

What Is the Evidence for Evidence-Based Treatments? A Hard Look at Our Soft Underbelly

Peter S. Jensen,^{1,2} Robin Weersing,¹ Kimberly Eaton Hoagwood,¹ and Eliot Goldman¹

In the rising quest for evidence-based interventions, recent research often does not give adequate attention to “nonspecific therapeutic factors,” including the effects of attention, positive regard, and therapeutic alliance, as well as the effects of treatment dose, intensity and actual processes mediating therapeutic change. To determine the extent to which recent clinical trial designs fully this problem, the authors conducted a systematic review of PsychLit/Medline of all controlled child psychotherapy treatment studies from 1995 to 2004. A total of 52 studies were identified that met review criteria: two or more therapy conditions and random assignment of participants to intervention groups. Of the 52 studies, one group ($n = 29$) compared a presumably active treatment with 1 or more similarly intensive treatments (often an attention control group) presumably not containing the active therapeutic ingredients. Of these, 14 studies found evidence of consistent differences between the two groups, whereas 15 did not. An additional group of studies ($n = 27$) compared therapy groups with different levels of intensity and “dose” of the putatively active treatment; 13 of these found evidence of the effects of different levels of treatment dose/intensity on outcomes and 14 did not. Four studies met criteria for inclusion in both groups. Across both groups of studies, when positive effects were found, few studies systematically explored whether the presumed active therapeutic ingredients actually accounted for the degree of change, nor did they often address plausible alternative explanations, such as nonspecific therapeutic factors of positive expectancies, therapeutic alliance, or attention. Findings suggest that many child psychotherapy treatment studies have not inadequately controlled for nonspecific factors such as attention and treatment intensity and have failed to assess specific mediators of change. Specific recommendations for future studies are offered, specifically:

1. Initial specification in study design how investigators will test if the intervention is efficacious over and above the effects of positive expectancies, positive regard, or attention;
2. Planned (rather than post hoc) analyses to explore whether and how a given treatment’s specific ingredients or the overall intensity of its “dose” are related to treatment outcomes;
3. More studies conducting head-to-head tests of different types but equally credible forms of treatment, with planned analyses testing different mechanisms of change
4. When positive effects of a treatment vs. a control are found, systematic elimination of specific, competing hypotheses of reasons for treatment efficacy (e.g., attention, therapeutic alliance, face validity of treatment and client’s treatment expectations, changes in self-efficacy), *and*

¹Center for the Advancement of Children’s Mental Health at Columbia University/NYSPI, 1051 Riverside Drive, New York, New York.

²Correspondence should be directed to Peter S. Jensen, MD, Center for the Advancement of Children’s Mental Health at

Columbia University/NYSPI, 1051 Riverside Drive, New York, New York 10032; e-mail: pj131@columbia.edu.

5. When negative effects are reported, appropriate analyses to address alternative explanations (lack of power, floor or ceiling effects, mediator analyses to address possible fidelity or adherence problems, therapist effects, absence of main effects but possible subgroup/moderator effects, etc.).

KEY WORDS: attention controls; efficacy; psychotherapy; controlled clinical trials.

INTRODUCTION

Over four decades of psychotherapy research studies have been conducted to determine the most efficacious treatments for childhood mental disorders. On one hand, this body of research, well over 1500 studies (Kazdin, 2000) and 500 within the last decade alone (Weisz & Jensen, 1999), has yielded impressive and mounting evidence for the effectiveness of several forms of cognitive behavior therapy for childhood internalizing disorders (anxiety, depression) and behavior therapies for externalizing disorders (conduct and oppositional disorders; ADHD, etc.). On the other hand, terms such as “evidence-based” and “empirically supported” have grown astronomically (see Hoagwood, Burns, Kiser, Ringeisen, & Schoenwald, 2001) to the point that they risk becoming such a familiar part of the mental health clinicians’ and policy makers’ lexicon that we too easily pass over all the difficult steps required in the chain of scientific proof to establish causal links between a proposed therapeutic procedure, its putative underlying mechanisms of change, and client outcomes. Thus, despite the advances in knowledge produced by this body of research, significant gaps remain in identifying the “active ingredients” (either specific or nonspecific therapeutic factors) that underpin actual behavior change and therapeutic effectiveness (Kazdin, 2000). Without convincing knowledge of what works why, we cannot know how best to ensure that proven procedures can effectively be translated from the ivory towers of our research laboratories to the “earthen trenches” of our real-world community clinics (Jensen, Hoagwood, & Trickett, 1999; Weisz, Weiss, & Donenberg, 1992, 1995).

Across this large body of research, a range of specific causal mechanisms for behavior change have been advanced, ranging from changes in arousal, motivation, learning, attributions, self-talk, self-concept, cognitive distortions, interpersonal relations, as well as nonspecific factors (see review by Weersing & Weisz, 2002). Yet in their review of 67 trials of empirically supported treatments, Weersing and Weisz noted that in only six instances were for-

mal tests of mediating mechanisms conducted, despite the fact that the majority of these trials did measure possible mediating mechanisms.

In all this research activity, too rarely do we step back to examine the extent to which all necessary elements of scientific “proof” have been woven together to provide the necessary, sturdy empirical fabric that satisfactorily demonstrates that the overarching theory has scientific support, that the theory-driven procedures have therapeutic effectiveness, and that postulated mechanisms for change have been identified (Jensen et al., 1999; Kazdin & Nock, 2003; Weersing & Weisz, 2002). Pausing for such reflection is necessary, however, since a number of competing explanations might equally be invoked to explain change in many instances. In many studies, for example, “non-specific therapeutic factors,” i.e., the effects of attention, positive regard, and/or therapeutic alliance, may not be sufficiently examined or ruled out as alternative causal explanations (e.g., Wiggs & Stores, 1999).

Taking a step back from current concepts of “efficacy” and “evidence-based,” as described now in numerous papers (Chambless & Hollon, 1998; Hoagwood et al., 2001; Weisz & Jensen, 1999), it is important at this juncture in the further expansion of research-based treatments to ask, “to what extent have recent studies used control groups even roughly comparable in intensity of exposure and face validity to the supposed active treatment?” Such studies are critical, if we are to conclude that elements or features of a given therapy are *specifically* effective, over and above simple compassion, friendliness, attention, and belief.

To address this question, we reviewed all studies in the last 10 years that included at least two potentially active interventions, even if one of them was of somewhat lesser intensity (such as a treatment-as-usual group or an attention-only group). Studies comparing therapy groups with different levels of intensity and “dose” of the putatively active treatment were also included. However, studies with only one treatment group (e.g., studies comparing an active treatment to a wait-list control), were

not included, because such studies cannot help us understand whether a given intervention is efficacious, over and above the amount of personal contact or exposure entailed in the intervention. Of note, some research conducted prior to the last 10 years (in particular studies of specific phobias and disruptive behaviors; see Weersing & Weisz, 2002, for a review) did often include attention controls as well as wait-list controls, but the focus of this review is the extent to which more recently developed interventions over the last decade, often based on manualized procedures and 12–16 psychotherapy sessions delivered in either individual or group format, give heed to these design issues when these treatments are evaluated.

Studies with active comparisons between treatments and attention controls are important to examine because to the extent that differences occur in favor of specific treatments, they may reveal that such effects cannot be explained away by the effects of attention alone. A finding of *no difference* on the other hand may result from the fact that the study is underpowered (i.e., less than 30 subjects per group), or may occur because both the treatment and the attention control arms were similarly efficacious or similarly ineffective. The possibility that an attention control group may actually have therapeutic elements highlights the difficulty of generating a truly inactive psychosocial control condition that could not be considered curative from one or more psychosocial theoretical perspectives (Baskin, Tierney, Minami, & Wampold, 2003).

To fully address the challenges in ferreting out the possible therapeutic effects of specific and nonspecific factors, then, would subject all studies of new psychotherapies to the burden of proof that they a) show differences in outcomes between the new treatment, an attention control treatment without the specific therapeutic ingredients, and a wait-list control group (to control for nonspecific therapeutic effects of attention, therapeutic alliance, etc.), b) measure and conduct formal tests of the differentiating and mediating roles of specific *and* nonspecific factors in the two treatment groups, and c) show differences in outcomes between the attention control and wait-list control groups. This third and final step would be especially key in those cases where mediator analyses comparing the two treatment groups show that a principal component of treatment outcomes is born by nonspecific factors, which may be true for some disorders or treatment types but not

others. Practically speaking for real-world studies, if it can be shown that nonspecific factors account for the lion's share of therapeutic outcomes for a given condition, it might be far less expensive to focus on teaching therapists how to maximize these positive aspects of therapy, rather than attempting to teach and encourage them to use the specific therapeutic procedures addressed in manual-based psychotherapies.

Other study designs may be informative as well. For example, studies that compare two active treatments or two dosages or intensities of the same treatment deserve our attention for two reasons: first, if an appropriately powered study comparing two theoretically similar treatments of differing intensity shows no differences, such findings may raise appropriate concerns as to whether the putative efficacious treatment ingredients are in fact active, since differences in dose yield no differences in outcomes. Of course, this hypothesis assumes a linear dose–response relationship that may or may not be relevant for certain psychosocial treatments. It is equally plausible that a “threshold” dose of a certain magnitude yields effects, and that any greater dose does not result in further improvement. Thus, interpretations of “no difference” should be made cautiously, since linear dose–response relationships may not be uniformly applicable. Nonetheless, interpretations of such studies should be attempted in the absence of strong theory articulating the merits of nonlinear or threshold effects: if an appropriately powered study yields no differences between the wait-list control group and either of the differentially dosed active treatments, one might cautiously assume that neither of the treatments is effective, nor are there any meaningful effects of attention. If both treatments are superior to the wait-list control, yet do not differ from each other, the notion of threshold effects as an explanation for this lack of difference is plausible.

Other possible variations in dose or intensity may have implications for the interpretation of effects. If the component added to the more intensive treatment is actually different in kind (e.g., adding a parent-focused treatment to a child anxiety treatment), the finding of “no difference” may not only be the result of a “ceiling” effect having been reached by the first presumably efficacious component, but it is also possible that the added component is simply inactive (*added component possibly inactive—ADI*) and adds nothing to the outcomes. On the other

hand, a positive difference in this case could also be legitimately interpreted as possibly due to an “*attentional effect only*” (AEO).

So the key question for our review was as follows: When rigorous studies comparing two comparably credible (but one evidence-based, one not) are conducted, what do we find? When a wait-list control group is also included, and differences are also found between it and the treatment groups, are alternative explanations for the findings explored in mediator analyses, such as the dose effects of therapeutic alliance, attention, treatment intensity, or belief in the credibility of the intervention?

METHODS

To address this question, a search of the literature using Psychlit over the last decade (1995–2004) was conducted, using *child(ren)* or *adolescent* and one of the other following terms: *comparison*, *clinical trial*, *controlled trial*, *psychotherapy*, *behavior therapy*, *cognitive behavior therapy*, or *treatment*. This search returned over 4000 abstracts, which were then hand reviewed for content based on the following criteria: a) two or more psychotherapy treatment conditions, including usual care or treatment as usual (TAU); b) random assignment of participants to intervention groups; and c) sample sizes in each group of 10 or more subjects. While sample sizes of 10 will invariably be underpowered in instances of finding “no difference,” we included them nonetheless, given the possibility that some studies might contain treatments of large effects, and to eliminate them might obscure meaningful comparisons in some instances. Nonetheless, “no difference findings” from such studies should be regarded with caution, because treatment group sizes ≥ 24 are needed for power to find an effect size of .4 between two treatment groups (power .81, 5% two-tailed test; Cohen, 1988; Kazdin & Bass, 1989).

For the purposes of this review, random assignment studies comparing a psychotherapy group versus a treatment-as-usual (TAU) comparison group were also accepted if the TAU was some form of psychotherapy, even though TAU psychotherapy is often (though not always) less intensive than the experimental psychotherapy treatment, but *only if*

post hoc analyses were done in the initial or follow-up reports to determine whether the dose of treatment ingredients (number and/or intensity) in the experimental group accounted for part or all of the group differences (*positive dose effect*, +PDE or –PDE).

Pharmacotherapy studies were *not* included, unless there was a comparison with at least two forms of psychotherapy in the several treatment arms. Only one study met this criterion (Abikoff et al., 2004).

Fifty-two studies were located in the inclusive years 1995–2004, then grouped into two categories: 1) those comparing two similarly intensive treatments, such as a putatively active treatment versus an attention control; 2) those comparing more intensive versus less intensive treatment groups (e.g., A vs. A+B, or A vs. A+, or A vs. TAU). Also included in this latter group were three studies comparing an active treatment versus a wait-list control, but *only because* the authors conducted additional analyses to determine if the hypothesized active ingredients of therapy were in fact related to treatment outcome (and their findings in fact revealed that to be the case (Forgatch & DeGarmo, 1999; Hoath & Sanders, 2002; Scott, Spender, Doolan, Jacobs, & Aspland, 2001). Four studies fell into both the first and second groups, because they contained some form of psychotherapy control group as well as two presumably active treatment groups of the same or different intensities.

In a few instances we included in the first group of studies some investigations comparing two or more similar experimental intervention conditions, where both might have been presumed to be active forms of treatment, but where it was unclear whether the contrasting treatment was intended to be an attention control per se.

RESULTS

Of the total of 52 identified studies that met our criteria, 29 fell into the first group where either an attention control was included or two similarly intense treatments were compared. Of these, 14 studies found evidence of consistent differences between the two groups, whereas 15 did not. In fact, in two of these studies, the presumed attention control showed outcomes superior to the presumed active

treatment under study (see Table 1; Merkel, 2000; Murray, Cooper, Wilson, & Romaniuk, 2003).

The second group included 27 studies comparing treatments of differing intensity. Four studies met both criteria (Dishion & Andrews, 1995; Mendlowitz et al., 1999; Murray et al., 2003; Wayland, 2003), thus yielding a total of 52 rather than 56 (29 + 27) studies. Thirteen of this latter group of studies found consistent evidence of additive outcome differences, possibly as a function of the enhanced intervention in the second group; whereas 14 studies revealed no additive differences. We describe each of the groups below, giving examples from each group where both positive and negative effects were found.

Group 1: Studies Comparing Similarly Intensive Treatments and/or Face Valid Attention Controls

Studies Showing Positive Results

Fourteen of twenty-nine studies conducted since 1995 showed a significant difference among similarly intensive treatment groups, suggesting that all treatments are not equal. See Table 1 for a complete listing of all located studies.

As an example of a study showing positive effects over and above attention factors alone, Beidel, Turner, and Morris (2000) randomized children to either a behavioral treatment designed to enhance their social skills and decrease social anxiety (*social effectiveness therapy for children—SET-C*) or to an active but nonspecific control. Results indicated that children treated with SET-C had significantly lower levels of social phobia at treatment endpoints versus those treated in the attention control group. This study is of marked importance in that it demonstrated that children not only benefited as a function of the intervention, but also that they maintained their gains at 6 months' follow-up. Note in Table 1, however, that the authors did not pursue mediational analyses, exploring whether there were positive dose effects (PDEs) between the posited active ingredients of therapy and outcomes. In addition, while they did meet the initial test for efficacy, various competing explanations (CEs) cannot be ruled out (true of most all studies, such as actual perceived differences in treatment credibility that may have affected subject ratings of outcomes).

Not uncommonly, studies may show initial differences among various treatments, followed by

later dissipation of differential effects. For example, Brent et al. (1997) conducted a comparative study of cognitive-behavioral therapy, systematic behavioral family therapy, and nondirective supportive psychotherapy in adolescent suicide attempters. Initial results indicated that the cognitive-behavioral therapy produced more rapid treatment gains than the behavioral family therapy and the nondirective supportive therapy. However, at 2-year follow-up of adolescents assigned to the three therapies, investigators found there were no differences in recurrence rates and levels of depression across the three treatments (Birmaher et al., 2000). Results suggest that maximal improvement within a short time frame may be best achieved by certain forms of therapies over others. Such improvements do not guarantee longer-term differences, though it can be argued that this may be due to the chronic, recurring nature of depression rather than to lack of treatment efficacy.

Sonuga-Barke, Daley, Thompson, Laver-Bradbury, and Weeks (2001) explored the efficacy of two psychosocial treatments for preschool children with attention-deficit/hyperactivity disorder, comparing two 8-week-long programs: 1) parent-delivered behavioral-therapy and education programs and 2) parent-counseling-and-support intervention. Subjects were randomly assigned to the parent-training group, the parent-counseling-and-support group, or to a waiting-list-control group. ADHD symptoms were significantly reduced in the parent-training group, relative to the two control conditions. While there were substantial differences between the parent-training group and the other two groups at week 8, at follow-up assessment at week 23, the two active intervention groups showed similar outcomes. At this point, most of the benefits of parent training relative to the parent-counseling-and-support group had dissipated, such that most of the benefit observed in the parent-training children was now observed in the parent-counseling-and-support children. Nonetheless, children in the parent-training group showed persistent gains and no further deterioration after the end of treatment. In contrast, children in the waiting-list-control groups showed minimal or no improvement over the course of the study.

Studies Showing Few or No Differences

Silverman et al. (1999) conducted one of the few studies that contrasted treatments in a head-to-head

Table 1. Group 1: Studies Comparing Experimental Therapy With Attention Controls or Other Treatments of Comparable Intensity

Study authors	Disorder	Design and overall findings	Total N	Adeq. power?	Do similarly intensive treatments differ? (Group 1 studies)	Alternative explanations for positive or negative findings? AEO, ADI, CE, UP	Are overall findings consistent with EBT, over and above effects of attention?
Barrett, Healy-Farrell, and March (2004)	Childhood OCD	Parallel groups, 3 arms: wait-list control vs. group cognitive behavioral therapy (CBFT) vs. individual CBFT. No differences found between two treatment groups, but both were better than wait-list controls; effects maintained at 6 months	77	Yes	No	PDEs not explored	No
Brent et al. (1997)	Major depressive disorder	Parallel groups, 3 arms: cognitive-behavioral therapy (CBT) vs. nondirective supportive therapy (NDST) vs. systemic family behavior therapy (SFBT) Follow-up showed later dissipation of effects	107	Yes	Yes	Unmeasured variables, PDEs not explored	Initially yes, CBT > SFBT > NDST; later no differences, CE
Also: Birmaher et al. (2000)							
Beidel et al. (2000)	Social phobia	Parallel groups, 2 arms: behavioral therapy to reduce social anxiety and enhance social skills, vs. attentional control group. Positive effects maintained at 6 months	67	Yes	Yes	Unmeasured variables, PDEs not explored	Yes, CE
Chadwick, Momiclovic, Rossiter, Stumbles, and Taylor (2001)	Behavior problems and severe learning disabilities	Case-control study, 3 London boroughs: group-based parent behavior training vs. individually based parent behavior training vs. wait-list controls. Parents in two intervention groups did better than controls. While parents reported better outcomes when receiving individual treatment, objective measures did not confirm this. Differences across all groups dissipated at 6 months	106	Yes	No, but some variables did in post hoc analyses	PDEs examined, effects seen for âtargetedâ behavioral outcomes	No. The few differences found dissipate at follow-up.
Cohen and Mannarino (1996)	S/P Sexual abuse	Parallel groups, 2 arms: CBT vs. nondirective supportive therapy (NST). CBT more efficacious than NST	77	Yes	Yes	Unmeasured variables, PDEs not explored	Yes, CBT > NST, CE
Cohen et al. (1999)	Stressed parents & their at risk infants	Parallel groups, 2 arms: Infant-contingent psychotherapy (Watch, Wait, and Wonder—WWW) vs. mother–infant psychotherapy (MIP)	67	Yes	Yes, on some measures	Unmeasured variables, PDEs not explored	Yes, WWW > MIP, CE

Table 1. Continued

Study authors	Disorder	Design and overall findings	Total <i>N</i>	Adeq. power?	Do similarly intensive treatments differ? (Group 1 studies)	Alternative explanations for positive or negative findings? AEO, ADI, CE, UP	Are overall findings consistent with EBT, over and above effects of attention?
Liddle et al. (2001)	Adolescent drug abusers	Parallel groups, 3 arms: adolescent group therapy, multidimensional family therapy (MDFT), and multi-family education and support. Improvement found in all 3 groups, with MDFT showing best results overall and at 1 year follow-up	182	Yes	Yes	Unmeasured variables, PDEs not explored	Yes, CE
McArdle et al. (2002)	Risk for behavior problems	Parallel groups, 2 arms: children assigned to a "drama therapy group" vs. children assigned to curriculum studies group. Drama group youth outcomes superior on multiple measures	122	Yes	Yes	Face validity, PDEs not explored	Yes, CE
Mendlowitz (1999) ^{a,b}	Anxiety disorders	Parallel groups, 3 arms: child-only CBT, parent-only CBT, or both. On a number of outcome measures, parents and children in the combined CBT intervention group fared better than the other two groups	62	No	No	AEO, UP	No
Merkel (2000)	Child victims of violent crime	Parallel groups, 2 arms: child CBT vs. a similarly intensive supportive intervention. In general, no differences found between groups, except that contrary to the study hypotheses, parents of children in the supportive intervention reported greater reductions in children's symptoms than parents of children in the CBT intervention	27	No	No. Occurred in the opposite direction of hypotheses	UP, CE	No
Muris, Meesters, and van Melick (2002)		Parallel groups, 3 arms: group CBT vs. attention control (emotional disclosure) vs. no-treatment controls (<i>n</i> = 10 in each group). Pretreatment-posttreatment comparisons showed that CBT was superior to ED and no treatment control in reducing anxiety and depression symptoms	30	No	Yes	Unmeasured variables, PDEs not explored	Yes, CE

Murray et al. (2003) ^{1,2}	Postpartum depressed mothers	193	Yes	No. Opposite direction for some variables	CE	No
Nicholson and Sanders (1999)	Child behavior problems in stepfamilies	42	No	No	UP, CE	No
Robin et al. (1999)	Anorexia nervosa	37	No	Yes	Unmeasured variables, PDEs not explored	Yes, CE
Rossello and Bernal (1999)	Major depression in Puerto Rican adolescents	71	No	Yes	Unmeasured variables, PDEs not explored	Yes, CE
Silverman et al. (1999)	Childhood phobic disorders	81	Yes	No	AEO	Mostly No
Snyder, Kymissis, and Kessler (1999)	Anger problems in youth	50	Yes	Yes	AEO. No PDE analyses	Yes, CE
Sonuga-Barke et al. (2001)	ADHD in preschool children	78	Yes	Yes	Unmeasured variables, PDEs not explored	Yes, CE

Table 1. Continued

Study authors	Disorder	Design and overall findings	Total <i>N</i>	Adeq. power?	Do similarly intensive treatments differ? (Group 1 studies)	Alternative explanations for positive or negative findings? AEO, ADI, CE, UP	Are overall findings consistent with EBT, over and above effects of attention?
Spoth, Redmond, and Shin (2001)	Substance abuse risk	Parallel groups, 3 arms: control condition vs. 2 family skill-strengthening programs: a 5-session Preparing for the Drug Free Years Program vs. a 7 session Iowa Strengthening Families Program. Both active programs achieved better outcomes than minimal contact controls	667	Yes	No	UP, no PDE analyses	No
Storr, Ialongo, Kellam, and Anthony (2002)	Risk for smoking onset	Parallel groups, 3 arms: children in control classrooms, vs. a classroom-centered (CC) intervention vs. a family school-partnership intervention (FSP). Both interventions more effective than control classrooms, but did not differ among themselves	678	Yes	No	Unmeasured variables, no PDE analyses	No
Verduyn, Barrowclough, Roberts, and Tarrrier, Harrington (2003)	Maternal depression and child behavior problems	Parallel groups, 3 arms: parental CBT vs. parent support group vs. wait-list control. No significant differences found across groups, though trends favored the CBT plus parent skills group at follow-up	119	Yes	No	UP, AEO, no PDE analyses	No
Wayland (2003) ^{a,b}	Childhood problem behaviors	Parallel groups, 3 arms: individualized behavioral consultation with the parent, videotape review with group discussion/consultation, vs. videotape alone. Results generally showed no differences among groups	35	No	No	ADI, UP	No
Wood, Harrington, and Moore (1996)	Depressive Disorders in Youth	Parallel groups, 2 arms: CBT vs. relaxation training. Significant improvements seen in CBT group over relaxation training group	48	Yes	Yes	Face validity, PDEs not explored	Yes, CE

Note. ADI: added component possibly inactive; NA: not applicable; AEO: findings can be attributed to an “attention effect only”; PDE: *positive dose effect* (present [+]) or absent [-]; CE: other *competing explanations* cannot be ruled out. UP: likely *under-powered* to truly ascertain “no difference.”
^{a,b}Study meets criteria for Groups 1 and 2. Coded in both Tables 1 and 2.

trial, comparing a behavioral approach utilizing contingency management, a self-control strategy, and educational support for children with phobic disorders. All groups were found to be similarly efficacious on most treatment-outcome measures, including the presumably inactive control (the educational supportive intervention). Findings underscored the point that a basic therapeutic relationship, family support, and education may yield substantial benefits for child and family outcomes. In several instances, however, there was evidence of some modest advantages of the self-control and/or contingency management strategy over the educational support condition.

Similarly, Last, Hansen, and Franco (1998) tested a manualized cognitive-behavioral treatment for school phobia in 56 children, one-half of whom were randomly assigned to the intervention over 12 weeks, while the other half were provided a similarly intensive educational-supportive intervention. Comparing the groups on school attendance, self-reports of anxiety and depression, and diagnostic status at endpoint, they found that both interventions were similarly efficacious in returning children to school and in reducing their anxiety and depressive symptoms. Likewise, follow-up assessments of children were during the next school year were done, with analyses again showing no differences between the two groups.

In another interesting, yet counterintuitive finding, Murray et al. (2003) conducted a longitudinal study of 193 postpartum depressed women and their newborn infants over 5 years. Mothers were randomly assigned to one of four groups: CBT, psychodynamic therapy, nondirective supportive therapy, or routine primary care. Few differences were found among the groups at 5 years, but some evidence emerged for the superiority of the nondirective counseling arm at 18 months.

Summary of Group 1 Studies

Thus, in this review we found that of those studies that compared two active treatments or otherwise attempted to provide some reasonable control for attention, only about one-half of studies indicated initial differential superiority of one form of psychotherapy over reasonably valid controls (with effects eventually dissipating in many of these studies). The other half of these comparative studies showed no advantage of the psychotherapy over some reasonable, similarly intensive control group. From the

total 29 studies, among only the 17 studies with sufficient power to determine differences with an ES of .4, 10 showed a significant effect of the presumed active treatment with effects greater than the impact of attention alone, while 7 did not. Without exception, among these 10 sufficiently powered studies with positive results, mediator analyses (i.e., examining possible dose effects of the presumed active ingredients) were *not* performed, paralleling the findings reported by Weersing and Weisz (2002) that such studies are rarely done. Likewise, in all these 10 instances, competing explanations for positive results (such as parent/family perceptions of the believability/credibility of the treatment, therapeutic alliance, etc.) could not be ruled out. These findings indicate that while our “evidence-based” treatments are often promising, the ingredients that make them so remain largely unknown.

One must therefore ask, “what do the presumed active elements add above and beyond the nonspecific elements of psychotherapy, and why are our evidence-based therapies often not more effective than attention controls? Certainly, one explanation is that many studies are underpowered to detect a clinically meaningful effect, but an alternative, complementary explanation is that the effects of our evidence-based therapies may often be only modestly better than the nonspecific effects of attention, positive regard, and support.

An examination of studies that directly compared treatments of differing intensities may provide some additional insights on this question, however. If similar therapies of differing intensity yield different results, but different therapies of similar intensities do not, one parsimonious conclusion would suggest that some common aspects of the equally intense treatments (including therapeutic alliance, rapport, and attention, as well as other factors such as explicit problem definition or psychoeducation) are powerful factors in and of themselves, perhaps exerting even greater effects than the “evidence-based” component—whatever that may be—of the therapy itself. Examples of such studies are illustrated below.

Group 2: Studies Comparing Treatments of Differing Intensity

Studies Showing Positive Effects of Intensity

Thirteen of twenty-seven studies conducted since 1995 showed a significant difference among similarly intensive treatment groups, suggesting that

differing intensity, i.e., treatment dose, makes a difference. See Table 2 for a complete listing of all located studies. Please note that we do not include studies that included a presumed active treatment against wait-list controls only, nor do we include studies that compared an experimental treatment versus TAU, *unless* that study *also* specifically examined whether treatment dose/intensity/compliance in either or both groups mediated outcomes within or across groups.

As an exemplar study showing additive effects of different doses (or types) of treatment, Kazdin and Whitley (2003) examined the impact of a parent problem-solving (PPS) intervention designed to augment the effects of other evidence-based therapies for children referred for aggressive and antisocial behavior. Thus, all children received problem-solving skills training (PSST), while their parents received parent management training (PMT). Families were randomly assigned to receive or not to receive the additional component (PPS) that targeted parental stress factors. Although all children improved in both treatment arms, those families receiving the additional PPS intervention showed greater behavioral change in youth, and parents/families experienced fewer barriers to care.

Similarly, Jocelyn, Casiro, Beattie, Bow, and Kneisz (1998) studied treatments of children with autism, conducting a randomized controlled trial to evaluate a caregiver-based intervention program versus TAU in community day-care centers. Families were assigned to one of two conditions: 1) day-care for the child, with a one-on-one child development worker, or 2) the same day care experience as 1) plus an intensive 12-week intervention program for the parents, teaching them skills in functional analysis and behavior therapy, and additional educational seminars on autism. Children whose parents received the intensive training showed substantially greater language gains than day-care-only children. Analyses exploring why this group showed greater gains were not conducted, so specific conclusions about whether these differences were due to attention or to parents' application of behavioral procedures over and above what was done in the therapeutic day-care setting cannot be determined.

Studies Showing Few or No Effects of Intensity

By and large, studies of negative effects are difficult to find, particularly since so many studies

have contented themselves with wait-list-only controls (i.e., treatment vs. nothing), or experimental treatment versus TAU, but without determining whether specific ingredients or dose (e.g., compliance) of the experimental treatment mediated positive outcomes.

In the child psychotherapy literature, there seems to be one exception to the general rule that "*no negative studies can be located*"; namely studies of the relative impact of adding one presumably efficacious component to another and then seeing if the combined therapy package is superior to a single efficacious component. For example, Spence, Donovan, and Brechman-Toussaint (2000) randomized 50 children to one of three arms: 1) a social-skills-based 12-session training program delivered principally in group settings, followed by booster sessions at 3 and 6 months after the completion of the course; 2) the same treatment augmented by parent involvement, where parents were taught to reinforce children's practice of target skills and to not reinforce socially anxious and avoidance behaviors; and 3) a wait-list control group. Study outcome data indicated that children in both active-treatment conditions, compared to the wait-list control, showed significant improvements in social phobia symptom severity ratings on multiple measures, as well as on the proportions of children who met full diagnostic criteria for social phobia at the end of treatment. However, there were no significant differences between the social-skills intervention programs with and without parental involvement.

Similar studies targeting anxiety or depression treatments have been conducted by Cobham, Dadds, and Spence (1998); Silverman et al. (1999); and Clarke, Rohde, Lewinsohn, Hops, and Seeley (1999); with all studies essentially not finding a significant main effect of adding parent-directed therapy to the already proven-efficacious child intervention. By way of caution, it should be noted that the Spence et al. (2000) study was underpowered to detect effect sizes of small to moderate effects, and that post hoc analyses in the Cobham et al. (1998) study found that CBT + parental anxiety management was more efficacious among a subgroup of parents who were anxious.

Studies Examining Mediators of Change or Dose Effects of Presumed Therapeutic Ingredients

Very few studies were found that explicitly examined not just *whether*, but how and why therapy

Table 2. Group 2: Studies Comparing Treatments of Differing Intensity

Study authors	Disorder	Design and overall findings	Total N	Adeq. power?	Do dose-dissimilar Tx differ? If yes, + or -PDE? (Group 2 studies)	Alternative explanations for positive or negative findings? AEO, ADI, CE, UP	Are overall findings consistent with EBT, over and above effects of attention?
Barrett, Dadds, and Rapee (1996)	Anxiety disorders	Parallel groups, 3 arms: CBT alone, CBT plus family management, and wait-list controls. CBT + family management had better outcomes on some variables	79	Yes	No	Face validity, AEO, ADI	No
Barrett et al. (1996)	Childhood anxiety disorders	Parallel groups, 3 arms: group CBT, group CBT plus a family management skills training, and wait-list controls. Both active groups better than WL, some evidence of modest superiority of the CBT plus family intervention	60	No	Yes. PDE not reported	AEO, UP	No
Chamberlain and Reid (1998)	Chronic juvenile offenders	Parallel groups; 2 arms: group foster care vs. multidimensional treatment foster care (MTFC); MTFC resulted in less criminality, more likely to return to live with relatives	79	Yes	Yes. +PDE reported in follow-up by Eddy and Chamberlain, 2000	AEO, CE	No
Also: Eddy, Whaley, and Chamberlain (2004) and Eddy and Chamberlain (2000)		MTFC also resulted in less violent offending over 2-year follow-up Mediator analyses suggested improvement related to degree that MTFC addressed known risk factors					
Clarke et al. (1999)	Youth with major depression	Parallel groups, 3 arms: wait-list controls, adolescent group CBT, vs. adolescent group CBT with separate parent sessions. No differences found between active treatment arms	123	Yes	No. PDE not examined	CE, UP, ADI	No
Clarke, Hornbrook, Lynch, and Polen (2002)	Youth with depressed parents	Parallel groups, 2 arms: usual care within the HMO vs. 16-session CBT. No differences found for youth receiving HMO care and those receiving HMO care plus CBT	88	Yes	No	ADI, CE	No

Table 2. Continued

Study authors	Disorder	Design and overall findings	Total N	Adeq. power?	Do dose-dissimilar Tx differ? If yes, + or -PDE? (Group 2 studies)	Alternative explanations for positive or negative findings? AEO, ADI, CE, UP	Are overall findings consistent with EBT, over and above effects of attention?
Cobham et al. (1998)	Anxiety disorders	Parallel groups, 2 arms: CBT alone or CBT plus a parent anxiety management component (PAM). No difference between groups found. However findings suggest CBT + FAM more effective for anxious children w/anxious parents	67	Yes	No, except by post hoc analyses	ADI, CE	No
Connolly, Sharry, and Fitzpatrick (2001)	Childhood behavior problems	Parallel groups, 3 arms: Wait-list control vs. Webster-Stratton videotape training program (VTP), vs. VTP plus additional components tailored to child's needs. Both active treatments better than wait-list, no differences found between two active treatment groups	105	Yes	No	ADI, CE	No
Dadds, Spence, Holland, Barrett, and Laurens (1997)	Childhood anxiety problems	Prevention study, 2 arms: a monitoring group vs. school-based parent- and child- CBT intervention. Both groups showed similar improvements after 10 weeks, but only the presumed active group maintained benefits at 6 months	128	Yes	Not initially, effects seen later	Unmeasured variables, PDEs not explored	Initially no, later yes, CE
Dishion and Andrews (1995) ^{a,b}	High-risk adolescents	Parallel groups, 4 arms: a) parent focused, b) teen focused, c) parent and teen focused, d) self-directed change (materials only). Parent-focused intervention > all other groups on school behavior problems, tobacco use	119	Yes	Yes	ADI, CE	Yes
Drew et al. (2002)	Autism	Parallel groups, 2 arms: treatment as usual vs. TAU plus parent training. Parent training group had better outcomes in general	24	No	Yes, PDE not reported	Face validity, AEO, unmeasured variables	No
Driskill (2000)	ADHD	Parallel groups, 3 arms: wait-list control, parent training group, or parent informational handouts only. Children received 7 sessions of social skills and behavioral management in all 3 groups. Parent training group participants showed greater changes on most measures than handout or wait-list participants	60	No	Yes, PDE not examined	AEO, CE	No

Forgatch and DeGarmo (1999)	Children of divorcing mothers	Parallel groups, 2 arms: Wait-list controls vs. 14 session group behavioral training, videotape modeling, and problem solving for parents. Positive differences between groups noted, with changes in parents' disciplinary and parenting practices mediating teacher and parent ratings of improved outcomes	238	Yes	Yes, +PDE examined, found to mediate	AEO	No
Hebert (2002)	Disruptive adolescents	Parallel groups, 2 arms: 8 week manualized CBT vs. TAU inside an intensive special school	21	No	No	CE, UP, ADI	No
Henggeler, Melton, Brondino, Scherer, and Hanley (1997)	Chronic juvenile offenders	Parallel groups, 2 arms: Multisystemic Therapy (MST) vs. usual juvenile justice services. MST resulted in decreased adolescent symptoms and incarceration at follow-up. Outcomes better with higher adherence	155	Yes	Yes, +PDEs reported (fidelity)	ADI, CE	No
Hoath and Sanders (2002)	ADHD	Parallel groups, 2 arms: wait-list control vs. an enhanced Triple P behavioral intervention. Parents of children receiving the active intervention also had changes in disciplinary style	20	No	Yes, +PDE	AEO	No
Jocelyn et al. (1998)	Autism	Parallel groups, 2 arms: traditional daycare with one-on-one worker, vs. the same plus a 12 week intensive program for parents teaching them functional analysis, beh therapy. Parents and children in the experiment group showed significantly greater gains than the traditional daycare group	35	No	Yes, +PDE	AEO	No
Kazdin and Whitley (2003)	Aggressive antisocial youth	Parallel groups, 2 arms: a) problem-solving skills training (PSST) and parent management training (PMT), or b) PSST, PMT, and a parental stress intervention. Added PPS component yielded greater change for children and parents and reduced family treatment barriers	127	Yes	Yes, but PDEs not examined	AEO, CE	No

Table 2. Continued

Study authors	Disorder	Design and overall findings	Total N	Adeq. power?	Do dose-dissimilar Tx differ? If yes, + or -PDE? (Group 2 studies)	Alternative explanations for positive or negative findings? AEO, ADI, CE, UP	Are overall findings consistent with EBT, over and above effects of attention?
King et al. (2000)	Sexually abused children	Parallel groups, 3 arms: wait-list control vs. individually directed CBT vs. individual CBT plus parent-directed CBT. Both treatment groups fared better than wait-list control, though parents found the combined program more satisfactory	36	No	For almost all measures no differences found	UP, CE, ADI	No
Mendlowitz (1999) ^{a,b}	Anxiety disorders	Parallel groups, 3 arms: child-only CBT, parent-only CBT, or both. On a number of outcome measures, parents and children in the combined CBT intervention group fared better than the other two groups	62	No	Yes. But PDE not explored	AEO, UP	No
Montgomery, Stores, and Wiggs (2004)	Mentally retarded children w/ sleep problems	Parallel groups, 3 arms: wait-list control vs. brief behavioral therapy vs. behavioral information in booklet form. Two active groups showed better outcomes than wait-list control, no differences found between active groups	66	No	No, PDE not reported	UP, CE	No
Murray et al. (2003) ^{a,b}	Postpartum depressed mothers	Parallel groups, 4 arms: CBT vs. dynamic therapy vs. nondirective counseling, vs. regular primary care follow-up. In general, no differences were seen on children's behaviors through 5 years of age, but some evidence of superiority of nondirective counseling was noted for 18-month outcomes	193	Yes	No	CE	No
Ost, Svensson, Hellstrom, and Lindwall (2001)	Specific phobias	Parallel groups, 3 arms: exposure condition vs. exposure plus parental involvement, vs. wait-list control. Both active conditions superior to WL. No differences between groups	60	No	No	ADI, CE, UP	No

Author(s)	Outcome	Intervention	N	Yes	Yes, +PDE reported	AEO	No
Scott et al. (2001)	Childhood antisocial behavior	Parallel groups, 2 arms: wait-list controls vs. Webster-Stratton videotape parent training program. Intervention group fared better than wait-list controls on multiple outcomes, including presumed mediators of outcomes	141	Yes	Yes, +PDE reported	AEO	No
Spence et al. (2000)	Childhood social phobia	Parallel groups, 3 arms: wait-list control vs. child-focused social skills training vs. child-focused CBT augmented by parent training. Both active interventions were superior to wait-list, but did not differ among themselves	50	No	No	UP, ADI, CE	No
Taylor, Schmidt, Pepler, and Hodgins (1998)	Preschool behavior problems	Parallel groups, 3 arms: wait-list control, vs. clinic treatment as usual (TAU) vs. the Parent and Child Webster-Stratton Series (PACS). PACS and TAU were superior to wait-list, and PACS was superior to TAU on several measures	108	Yes	Yes	Unmeasured variables, PDEs not explored	No
Wayland (2003) ^{a,b}	Childhood problem behaviors	Parallel groups, 3 arms: individualized behavioral consultation with the parent, videotape review with group discussion/consultation, vs. videotape alone. Results generally showed no differences among groups	35	No	No	ADI, UP	No
Wolchik et al. (2002)	Children postdivorce	Parallel groups, 3 arms: a control condition (reading materials on divorce only), an 11 session problem-solving group for mothers, and a group that augmented the mother-only program with 11 sessions of youth directed content. Both groups were found to be superior to the wait list, with no meaningful differences among groups	218	Yes	No	ADI	No

Note. ADI: added component possibly inactive; NA: not applicable; AEO: findings can be attributed to an "attention effect only"; PDE: positive dose effect (present [+]) or absent [-]; CE: other competing explanations cannot be ruled out. UP: likely under-powered to truly ascertain "no difference."
^{a,b}Study meets criteria for Groups 1 and 2. Coded in both Tables 1 and 2.

worked. In one of the few exceptions from Group 1 studies, Kaminer, Blitz, Burlison, Kadden, and Rounsaville (1998; Kaminer, Burlison, Blitz, Sussman, & Rounsaville, 1998) compared two 12-week manualized treatments for adolescents with substance abuse, hypothesizing that adolescents with comorbid externalizing disorders would be more likely to respond to CBT, compared to adolescents with internalizing disorders, presumed more likely to respond to an interactional group treatment. Kaminer and colleagues demonstrated that both treatments could be carefully specified in terms of the actual delivery of specific therapeutic procedures and that therapists were able to deliver the therapy-specific interventions to research participants with good fidelity to the treatment principles, but contrary to their initial hypotheses, they found no relationship between patient problem type (internalizing vs. externalizing problems), hypothesized therapeutic processes, and specific outcomes.

From Group 2 studies, Forgatch and DeGarmo (1999) developed a theory-based 14-session parent-training program for 238 divorcing mothers whose sons were in grades one to three. The intervention relied upon videotape vignettes as well as specific training in noncoercive discipline, contingent reinforcement, monitoring, and parental problem solving. Parents were randomly assigned to the intervention versus a wait-list control. Over the course of the intervention, the investigators found that the postulated mechanism of adverse outcomes, coercive interactions, were significantly reduced between mothers and children, and that these improved parenting practices correlated significantly with improvements in teachers' and mothers' reports of children's adjustment. These results support the overall model developed by Patterson and colleagues concerning the role of coercion in the development of behavioral and conduct problems. While this study would have been improved by an equally intensive treatment control group that included none of the ingredients postulated to change coercive interactions, the study is nonetheless well-executed and valuable in shedding light on what may have brought about change.

Similarly from Group 2, Eddy and Chamberlain (2000), following on the earlier study by Chamberlain and Reid (1998) examined the impact of factors hypothesized to be related to improved treatment outcomes in multidimensional-treatment foster-care settings, including family management skills, supervision, discipline, positive reinforcement, and the extent of deviant peer associations.

Results bore out investigators' hypotheses: Youth participating in the multidimensional-treatment foster-care placement had significantly improved youth outcomes compared to the usual group settings. Moreover, mediational analyses indicated that the impact of treatment was carried through the processes of increased parental supervision, positive adult-youth relationships, and reduced contact with deviant peers. The combination of family management skills and association with deviant peers accounted for 32% of the variance in later antisocial behavior. This study is particularly important in that it explicitly examined the active therapeutic ingredients most relevant to producing change in youth delinquency and the prevention of future behavioral problems.

IMPLICATIONS

Our review of studies over the last decade reveals significant limitations in the knowledge base about evidence-based psychosocial treatments. First, only 5 of the 52 studies reviewed deliberately conducted a systematic examination of the mediators of change, paralleling the findings of Weersing and Weisz (2002) that relatively few investigators undertake these analyses. Without further attention to the mediators and processes of change, accumulation of evidence from clinical trials will only add to the generation of more questions, not policy-relevant answers. What are the therapeutic ingredients that activate improvement? Lack of knowledge in this area is seriously impeding the efficient deployment of higher quality treatments into the public mental health system.

Secondly, roughly half of the studies contrasting active attention to active therapy indicated no preferential effect of therapy. Consequently, in many instances the positive and perhaps "therapeutic" effects of attention, belief, expectancies, and values have not been systematically eliminated as potential ingredients of specific therapies. This review suggests that we must sharpen our theories of psychotherapy, beginning first by better specifying the active ingredients of therapeutic change. Research attention should be paid not simply to specific psychological procedures, but also to the "nonspecific" factors such as engagement, empathy, therapeutic alliance, belief, and hope. More specific, experimental tests of how therapists can apply these interpersonal qualities more effectively and consistently are needed, regardless of the other specific techniques that they use.

In terms of possibly achieving improvements in psychotherapy outcomes in the “real world,” such efforts might be more effective and cost-effective than too great a focus on evidence-based, manualized therapies that may have only modest benefits over and above the effects of attention, and that quite possibly may even be *less effective* than well-timed and delivered empathy and support, if these factors are not carefully specified and controlled as a crucial component of our evidence-based therapies. Moreover, our theories of therapy would be strengthened through better linkage to known basic psychological mechanisms that have been examined in studies of behavioral change, including decision-making, social cognition, motivation, self-efficacy, expectancies, etc. To what extent might these factors be operative in psychotherapy? For example, Jocelyn et al. (1998) found that their intervention when added to TAU increased parents’ sense of self-efficacy, yet this type of measure seems relatively rare in the current psychopathology treatment literature, as is exploring such variables’ possible mediating effects on behavioral outcomes (see Bickman, Heflinger, Northrup, Sonnichsen, & Schilling, 1998; Heflinger, Bickman, Northrup, & Sonnichsen, 1997 for another exception).

Third, it should be noted that even in trials utilizing an attention control group and showing that treatment benefits are superior in the group receiving the posited active therapeutic ingredients, competing hypotheses explaining the findings often remain viable and are rarely examined, such as patients’ perceptions of treatment credibility, and so forth. These will remain difficult challenges to address, though they can be, and have been, partly handled by the use of blinded, independent outcome raters and reliable measures of these alternative explanations.

In our next generation of psychotherapy clinical trials, we suggest that the time has come to raise the bar on the “standard of evidence” required for our evidence-based treatments. Building upon the seminal recommendations of Chambless and Hollon (1998), these standards, and studies based on them should strive to include:

- 1) Initial specification in study design how investigators will test if the intervention is efficacious over and above the effects of positive expectancies, positive regard, or attention,
- 2) Planned (rather than post hoc) analyses to explore whether and how a given treatment’s

- specific ingredients or the overall intensity of its “dose” are related to treatment outcomes,
- 3) More studies conducting head-to-head tests of different types but equally credible forms of treatment, with planned analyses testing different mechanisms of change.
- 4) When positive effects of a treatment versus a control are found, systematic elimination of specific, competing hypotheses of reasons for treatment efficacy (e.g., attention, therapeutic alliance, face validity of treatment and client’s treatment expectations, changes in self-efficacy), *and*
- 5) When negative effects are reported, appropriate analyses to address alternative explanations (lack of power, floor or ceiling effects, mediator analyses to address possible fidelity or adherence problems, therapist effects, absence of main effects but possible subgroup/moderator effects, etc.).

Given monetary and practical constraints, not every trial will be able to address all of these areas, but we hope this list can serve as a challenge to substantially improve the specificity and explanatory power of our treatment research. To address these considerations, psychotherapy studies will often need to extend beyond designs that employ only a single treatment arm (generally compared to waitlist controls), and instead consider *differential* treatment contrasts for study design. Even within single treatment groups, differential comparisons can often be pursued, such as moderator analyses testing the effects a priori hypotheses about initial treatment acceptability and match to patient’s wishes, demographic or psychopathology variables (e.g., presence or absence of baseline comorbidity, presence or absence of accompanying parental psychopathology, etc.), or mediator analyses exploring dose, duration, adherence, therapeutic alliance, changes in perceived treatment credibility, etc.

In sum, the current state of the art in psychotherapy studies suggests that what is deemed “evidence-based” or “empirically supported” may be clearing a very low threshold. In many instances, alternative/competing explanations are also consistent with much of the current literature’s findings. We also concur with the recent review by Westen, Novotny, and Thompson-Brenner (2004) about the need for evaluating our standards of evidence and how we apply them to determine whether an intervention meets these seemingly “all-or-none” criteria. As they note,

much needed are more sophisticated, nuanced approaches to how we view and review our accumulating database.

Greater caution in the use of the terms “evidence-based” and “empirically supported” are clearly warranted. If by these terms one means that a given treatment has been tested and found more effective than nothing or a waiting list, then that statement is technically accurate but may be misleading. In fact, by and large our current state of knowledge does not allow unequivocal statements that most interventions considered to be “evidence-based” are better than similarly intense, equally face-valid interventions. And in almost all cases it is not clear that the essential ingredients of change are understood, as is, for example, possible in medicine through the testing of serum levels and dose–response analyses of certain medications. If true progress is to be made in developing and deploying evidence-based treatments, greater circumspection, caution, and qualification are needed in our clinical trials and the conclusions we draw from them, and more attention to the mechanisms of change and to the active ingredients of therapy must occur within these studies. Only in this way will our field more systematically acquire the evidence for “evidence-based” treatments.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support and intellectual contributions of the Youth Mental Health Network. Network members include John Weisz, PhD (Principal Investigator); Bruce Chorpita, PhD; Robert Gibbons, PhD; Charles Glisson, PhD; Evelyn Greene, Kimberly Hoagwood, PhD; Peter S. Jensen, MD; Kelly Kelleher, MD; John Landsverk, PhD; Steven Mayberg, PhD; Jeanne Miranda, PhD; Lawrence Palinkas, PhD; and Sonja Schoenwald, PhD.

REFERENCES

- Abikoff, H., Hechtman, L., Klein, R. G., Weiss, G., Fleiss, K., Et-covitch, J., et al. (2004). Symptomatic improvement in children with ADHD treated with long-term methylphenidate and multimodal psychosocial treatment. *Journal of the American Academy of Child and Adolescent Psychiatry*, 43, 802–811.
- Barrett, P. M. (1998). Evaluation of cognitive-behavioral group treatments for childhood anxiety disorders. *Journal of Clinical Child Psychology*, 27, 459–468.
- Barrett, P., Healy-Farrell, L., & March, J. S. (2004). Cognitive-behavioral family treatment of childhood obsessive-compulsive disorder: A controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, 43, 46–62.
- Barrett, P. M., Dadds, M. R., & Rapee, R. M. (1996). Family treatment of childhood anxiety: A controlled trial. *Journal of Consulting and Clinical Psychology*, 64, 333–342.
- Baskin, T. W., Tierney, S. C., Minami, T., & Wampold, B. E. (2003). Establishing specificity in psychotherapy: A meta-analysis of structural equivalence of placebo controls. *Journal of Consulting and Clinical Psychology*, 71, 973–979.
- Beidel, D. C., Turner, S. M., & Morris, T. L. (2000). Behavioral treatment of childhood social phobia. *Journal of Consulting and Clinical Psychology*, 68, 1072–1080.
- Bickman, L., Heflinger, C. A., Northrup, D., Sonnichsen, S., & Schilling, S. (1998). Long-term outcomes of family caregiver empowerment. *Journal of Child and Family Studies*, 7, 269–282.
- Birmaher, B., Brent, D. A., Kolko, D., Baugher, M., Bridge, J., Holder, D., et al. (2000). Clinical outcome after short-term psychotherapy for adolescents with major depressive disorder. *Archives of General Psychiatry*, 57, 29–36.
- Brent, D. A., Holder, D., Kolko, D., Birmaher, B., Baugher, M., Roth, C., et al. (1997). A clinical psychotherapy trial for adolescent depression comparing cognitive, family, and supportive therapy. *Archives of General Psychiatry*, 54, 877–885.
- Chadwick, O., Momcilovic, N., Rossiter, R., Stumbles, E., & Taylor, E. (2001). A randomized trial of brief individual versus group parent training for behaviour problems in children with severe learning disabilities. *Behavioural and Cognitive Psychotherapy*, 29, 151–167.
- Chamberlain, P., & Reid, J. B. (1998). Comparison of two community alternatives to incarceration for chronic juvenile offenders. *Journal of Consulting and Clinical Psychology*, 66, 624–633.
- Chambless, D., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 66, 7–18.
- Clarke, G. N., Hornbrook, M., Lynch, F., & Polen, M. (2002). Group cognitive-behavioral treatment for depressed adolescent offspring of depressed parents in a health maintenance organization. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41, 305–313.
- Clarke, G. N., Rohde, P., Lewinsohn, P. M., Hops, H., & Seeley, J. R. (1999). Cognitive-behavioral treatment of adolescent depression: Efficacy of acute group treatment and booster sessions. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38, 272–279.
- Cobham, V. E., Dadds, M. R., & Spence, S. H. (1998). The role of parental anxiety in the treatment of childhood anxiety. *Journal of Consulting and Clinical Psychology*, 66, 893–905.
- Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd edn.). Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, J. A., & Mannarino, A. P. (1996). A treatment outcome study for sexually abused preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 35, 42–50.
- Cohen, N. J., Muir, E., Parker, C. J., Brown, M., Lojkasek, M., Muir, R., et al. (1999). Watch, wait and wonder: Testing the effectiveness of a new approach to mother-infant psychotherapy. *Infant Mental Health Journal*, 20, 429–451.
- Connolly, L., Sharry, J., & Fitzpatrick, C. (2001). Evaluation of a group treatment programme for parents of children with behavioural disorders. *Child and Adolescent Mental Health*, 6, 159–165.
- Dadds, M. R., Spence, S. H., Holland, D. E., Barrett, P. M., & Laurens, K. R. (1997). Prevention and early intervention for anxiety disorders: A controlled trial. *Journal of Consulting and Clinical Psychology*, 65, 627–635.
- Dishion, T. J., & Andrews, D. W. (1995). Preventing escalation in problem behaviors with high-risk young adolescents:

- immediate and 1-year outcomes. *Journal of Consulting and Clinical Psychology*, 63, 538–548
- dos Santos, E. L. C., Marturano, E. M., de Almeida Motta, A. M., & Giurlana, A. G. (2003). Treating boys with low school achievement and behavior problems: Comparison of two kinds of intervention. *Psychological Reports*, 92, 105–116.
- Driskill, J. D. (2000). Structured child and parent groups with ADHD children: Evaluation of varying levels of parent involvement. *Dissertation Abstracts International: B*, 60 (9–B), 4884.
- Eddy, J. M., & Chamberlain, P. (2000). Family management and deviant peer association as mediators of the impact of treatment condition on youth antisocial behavior. *Journal of Consulting and Clinical Psychology*, 68, 857–863.
- Eddy, J. M., Whaley, R. B., & Chamberlain, P. (2004). The prevention of violent behavior by chronic and serious male juvenile offenders: A 2-year follow-up of a randomized clinical trial. *Journal of Emotional and Behavioral Disorders*, 12, 2–8.
- Eisler, I., Dare, C., Hodes, M., Russell, G., Dodge, E., & Le Grange, D. (2000). Family therapy for adolescent anorexia nervosa: The results of a controlled comparison of two family interventions. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41, 727–736.
- Flannery-Schroeder, E. C., & Kendall, P. C. (2000). Group and individual cognitive-behavioral treatments for youth with anxiety disorders: A randomized clinical trial. *Cognitive Therapy and Research*, 24, 251–278.
- Forgatch, M. S., & DeGarmo, D. S. (1999). Parenting through change: An effective prevention program for single mothers. *Journal of Consulting and Clinical Psychology*, 67, 711–724.
- Hebert, D. A. (2002). Evaluating cognitive-behavioral group treatment for disruptive adolescents in a special school. *Dissertation Abstracts International: B*, 63(1–B), 528.
- Heflinger, C. A., Bickman, L., Northrup, D., & Sonnichsen, S. (1997). A theory-driven intervention and evaluation to explore family caregiver empowerment. *Journal of Emotional and Behavioral Disorders*, 5, 184–191.
- Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology*, 65, 821–833.
- Hoagwood, K., Burns, B., Kiser, L., Ringeisen, H., & Schoenwald, S. K. (2001). Evidence-based practice in child and adolescent mental health services. *Psychiatric Services*, 52, 1179–1189.
- Hoath, F. E., & Sanders, M. R. (2002). A feasibility study of enhanced group triple P—Positive Parenting Program for parents of children with attention-deficit/hyperactivity disorder. *Behaviour Change*, 19, 191–206.
- Huey, S. J., Jr., Henggeler, S. W., Rowland, M. D., Halliday-Boykins, C. A., Cunningham, P. B., Pickrel, S. G., et al. (2004). Multisystemic therapy effects on attempted suicide by youths presenting psychiatric emergencies. *Journal of the American Academy of Child and Adolescent Psychiatry*, 43, 83–90
- Jensen, P. S., Hoagwood, K., & Trickett, E. (1999). Ivory Towers or Earthen Trenches? Community collaborations to foster real world research. *Applied Developmental Science*, 3, 306–212.
- Jocelyn, L. J., Casiro, O. G., Beattie, D., Bow, J., & Kneisz, J. (1998). Treatment of children with autism: A randomized controlled trial to evaluate a caregiver-based intervention program in community day-care centers. *Journal of Developmental and Behavioral Pediatrics*, 19, 326–334.
- Kaminer, Y., Blitz, C., Bursleson, J. A., Kadden, R. M., & Rounsaville, B. J. (1998). Measuring treatment process in cognitive-behavioral and interactional group therapies for adolescent substance abusers. *Journal of Nervous and Mental Diseases*, 186, 407–413.
- Kaminer, Y., Bursleson, J. A., Blitz, C., Sussman, J., & Rounsaville, B. J. (1998). Psychotherapies for adolescent substance abusers: A pilot study. *Journal of Nervous and Mental Diseases*, 186, 684–690.
- Kazdin, A. E. (2000). Developing a research agenda for child and adolescent psychotherapy. *Archives of General Psychiatry*, 57, 829–837.
- Kazdin, A. E., & Bass, D. (1989). Power to test differences between alternative treatments in comparative psychotherapy outcome research. *Journal of Consulting and Clinical Psychology*, 57, 138–147.
- Kazdin, A. E., & Nock, M. K. (2003). Delineating mechanisms of change in child and adolescent therapy: Methodological issues and research recommendations. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 44(8), 1116–1129.
- Kazdin, A. E., & Whitley, M. K. (2003). Treatment of parental stress to enhance therapeutic change among children referred for aggressive and antisocial behavior. *Journal of Consulting and Clinical Psychology*, 71, 504–515.
- King, N. J., Tonge, B. J., Mullen, P., Myerson, N., Heyne, D., Rollings, S., et al. (2000). Treating sexually abused children with posttraumatic-stress symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39, 1347–1355.
- Last, C. G., Hansen, C., & Franco, N. (1998). Cognitive-behavioral treatment of school phobia. *Journal of the American Academy of Child and Adolescent Psychiatry*, 37, 404–411.
- Liddle, H. A., Dakof, G. A., Parker, K., Diamond, G. S., Barrett, K., & Tejada, M. (2001). Multidimensional family therapy for adolescent drug abuse: Results of a randomized clinical trial. *American Journal of Drug and Alcohol Abuse*, 27, 651–688.
- McArdle, P., Moseley, D., Quibell, T., Johnson, R., Allen, A., Hammal, D., et al. (2002). School-based indicated prevention: A randomized trial of group therapy. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 43, 705–712.
- Mendlowitz, S. L., Manassis, K., Bradley, S., Scapillato, D., Miezitis, S., & Shaw, B. F. (1999). Cognitive behavioral group treatments in childhood anxiety disorders: The role of parental involvement. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38, 1223–1229.
- Merkel, K. M. (2000). Effect of treatment of distress symptoms in child victims of violent crime: A comparison of supportive therapy and cognitive-behavioral therapy. *Dissertation Abstracts International: B*, 60(9–B), 4897.
- Montgomery, P., Stores, G., & Wiggs, L. (2004). The relative efficacy of two brief treatments for sleep problems in young learning disabled (mentally retarded) children: A randomised controlled trial. *Archives of Disease in Childhood*, 89, 125–130.
- Muris, P., Meesters, C., & van Melick, M. (2002). Treatment of childhood anxiety disorders: A preliminary comparison between cognitive-behavioral group therapy and a psychological placebo intervention. *Journal of Behavior Therapy and Experimental Psychiatry*, 33, 143–158.
- Murray, L., Cooper, P. J., Wilson, A., & Romaniuk, H. (2003). Controlled trial of the short- and long-term effect of psychological treatment of post-partum depression. 2. Impact on the mother-child relationship and child outcome. *British Journal of Psychiatry*, 182, 420–427.
- Nicholson, J. M., & Sanders, M. R. (1999). Randomized controlled trial of behavioral family intervention for the treatment of child behavior problems in stepfamilies. *Journal of Divorce and Remarriage*, 30, 1–23.
- Ost, L. G., Svensson, L., Hellstrom, K., & Lindwall, R. (2001). One-session treatment of specific phobias in youths: A randomized clinical trial. *Journal of Consulting and Clinical Psychology*, 69, 814–824.
- Robin, A. R., Siegel, P. T., Moye, A. W., Gilroy, M., Dennis, A. B., & Sikand, A. (1999). A controlled comparison of family versus individual therapy for adolescents with anorexia nervosa.

- Journal of the American Academy of Child and Adolescent Psychiatry*, 38, 1482–1489.
- Rossello, J., & Bernal, G. (1999). The efficacy of cognitive, behavioral, and interpersonal treatments for depression in Puerto Rican adolescents. *Journal of Consulting and Clinical Psychology*, 67, 734–745.
- Scott, S., Spender, Q., Doolan, M., Jacobs, B., & Aspland, H. (2001). Multicentre controlled trial of parenting groups for childhood antisocial behaviour in clinical practice. *British Medical Journal*, 323, 194.
- Silverman, W. K., Kurtines, W. M., Ginsburg, G. S., Weems, C. F., Rabian, B., & Sarafini, L. T. (1999). Contingency management, self-control, and education support in the treatment of childhood phobic disorders: A randomized clinical trial. *Journal of Consulting and Clinical Psychology*, 67, 675–687.
- Snyder, K. V., Kymissis, P., & Kessler, K. (1999). Anger management for adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38, 1409–1416.
- Sonuga-Barke, E. J. S., Daley, D., Thompson, M., Laver-Bradbury, C., & Weeks, A. (2001). Parent-based therapies for preschool attention-deficit/hyperactivity disorder: A randomized, controlled trial with a community sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, 402–408.
- Spence, S. H., Donovan, C., & Brechman-Toussaint, M. (2000). The treatment of childhood social phobia: The effectiveness of a social-skills-training-based, cognitive-behavioral intervention, with and without parental involvement. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41, 713–726.
- Spoth, R. L., Redmond, C., & Shin, C. (2001). Randomized trial of brief family interventions for general populations: Adolescent substance-use outcomes four years following baseline. *Journal of Consulting and Clinical Psychology*, 69, 662–642.
- Storr, C. L., Ialongo, N. S., Kellam, S. G., & Anthony, J. C. (2002). A randomized controlled trial of two primary intervention strategies to prevent early onset tobacco smoking. *Drug and Alcohol Dependence*, 66, 51–60.
- Taylor, T. K., Schmidt, F., Pepler, D., & Hodgins, C. (1998). A comparison of eclectic treatment with Webster-Stratton's parents and children series in a children's mental health center: A randomized controlled trial. *Behavior Therapy*, 29, 221–240.
- Verduyn, C., Barrowclough, C., Roberts, J., Tarrier, N., & Harrington, R. (2003). Maternal depression and child behaviour problems: Randomised placebo-controlled trial of a cognitive-behavioural group intervention. *British Journal of Psychiatry*, 183, 342–348.
- Wayland, L. A. L. (2003). Treatment of children with problem behaviours: The efficacy of conjoint behavioural consultation versus videotape therapy and the impact on parent-teacher collaboration. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 64(1-B), 404.
- Weersing, V. R., & Weisz, J. R. (2002). Mechanisms of action of youth psychotherapy. *Journal of Child Psychology & Psychiatry and Allied Disciplines*, 43, 3–29.
- Weisz, J. R., & Jensen, P. S. (1999). Efficacy and effectiveness of psychotherapy and pharmacotherapy with children and adolescents. *Mental Health Services Research*, 1, 125–158.
- Weisz, J. R., Weiss, B., & Donenberg, G. R. (1992). The lab versus the clinic: Effects of child and adolescent psychotherapy. *American Psychologist*, 47, 1578–1585.
- Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in clinical trials. *Psychological Bulletin*, 130, 631–663.
- Wiggs, L., & Stores, G. (1999). Behavioural treatment for sleep problems in children with severe learning disabilities and challenging daytime behaviour: Effect on daytime behaviour. *Journal of Child Psychology and Psychiatry (Allied Disciplines)*, 40, 627–635.
- Wolchik, S. A., Sandler, I. N., Millsap, R. E., Plummer, B. A., Greene, S. M., Anderson, E. R., et al. (2002). Six-year follow-up of preventive interventions for children of divorce: a randomized controlled trial. *JAMA*, 288, 1874–1881.
- Wood, A., Harrington, R., & Moore, A. (1996). Controlled trial of a brief cognitive-behavioural intervention in adolescent patients with depressive disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 37, 737–746.