; FSN = family support network; ACRA = adolescent community reinforcement approach; AGT = adolescent group therapy; MEI = multifamily educational intervention; SET = structural ecosystems therapy; FAM = family therapy; TLFB = timeline followback; BSFT = behavioral strategic family therapy; 7C = seven challenges; SOFT = strengths oriented family therapy; FFT = transition family therapy; FFT, CBT-G, and IBFT all showed significant reductions in substance use at posttreatment and follow-up; FFT and IBFT superior to CBT-I.

is needed by an independent investigator to consider BFT a well-established treatment.

Across all studies that have evaluated MDFT, MST, and FFT, these approaches have emerged as well-established treatments. Since 1998, randomized trials involving MDFT have been implemented across independent research teams and across multiple sites (Dennis et al., 2004; Liddle et al., 2001; Liddle, Dakof, Turner, & Tejeda, 2003; Liddle, Rowe, Dakof, Ungaro, & Henderson, 2004), with comparisons to one or more active treatment conditions. Similarly, findings for FFT have been replicated by independent investigators and at multiple sites (Friedman, 1989; Hops et al., 2007; Waldron et al., 2001; Waldron, Ozechowski, Turner, & Brody, 2005; Waldron et al., 2007; see also Waldron & Kaminer, 2004). Studies of MDFT and FFT have involved relatively larger samples, clearly specified sample characteristics, careful measurement, extensive model specification, manually guided treatments, adherence monitoring, and adequate follow-up periods. In addition, FFT has been widely transported into communities across the United States, with independent international replications for juvenile delinquent populations. Significant and clinically meaningful reductions in substance use have been associated with both interventions. Although the Liddle et al. (2001) study had a relatively large pre- to posttreatment effect size, the analysis includes only those youths who completed treatment. In another investigation of MDFT (Dennis et al., 2004), the treatment dropouts and completers were included in the analysis, and the MDFT intervention had comparable effect sizes to other major treatments. Based on the strengths of the investigations, in combination with the results of the meta-analysis presented below demonstrating MDFT and FFT to be superior to the minimal treatment control condition, MDFT and FFT emerge as well-established treatments for adolescent substance use disorders.

MST is another model that has been associated with significant reductions in substance use across studies, with a number of replications conducted by the same research team (Henggeler et al., 1991, Henggeler,
Clingempeel, Brondino, & Pickrel, 2002; Henggeler et al., 1999). MST is also widely recognized as an effective treatment for conduct disorder, and it has been widely disseminated for the treatment of this problem. There has been at least one replication by an independent research team for antisocial youth (Ogden & Halliday-Boykins, 2004). However, because we did not find any independent replications for substance abuse treatments, we classified MST for adolescent substance abuse treatment as probably efficacious. Because conduct disorder and substance abuse are so highly co-occurring problems in these treatment samples, it is likely that in an independent evaluation of MST to demonstrate its effectiveness, MST would emerge as a well-established treatment for substance abuse. It is noteworthy that MST studies are among the only randomized trials to compare an active treatment to a usual care comparison condition.

The MST model is well-specified in the clinical literature, and the investigators have led the field in their careful measurement and reporting of treatment adherence and have examined outcomes over as much as 4 years posttreatment. MST is a widely disseminated, highly transportable intervention as well. The reliance on substance use outcome measures that were relatively insensitive to change (e.g., dichotomous urine drug screens that are important for establishing convergence with self-reports but are relatively insensitive as stand-alone measures; categorical “summary” measures of use) may have limited the ability of the studies to detect MST-related treatment gains for adolescent substance use disorders and may have resulted in smaller effect sizes than might otherwise have been realized. Additional research has now been conducted using the TLFB to measure substance use measures (Henggeler et al., 2006). Replication of findings specifically for substance abusing youth by an independent team of researchers is still needed.

We did not include the Henggeler et al. (2006) study in our meta-analysis because the research involved much more intensive interventions (100 hr of treatment or weekly urine screens monitored by the court) than any of the other studies reviewed and was systematically different in several ways from the other efficacy and effectiveness trials evaluated. The study is important, however, in that it evaluated MST and contingency management (CM) in the context of a family court and a juvenile drug court for 161 outpatient substance-abusing adolescents. Adolescents were randomly assigned to the following conditions: (a) a family court with usual community services, (b) a drug court also using typical community services, (c) drug court supplemented with MST, and (d) a drug court plus MST supplemented with CM. The findings at the end of the 12-week treatment period indicated that the three drug court conditions had lower drug use at the end of treatment than the family court condition. However, neither MST alone nor MST plus CM improved the substance use outcomes as compared to drug court alone. After 1 year, the two MST conditions had lower drug use than either the family court or drug court alone. Of note is that the effect size for the comparison of MST plus drug court versus drug court alone at 1 year is about the same order of magnitude as the value obtained for the remaining studies of family therapy included in the meta-analysis next.

CBT-I interventions were evaluated in seven studies (Azrin et al., 2001; Dennis et al., 2004; Hops et al., 2007; Liddle et al., 2003; Waldron et al., 2007; Waldron et al., 2005; Waldron, Slesnick, et al., 2001; see also Waldron & Kaminer, 2004). In three studies, CBT-I included MET as a treatment component and in one study CBT-I was based on the ACRA model. Again, despite some variability, six of the seven studies found significant pre- to posttreatment reductions in substance use for CBT-I. The overall quality of the studies, including large sample sizes, manual-guided treatments, sensitive measurement of substance use, treatment adherence ratings, appropriate analyses, and the significant and clinically meaningful reductions in substance use associated with the intervention approach, results in a classification of well-established for CBT-I as a treatment for adolescent substance abuse.

CBT-G interventions were evaluated in 11 randomized trials involving 14 unique study replications (Battjes et al., 2004; Dennis et al., 2004; Hops et al., 2007; Kaminer et al. 1998; Kaminer & Burleson, 1999; Kaminer, Burleson & Goldberger, 2002; Liddle et al., 2001; Liddle et al., 2004; D. C. Smith et al., 2006; Stanton, Rempala, & Conway, 2007; Waldron et al., 2007; Waldron et al., 2005; Waldron, Slesnick, et al., 2001; see also Waldron et al., in press; Waldron & Kaminer, 2004). Some study limitations were noted. For example, two studies had considerable attrition, one that supported CBT-G (Kaminer et al., 1998) and one showing CBT-G fared more poorly than family therapy (Liddle et al., 2001). A number of studies implementing CBT-G were methodologically rigorous Type I studies (Nathan & Gorman, 2002), in which significant and clinically meaningful reductions in substance use were achieved. Although the number of CBT-G sessions varied considerably across studies and research sites, with number of sessions ranging from 5 to 19 sessions, the CBT-G therapy manuals used in studies examining interventions with 12 or more sessions (e.g., Dennis et al., 2004; Kaminer et al., 1998; Waldron, Slesnick, et al., 2001) were highly similar, with the majority of treatment modules shared across programs. In light of the consistent support for CBT-G, the maintenance of gains over relatively long follow-up periods, the multiple replications across independent research teams, the
inclusion of CBT-G in the one effectiveness study conducted (i.e., Dennis et al., 2004) and the high transportability of manual-guided group intervention, we have classified CBT-G as a well-established treatment.

The CBT-G findings for substance-abusing adolescents in the studies reviewed here are consistent with those in other reviews (Maccowan & Wagner, 2005). The results directly contradict earlier claims of iatrogenic effects (Dishion, McCord, & Poulin, 1999; Dishion, Poulin, & Burraston, 2002; Poulin, Dishion, & Burraston, 2001) associated with adolescent group interventions, which have led to questions about the appropriateness of group-based treatments for substance-abusing youth. Specifically, Dishion and others suggested that unplanned, incidental interactions among adolescents, referred to as “deviancy training,” may be more powerful in influencing an adolescent’s future than interactions structured by a group treatment curriculum (Dishion et al., 1999). Deviancy training has been linked to an increased probability of the initiation of substance use, delinquency, and for police-reported violent behavior (Dishion, Spracklen, Andrews, & Patterson, 1996), although the magnitude of the effect of deviancy training on growth in addictive behaviors has been marginal and shown only for youth who are at-risk for substance use (Dishion, Poulin, & Burraston, 2001; Dodge, Dishion, & Lansford, 2006).

To address the potential for iatrogenic effects of group substance abuse treatments, Burleson and Kaminer (2005) examined 400 group therapy participants in the CYT study and found that adolescents with more severe conduct problems benefited from participating in a group with others who had few conduct problems. Moreover, they found that the more severely conduct disordered youth did not appear to influence those youth with lower levels of such problems. This finding runs counter to the deviancy training assertion that highly disordered youth adversely affect youth with lower problem severity (Dishion et al., 1999). Alternatively, the Burleson and Kaminer (2006) study is consistent with a host of others that have not supported iatrogenic or deviancy training effects for group treatments for anger, aggression, and other antisocial behavior (Ang & Hughes, 2001; Hoag & Burlingame, 1997; Lipsey, 2006; Sukhodolsky, Kassionve, & Gorman, 2004; van Manen, Prins, & Emmelkamp, 2004; Weiss et al., 2005).

In recent writings, Dishion and Dodge (2005) concluded that the concept of peer contagion may not apply to all individuals and/or contexts. They noted that a variety of other variables need to be considered, including level of maturity, severity of deviance, and skill of the therapist. They pointed out that recent evidence cited in Waldron and Kaminer (2004) supports the effectiveness of group interventions for adolescents with substance abuse disorders, without evidence for peer contagion. Thus, deviant peer influences may operate more strongly on adolescents who are only marginally deviant, whereas well-adjusted youths may be able to resist deviant peer influences and those youth with severe levels of deviance may be beyond the influence of others (Dishion, Dodge, & Lansford, 2006).

Moreover, Dennis et al. (2004) included an economic analysis in the report of their study findings, which established a cost advantage for CBT-G that was second only to ACRA. Similarly, our economist colleagues examined treatment outcomes in the Waldron et al. (2001) study. The analysis demonstrated that our CBT-G was more cost-effective than other efficacious treatments when evaluated at a 7-month follow-up point (French et al., in press). Based on the evidence available, we conclude that CBT-G is a well-established, empirically supported treatment for adolescent substance use disorders and could be considered a treatment of choice in the context of limited treatment resources. Further research is needed to evaluate the economic benefits of CBT-G on the common comorbid diagnoses that accompany substance abuse.

The study conducted by Winters and his colleagues (2000) represents the first study evaluating the Minnesota Model 12 Step approach for adolescents. Although the study was not a randomized trial, the study used a carefully selected comparison sample of adolescents who were placed on a waitlist to receive treatment. The model is steeped in the tradition of treatments for addictive behaviors, and the evaluation of the model with adolescents is an important addition to the field. Although the model has not yet been studied with the same intensity as other models, the findings from our study suggest that this is a promising approach worthy of additional investigation. Similarly, the Transitional Family Therapy (Stanton, Rempala, & Conway, 2007) and SOFT (D. C. Smith et al., 2006) appear promising and merit further study.

META-ANALYSIS RESULTS

The use of meta-analysis was designed to capitalize on the evidence provided by investigations in which hypothesized “nonactive” interventions were implemented (e.g., usual care, psychoeducation). The meta-analysis examined four questions. First, we assessed whether the pre- to posttreatment effect size index in any of three types of active treatment conditions—CBT-I, CBT-G, or family therapy (FT)—were significantly different from 0.0, a value which reflects no change in drug use. Thus, these analyses determined whether the average effect size across replications within each treatment type was associated with improvement or deterioration in drug use. Second, we assessed whether any of the three types of active treatment conditions...
were significantly different from the minimal treatment control condition in pre- to posttreatment change. Third, we examined the possible role of demographic moderator effects (i.e., age, sex, ethnicity) on the outcomes. Fourth, we assessed whether significant heterogeneity existed among the various replications within each of the treatment modalities (i.e., control, CBT-I, FT, CBT-G) after controlling for demographic moderator variables. This last research question emerged from the recognition that our classification system pooled together a variety of interventions that differ from one another in the specific approach to behavior change for adolescent substance abuse.

On one hand, if the various approaches within each modality were differentially effective, then we would expect significant effect size heterogeneity among the replications within each modality. On the other hand, if the replications within each modality are comparable, such that the heterogeneity test is not significant, then the variability among the replications within a modality would be no greater than would be expected based on the within-treatment variability. In other words, a nonsignificant heterogeneity test would be consistent with the premise that the effect sizes from the various replications within a treatment type are all sampled from the same population.

The current meta-analysis (see appendix for more details) used empirical Bayes estimates within the context of HLM procedures (Raudenbush & Bryk, 2002, chap. 7). A number of different approaches have been advocated for creating an effect size estimate (see Rosenthal, 1994). A common goal of these different approaches is the creation of a standardized index that permits the quantitative synthesis of findings from different studies into a single analysis. By using standardized effect size indexes, the results can be combined, even though the original studies used somewhat different measurement procedures and study samples.

The primary effect size index used in the present analysis was a variation on Cohen’s d (Lipsey, 1990; Lipsey & Wilson, 2001; Morris & DeShon, 2002; Raudenbush & Bryk, 2002; Rosenthal, 1994). This index is usually calculated as the mean difference between a treatment and a control condition divided by an estimate of the standard deviation. Nearly all of the studies in this review compared two or more active treatment conditions. For the reasons noted previously, few of the studies had a no-treatment control condition (Stanton & Shadish, 1997; Szapocznik et al., 1986). Moreover, because few of the studies included the same combinations of active treatments, the specific type of between-treatment comparisons varied substantially across studies. To create a standardized effect size for each of the 46 unique replications, we used the pretreatment drug use mean and standard deviation for each study to control for initial differences in these studies.

By using the pretreatment standard deviation, we were also able to create an estimate of the size of the pre-to-post change relative to the variability that existed among the as-of-yet untreated sample for each replication (see Lipsey & Wilson, 2001; Morris & DeShon, 2002).

A composite effect size was calculated for each of 46 study conditions, and the index combined both a treatment effect and a posttreatment follow-up effect. The treatment effect was calculated as the mean difference between baseline and an observation point near the end of treatment, divided by the baseline standard deviation. The posttreatment effect was calculated as the mean difference between baseline and the first posttreatment follow-up observation point (usually about 3 months after treatment initiation), divided by the baseline standard deviation. The composite index was the average of the two effect sizes. Because the results for the treatment effect and the posttreatment effect size are very similar, we report only the findings for the composite index. This composite effect size provided a global index reflecting change during treatment and immediately following treatment. The individual studies used somewhat different measures of drug use, and the research participants had different initial levels of drug use.

Rosenthal (1994, p. 23), described circumstances in which only the standard deviation of the control condition is used in the effect size index (also see Lipsey & Wilson, 2001; Morris & DeShon, 2002). For example, the control condition might be preferred if the experimental treatment could create a standard deviation that is too small or too large relative to the control condition. By computing effect sizes based on the difference between baseline and subsequent observation points, we were able to control for initial differences among replications in the mean levels of drug use. We divided these mean differences by the baseline standard deviations, in order to control for possible differences in the way that the various scales of measurement were created for each study (Morris & DeShon, 2002). For example, one scale may be based on numbers of days use, either in the past 30 days or the past 90 days. Another scale may be based upon percentage of days of use.

Using the baseline standard deviation, each pre-to-post difference was converted into a standard metric similar to a z score. We used only the baseline standard deviation in our calculations, so that the effect size reflected change relative to the pretreatment variability that exists in drug use for each replication. When a second follow-up assessment was not available, we used Expectation-Maximization imputation procedures to estimate means. See Table 2 for a summary of the treatment outcome findings for the studies reviewed.

Hypothesis testing. The first hypotheses to be evaluated were (a) whether the mean of the minimal
treatment control effect size was significantly different from 0.0, and (b) whether the between-replication (level-2 in HLM terms) differences can be explained, in part, by the types of active treatment condition (minimal treatment control, CBT-I, FT, and CBT-G) implemented in each replication. In this analysis, we pooled all three active treatment conditions together and compared the pooled studies to the minimal treatment control condition on the composite effect sizes index. The results of this analysis indicated the following: First, the pre- to posttreatment changes in the minimal treatment control conditions (M = 0.19) were significantly different from 0.00 (B = .19), t(44) = 2.00, p < .05. Second, this minimal treatment control condition effect size was significantly lower than the mean effect size (M = 0.45) of the three active treatment conditions (B = 0.26), t(44) = 2.60, p < .01. Although the minimal treatment control conditions revealed a significant reduction in marijuana use, the findings suggest that the active treatment conditions produced a significantly greater reduction in marijuana use. The results also indicated that significant residual heterogeneity existed among the effect size estimates (V = 0.16), \( \hat{\chi}^2(45) = 73.82, p < .005 \), indicating an unexplained source of variation among the effect sizes. The next analyses examined this heterogeneity.

One possible explanation for the unexplained variation is that the three active treatments were not all similarly effective in reducing marijuana use as compared to the minimal treatment control condition. To assess this possibility, we created three dummy vectors that contrasted each active treatment condition (CBT-I, FT, or group therapy) with the minimal treatment control condition replications. The control condition was coded as 0.00 in each dummy vector. We also tested the hypothesis that any additional heterogeneity of effects remaining after controlling for the treatment conditions could be explained by demographic characteristics of the treatment samples (e.g., age and sex of adolescents). In addition to covariates that assessed the average age of each replication and the percentage of males/females, we also included a Level 2 covariate, which reflected the percentage of the sample that was of Hispanic origin. This variable was included because an inspection of the demographic characteristics of samples revealed great variability (M = 29.2%, SD = 35.3%) with respect to this attribute, which ranged from 0% to 100%, whereas little variability existed on other demographic variables (e.g., percentage of the sample that was non-Hispanic White or African American).

These three contrast vectors and the three covariates were added simultaneously to the HLM random effects model, and the results of the analysis indicated that the dummy contrasts for family, (\( \gamma = 0.31 \), t(41) = 2.87, p < .007, and for CBT-G, (\( \gamma = 0.22 \), t(41) = 1.99, p < .05, were statistically significant, but CBT-I (\( \gamma = 0.21 \), t(41) = 1.67, p < .10, was not significantly different from the minimal treatment control condition.\(^1\)

Furthermore, the Hispanic covariate was significant (\( \gamma = -0.34 \), t(41) = 3.11, p < .004. The age and sex covariates were not, however, significant. The results for the Hispanic covariate indicate that the distinct replications with a higher percentage of Hispanic participants had smaller effect sizes. Finally, the test for heterogeneity of effect sizes was not statistically significant (V = 0.11), \( \hat{\chi}^2(41) = 55.33, p < .07 \), after controlling for the proportion of Hispanic participants.

We explored other possible explanations, such as the percentages of the treatment samples that were non-Hispanic White or African American origin, in addition to the male:female proportion or the average age of adolescents. None of these variables accounted for the residual variation among the replications: adolescent age (\( \gamma = -0.001, SE = 0.001 \), sex (\( \gamma = -0.008, SE = 0.007 \), percentage of sample that was non-Hispanic White, (\( \gamma = -0.000, SE = 0.002 \), or African American (\( \gamma = -0.001, SE = 0.002 \). We recognized that the differential effects of the Hispanic covariate might occur because the proportion of Hispanics in each replication might differ across the treatment modalities. To assess this possibility, we conducted a multivariate analysis of variance with treatment condition as the independent variable, and each demographic variable as one of multiple dependent variables. The results indicated that the treatment conditions did not differ significantly across this set of dependent variables using Wilks’s lambda criteria, \( MVA(15, 105.3) = 0.72, p = .76, \eta^2 = 0.085 \). Thus, the findings are not consistent with an explanation based on differential proportions of Hispanic youth in the treatment modalities.

**Hispanic proportion within treatment conditions on treatment effect size.** A final analysis explored the differential association of the Hispanic covariate with changes in each of the treatment conditions. The results demonstrated that the covariate was significant for the CBT-G condition (\( \gamma = -0.79 \), t(11) = 2.96, p < .01, indicating that studies with more Hispanic participants had smaller effect sizes across the CBT-G replications. The Hispanic covariate was not significant in family

\(^1\)The current analysis did not control for initial within-replication differences in drug use, as we did not have a direct estimate for each study condition. However, if we had statistically controlled for these initial differences (assuming a within replication correlation of .55, a value which was based on our analysis of more than 500 adolescents receiving similar substance abuse treatments), the between treatment modality would have been increased by approximately 30%. We derived the .55 correlation from a pooled estimate from our own study of nearly 500 adolescents in substance abuse treatment programs. As a consequence, the family, group CBT, and individual CBT modalities would each be significantly different from the minimal treatment control condition.
Test for residual heterogeneity. The meta-analysis procedure involves the pooling together of various forms of treatment conditions within each of the major treatment modalities. Furthermore, the originators of these treatment variants advocate somewhat different mechanisms of behavior change, and these could have quite different effects on outcomes. By pooling these studies together, we could be obscuring the differential effectiveness of these treatments. One method for assessing the premise that these variations can be pooled together is to test the heterogeneity of treatment effects within each treatment modality. The test for heterogeneity of effects was significant for the CBT-G modality ($V = 0.18$), $\chi^2(11) = 21.40, p < .02$, suggesting that some additional unexplained source of variance existed among CBT-G replications. However, the heterogeneity test was not significant in the family ($V = 0.004$), $\chi^2(15) = 15.51, p < .42$; the CBT-I ($V = 0.06$), $\chi^2(5) = 4.37, p > .50$; or the control conditions ($V = 0.16$), $\chi^2(7) = 9.01, p < .25$.

We have estimated the effect sizes assuming that the percentage of Hispanics in the sample was either 0%, 25%, 50%, or 75%, with results presented in Figure 1. An inspection of the figure indicates that the group condition is one of the most effective treatments when the study contains a small percentage of Hispanic participants but it is one of the least effective with a relatively large percentage of Hispanic participants.

Although it is always difficult to interpret a nonsignificant finding, our results are consistent with the premise that the variability in treatment outcomes within three of the treatment modalities (control, CBT-I, and FT) can be explained by random sampling principles. Although the distinct implementations within each modality do not appear to produce different outcomes on the substance use measure, these variants may have quite different effects on other problem behaviors such as conduct disorder and delinquency. Moreover, the present study focused only on relatively brief follow-up periods. The treatment variants may have quite different effects in terms of the long-term maintenance of drug abuse reductions.

The variation among CBT-G replications appears to involve other sources of unexplained variability beyond the effect that can be attributed to the Hispanic sample proportion. We know that the CBT-G conditions may vary on a number of factors other than the proportion of the sample that is of Hispanic origin. For example, the CBT-G treatments are likely to involve a mixture of different ethnic groups within each therapy group. When these mixtures occur, the group leader may find it difficult to address the differences that may occur among the group members in their cultural values and the sources of their behavior problems. Tension among the ethnic group members also may impede change processes within the groups. Ethnic differences among group participants may negatively influence the outcomes of treatments for these heterogeneous groups. By contrast, the CBT-I, FT, and control conditions do not involve the mixture of treatment samples from different ethnic groups in the same intervention sessions. Of course, ethnic differences between the therapist and the treatment clients may also have an impact on the effectiveness of any of the interventions. Further research is needed to evaluate this and other possible sources of unexplained variability associated with CBT-G interventions.

EVALUATIONS OF TREATMENT MODERATORS

Although randomized controlled trials may provide evidence for treatment efficacy, such studies contribute little to untangling what aspects or features of these approaches may affect outcome (Clingempeel & Henggeler, 2002), and few of the studies reviewed have evaluated either moderators or mediators linked to specific interventions. Moderators are pretreatment characteristics that influence the relation between treatment type or other independent variable and treatment outcome (Kazdin, 2007). Kaminer et al. (1998), for example, found that older male adolescents had better outcomes in CBT-G than in the comparison group treatment condition. Other potential moderators for adolescent substance abuse treatment include sex, co-occurring conditions (e.g., delinquency, comorbid disorders), adolescent motivation for change, parenting and family factors, baseline impairment in coping skills deficits, traumatic life events, opportunity for exposure.
to different situations posing high-risk for relapse, self-efficacy, negative moods, and treatment expectancies. Note that these potential moderating variables may also be mediating mechanisms, if changes from baseline in these variables are linked to treatment outcomes.

Comorbid Disorders

The majority of untreated adolescents with a substance abuse disorder are likely to have a comorbid psychological disorder (Dennis et al., 2004; Flynn, Craddock, Luckey, Hubbard, & Dunteman, 1996; Kaminer & Bukstein, 2007), as well as a history of physical, emotional, or sexual victimization (Dennis et al., 2004). Kaminer and Waldron (2006) noted that these co-occurring conditions may influence the onset, identification, course, and treatment of substance abuse problems. Substance using youths who also display conduct disorders are at increased risk of not completing treatment, with lower retention rates linked to poorer treatment response (Kaminer, Burleson, Goldston, & Burke, 2006; Kaminer, Tarter, Bukstein, & Kabene, 1992; Myers, Stewart, & Brown, 1998). Findings in the literature also include some inconsistencies. For example, Kaminer et al. (1992) found poorer outcomes for individuals whose comorbid status does not include either depression or anxiety disorders, and another study found that youth with both externalizing and internalizing disorders have the poorest outcomes (Rowe, Liddle, Greenbaum, & Henderson, 2004). One study also suggested that CBT may be especially beneficial to adolescents with a dual diagnosis of a substance use disorder and depression (Birmaher et al., 2000). Comorbidity in adolescent substance abuse is a critical area for further study (Kaminer et al., 2006).

Treatment Motivation

A pervasive problem in the treatment of adolescent substance abuse is the risk that these individuals will not complete the treatment programs to which they have been assigned. Many of these adolescents are in treatment because of external pressures from the school or juvenile justice systems. They typically have low motivation to change and they do not perceive treatment as suitable for themselves (Jainchill, Hawke, DeLeon, & Yagelka, 2000). Although several interventions have specific motivational components (e.g., MET/CBT5, MET/CBT12, FFT), no studies have yet examined treatment motivation as a moderator of treatment outcome for the intervention approaches included in this review. Because motivation is an important target of change in most of the interventions described in this review, changes from baseline on this variable also may be an important mediator of changes.

Ethnicity

The studies reviewed varied markedly in terms of the ethnic or racial composition of the treatment samples. Collectively, however, the total sample is relatively diverse, including approximately 45% non-Hispanic Whites, 25% Hispanics, 25% African Americans, and 5% others. These findings generally seem to indicate that the treatments are effective for the three largest ethnic groups represented in the treatment samples. Although replication is needed, some findings suggest that family based interventions may be more efficacious than CBT interventions for Hispanic participants (Hops et al., 2007; Waldron et al., 2007). Moreover, in one study of therapist-client ethnic matching, Hispanic adolescents with Hispanic therapists demonstrated greater reductions in substance use than those with Anglo therapists (Flicker, Waldron, Turner, Brody, & Hops, in press). Ethnic match status was not related to treatment outcome for Anglo clients. In addition, the results of the aforementioned meta-analysis suggest that group-based interventions may not be as efficacious as other interventions for Hispanic youth. Based on the findings presented, ethnicity and therapist-client ethnic matching may emerge as important moderators of substance abuse treatment outcome.

EVALUATION OF TREATMENT MEDIATORS

Mediators are intervening variables that may account for the relationship between an intervention and treatment outcome (Kazdin, 2007). No studies published to date have been able to provide clear links between mediators of change and outcomes for adolescent substance abuse treatments. One unknown is whether similar mediating factors operate across models or whether there may be multiple pathways to change. For example, although different family therapy approaches describe many techniques and procedures that are unique to a particular model, family therapists may display essentially the same behaviors across models. Therapy process research has shown promise in isolating therapist, client/family, and interactional mediators of outcome for family therapy (Alexander, Holtzworth-Munroe, & Jameson 1994), including family mediators of change in FFT with delinquent youth (Robbins, Alexander, & Turner, 2000) and family mediators of change in MST with delinquent youth (Huey, Henggeler, Brondino, & Pickrel, 2000). Clearly, research on mediators of substance abuse treatment is still in its infancy.

Some research has also provided clues to mechanisms of change associated with CBT interventions. Myers and Brown (1990a, 1990b), for example, found that after receiving CBT, adolescent alcohol abstainers and minor relapers were more likely to utilize problem-solving coping strategies than were major relapers. Further, youth
coping behaviors have been identified as significant predictors of treatment outcome (Myers, Brown, & Mott, 1993). Thus, the improvement of these skills may be a particularly important mediating variable in CBT. However, a limitation of most youth treatment studies employing CBT is a lack of empirical measurement of deficits in coping skills and changes in coping skills, compared to self-efficacy ratings and treatment outcome. In fact, no studies of adolescent substance abuse treatment have examined critically the assumption that coping skills are actually acquired or enhanced in treatment. There is a need to determine to what extent outcome of adolescent treatment is a function of acquired or improved coping skills, and whether coping skills acquisition is a function of specific treatment approaches. It is plausible that one, several, or an interaction of mediators, such as readiness to change, expectancy, therapeutic alliance, and engagement in treatment, are responsible for change in self-efficacy and, in turn, reductions of substance use.

Recent investigations have examined the role of therapeutic alliance on treatment engagement and outcomes (Diamond et al., 2006; Hogue, Dauber, Stambaugh, Cecero, & Liddle, 2006; Robbins, Turner, Alexander, & Perez, 2003; Robbins et al., 2006; Shelef, Diamond, Diamond, & Liddle, 2005) within different treatment approaches. These studies demonstrate that early alliance formation predicts the likelihood that adolescents will stay in treatment and that they will have better outcomes on drug abuse, internalizing, and externalizing behaviors (Diamond et al., 2006; Hogue et al., 2006).

**PRACTICE GUIDELINES**

As noted, Chambless et al. (1996) specify criteria for classifying treatment approaches as well-established or as probably efficacious. These criteria require manual-guided treatment and clearly specified sample characteristics. A well-established intervention, however, specifically requires at least two well-conducted group-design studies, conducted by different investigative teams. The research must show that the treatment is either superior to a placebo or to alternative treatment, or that the treatment is equivalent to an already established treatment. To be classified as probably efficacious, an intervention requires two studies showing the intervention to be more effective than a no-treatment control group or two group-design studies that would meet criteria for well-established interventions but conducted by the same investigator. We used the Chambless et al. (1996) criteria for classifying treatment approaches in this review, except that we used the findings from a meta-analysis (instead of between-group comparisons) to evaluate whether specific treatments were superior to a minimal treatment control condition derived from the pooled findings from the studies included in the meta-analysis (i.e., approximately half) that implemented a usual care condition or other minimal treatment control.

Based on the quality of the studies and replications in the field, two family-based approaches, MDFT and FFT, as well as group CBT, are well-established for adolescent substance abuse treatment. Other family models, including MST, BSFT, and BFT, are probably efficacious, pending replications by independent research teams. ACRA and other individual CBT approaches appear promising, but additional research is needed. Despite the collective evidence, however, no clear pattern emerged for the superiority of one treatment model over another. As noted in the previous discussion of treatment mediators and moderators, specific treatment models may be differentially efficacious for particular subgroups of youth and/or may be associated with different patterns of outcomes across different domains of functioning. More research is needed to identify which adolescents may be more likely to respond to specific interventions and how treatments can be adapted or tailored to the individual needs of adolescents to improve substance use outcomes. Until these issues have been addressed more adequately, clinicians have the flexibility to choose from among the well-established treatments, depending on how the approaches fit within their current treatment environments and staffing resources.

**FUTURE RESEARCH DIRECTIONS**

Apparent from this review, a number of controlled clinical trials have demonstrated the efficacy of several CBT and family therapy approaches for adolescent substance abuse. It should be noted, however, that the mean within-subject effect size for the active treatments under consideration ($M = .45$) is not quite medium in magnitude, revealing that the treatments under consideration have not, on average, produced large effects on substance-abuse-related outcomes. In addition, individual adolescent outcomes vary widely within the intervention models, and because no one treatment approach appears to be similarly efficacious for all youth, research focusing on understanding who might benefit from particular treatments is urgently needed.

Several studies have begun to address this variation using cluster analysis and other analytic strategies to identify subgroups of youth who share similar change profiles or treatment response profiles over the course of treatment (Brown et al., 2001; Godley, Dennis, Godley, & Funk, 2004; Rowe et al., 2004; Waldron, Turner, & Ozechowski, 2005, 2006). These unique response trajectories represent a key challenge in the field for understanding treatment moderators and mediators and how treatment may work for subtypes of youth exhibiting distinct profiles and levels of substance use.
and associated behavioral symptoms. Such research could guide the adaptation, modification, or tailoring of treatments to enhance their effectiveness.

One approach to addressing problem associated with variability in treatment response is adaptive, progressive treatments (Collins, Murphy, & Bierman, 2004; Compton & Pringle, 2004; Sobell & Sobell, 2000). Traditional experimental designs have emphasized the comparison of standardized, fixed, interventions as the primary method for evaluating treatment efficacy, using a “one-size-fits-all” perspective. However, the emphasis on fixed interventions does not take into account the breadth of individual differences in the presentation of adolescent substance use disorders that might make aspects of a traditional intervention less relevant, inefficient, or even counterproductive for certain individuals. Problems with attrition and nonadherence may be linked to treatment regimens being suboptimal for some participants. In adaptive research designs, treatment algorithms that vary treatment components and dosages in response to the needs of the individual are used to guide clinical decision making (Collins, McKay, Oslin, Rosenberg, & Murphy, 2006; Collins et al., 2004; Murphy & McKay, 2004; Sobell & Sobell, 2000; Ten Have, Coyne, Salzer, & Katz, 2003). For example, the Sequential Multiple Assignment Randomized Trial outlines a process whereby a series of randomizations within an individual study is used to create an optimal adaptive intervention. Each randomization point is designed to address the question of how best to proceed, given a specific outcome. The goal of the Sequential Multiple Assignment Randomized Trial process overall is to build and refine adaptive treatments using both empirical research and clinical judgment (Collins et al., 2006).

Another area of much needed research is the relation between adolescent substance abuse and brain functioning and how the toxic effects of drug use may influence both the course of treatment and subsequent treatment response outcomes. Structural and functional brain imaging studies have shown the deleterious disruption in neural substrates resulting from drug abuse (e.g., Ernst et al., 2005; Jernigan et al., 2005). Neurotoxicity associated with drug use, as identified by functional magnetic resonance imaging, has been shown to powerfully predict posttreatment response for adult substance abusers (Paulus, Tapert, & Schuckit, 2005). These links may be identified in adolescent substance abusers as well, and the use of neuroimaging procedures could inform treatment planning.

Future research should also include an emphasis on mechanisms of change associated with treatment effectiveness. For example, the scant evidence of differences on family functioning variables between modalities challenges traditional views such as family interaction perspectives on the role of the family in problem substance use. Family functioning and other potential mechanisms of action across CBT and family-based interventions have not been adequately examined. A variety of research designs have been proposed as adjuncts to the traditional randomized controlled trial for the purpose of illuminating the core components and mechanisms of efficacious treatments (Kazdin, 2003). Among such designs is the dismantling strategy (Barrios & Hartmann, 1988; Kazdin, 2003; McFall, 1978; Strayhorn, 1987).

Briefly, a dismantling design strategy compares a total treatment package to one or several reduced versions of treatment in which a single component has been eliminated. A difference in outcome between treatment packages in which a particular component is either present or absent suggests a connection between the outcome and that particular component of treatment. Dismantling design strategies are appropriate for addressing research questions regarding the independent contributions to overall clinical change associated with particular components of an efficacious treatment. Such studies may provide a critical pathway leading to the refinement and strengthening of brief interventions for adolescents by enhancing exposure to active components and eliminating inactive components (Clingempeel & Henggeler, 2002; Miranda & Borkovec, 1999; Shirk, 2005). Moreover, the ability to streamline or consolidate treatment could enhance cost effectiveness, facilitate dissemination, and foster the identification of components that are useful across different disorders, which could reduce clinician burden for learning multiple intervention approaches (Shirk, 2005).

Other areas for increased attention include a focus on improving adolescent motivation and engagement techniques, especially for conduct-disordered and other severely disturbed individuals, and enhancing treatments for youth with other comorbid disorders. In addition, the problem of relapse and lack of maintenance of treatment gains make research on continuing care following a course of treatment, as a single modality or part of an integrated approach, an important focus of investigation.

More efforts also need to be made with respect to therapy process to focus more directly on interactions between the therapist and the client(s) that influence change and on linking hypothesized causal mechanisms to outcome. Adolescent substance abuse treatment outcome research is poised to move to the next level of person-centered approaches to treatment. Concerted efforts on key research fronts would provide a potential path for taking this important next step.

REFERENCES

*References marked with an asterisk indicate studies included in the meta-analysis.


Arean, P., & Alvídez, J. (2002). Ethical considerations in psychotherapy effectiveness research: Choosing the comparison group. Ethics and Behavior, 12, 63–73.


Kaminer, Y., Burleson, J. A., & Goldberger, R. (2002). Cognitive-behavioral coping skills and psychocducation therapies for...
adolescent substance abuse. The Journal of Nervous and Mental Disease, 190, 737–745.


NIDA Director. (2002, February). The CTN’s first three waves of research protocols. NIDA Notes. 16(6).


APPENDIX

Hierarchical Linear Modeling Meta-Analysis Model

An important goal of meta-analysis is to model the heterogeneity of effect sizes across the replications. In Hierarchical Linear Modeling terms (Raudenbush & Bryk, 2002, p. 209), the observed effect size statistic $d_j$ (Eq. 1) is a standardized mean difference between a treatment and a control group.

$$d_j = (\overline{Y}_{i,j} - \overline{Y}_{c,j})/s_j.$$  (1)

The observed $d_j$ provide an estimate of the population parameter $\delta_j$, and the subscript $j$ refers to the individual studies with $j = 1, \ldots, J$. For the Level 1 (within studies) model (Eq. 2), the sampling error is represented as $e_j$, and these errors are assumed to be normally distributed $\sim N(0, V)$.

$$d_j = \delta_j + e_j.$$  (2)

The between study replication component is represented with a Level 2 model (Eq. 3) where the unknown effect size, $\delta_j$, is a function of individual study characteristics.

$$\delta_j = \gamma_0 + \gamma_1 W_{1j} + \gamma_s W_{sj} + \mu_j.$$  (3)

The $W_{sj}$ in this model are unique study characteristics such as dummy codes to represent comparisons between treatment conditions or indicators of demographic characteristics reflecting differences between replications. The $\gamma_s$ are regression coefficients for the $W_{sj}$ whereas the $\mu_j$ are Level 2 random errors that are assumed to be approximately normally distributed.