

Evidence-Based Psychotherapy for Children and Adolescents: Data from the Present and a Model for the Future

John R. Weisz & Jane Simpson Gray

Judge Baker Children's Center, Harvard Medical School, 53 Parker Hill Avenue Boston, MA 02120-3225, USA. E-mail: jweisz@jbcc.harvard.edu

What is the best way to help children cope with behavioural and emotional problems? This question has been a focus of concern — to children and the adults who care for them — across many centuries and in cultures around the world. In this article we examine the current state of efforts to help children by means of an array of non-medical interventions designed to alleviate psychological distress, reduce maladaptive behaviour, and/or increase adaptive behaviour. We refer to these interventions, collectively, as 'psychotherapy'. In the article, we will note some accomplishments of the effort to develop effective interventions through clinical research. Then we will note some concerns about these efforts, and suggest ways to address the concerns through an alternate model of intervention development and testing.

Keywords: Evidence-based psychotherapy; children; adolescents; evidence-based treatment

Background: historical notes on child psychotherapy

Efforts to help children are almost certainly as old as parenthood. At some point, however, such efforts morphed into a set of strategies now known collectively as psychotherapy. The historical roots of psychotherapy are difficult to trace, but some elements may be found in the work of early Greek philosophers, particularly Socrates (469–399 BCE), whose Socratic method is a core element of some current forms of psychotherapy. Of course, psychotherapy did not become a profession or even an area of study in its own right until centuries after Socrates' time. Some link the beginning of the profession to the work of Sigmund Freud (1856–1939), whose ideas about the importance of early experience opened the door for therapy with children, and later the study of therapy outcomes in children. An early historical marker was Freud's application of psychoanalysis to 'Little Hans', a boy whose fear of horses seemed irrational, and to his own daughter, Anna Freud (1895–1982), who later became a prominent child analyst in her own right. Child psychotherapy soon expanded to include other approaches. Behavioural therapy emerged in the 1920s, illustrated by Mary Cover Jones' (1924) work with Peter, a 2-year old whose fears of a white rabbit were treated using modelling and 'direct conditioning'. The research and writing of Skinner (1938, 1974) on operant conditioning fuelled diverse applications of behaviourism to child treatment, from token economies to behavioural parent training. Later, Aaron Beck (1964) introduced the idea that affect originates from cognition, and developed cognitive therapy to alleviate emotional distress, mainly in adults. These ideas, combined with the child-focused work of Donald Meichenbaum and others (Meichen-

baum & Goodman, 1971), contributed to the creation of cognitive-behavioural therapy for children.

In the more than three decades since, practice and research on child psychotherapy have continued to expand dramatically in form and scope. At the beginning of the 21st century, Kazdin (2002) identified 551 different named therapies that are used with children. Even this remarkable figure underestimates the range of approaches employed, because it excludes countless unnamed eclectic approaches used by practitioners around the world.

For a variety of reasons, the practice of psychotherapy has grown much faster than research on its effects. Moreover, early findings of the research that *was* done were not encouraging, either for therapy with adults (Eysenck, 1952) or children (Levitt, 1957, 1963). Research methods since the time of Eysenck and Levitt have grown more rigorous, with treatment procedures documented in written protocols or 'manuals' and tested via research designs that offer genuine inferential power. This sets the stage for useful evaluation of the evidence on child psychotherapies and their effects.

Assessing the effects of psychotherapies

Questions about therapy effects are answered using a number of research methods, including within-group, single-subject, and multiple baseline designs. The randomised controlled trial (RCT), a group comparison design in which participants are randomly assigned to groups, is the most widely accepted way to assess therapy outcomes. Most of the evidence presented in this article is based on RCTs and summarised in terms of effect size (ES), an index of the magnitude and direction of treatment impact. The ES is a standardised

mean difference—i.e. the difference on a particular outcome measure between the treatment and comparison groups after the treatment phase has ended, divided by the standard deviation of the measure. ES values are typically calculated so that positive values reflect an advantage of the treatment group over the comparison group, and negative values indicate that the group receiving the target treatment fared worse after treatment than the comparison group. Many researchers follow Cohen's (1988) designation of .20 as threshold for a 'small' effect, .50 for a 'medium' effect, and .80 for a 'large' effect.

ES values can be averaged across multiple outcome measures within a single study, to reach conclusions about the mean impact of a tested treatment within that study. Such study mean effects can, in turn, be averaged across multiple studies to generate a picture of the mean ES for a body of evidence — say, on treatment of depression, or even treatment across multiple conditions. When ES values are averaged across studies in this way, the process is called meta-analysis. Meta-analyses can help us understand how well psychotherapy works, and it can guide critical evaluation of the research literature.

A long and expanding list of meta-analyses has helped to build an understanding of child psychotherapy effects over the past 25 years. Four of these meta-analyses are particularly broad in their inclusion criteria, providing a picture of treatment effects averaged across a range of treated problems and treatment methods. In the earliest of these broad meta-analyses, Casey and Berman (1985) included outcome studies published between 1952 and 1983, involving children aged 12 and younger. The mean effect averaged across outcome measures used in treatment-control comparisons was .71, indicating that the average child who received a tested treatment fared better at post-treatment than 76% of children in control groups. Weisz et al. (1987) also analysed studies published between 1952 and 1983, but included adolescent studies. An effect size of .79 found in this meta-analysis indicated the average treated youth functioned better after treatment than 79% of youth in the control groups. In a third analysis including studies published between 1970 and 1988 with children between the ages of 4 and 18, Kazdin et al. (1990) found an effect size of .88; the average treated child outperformed 81% of control group children. Finally, Weisz et al. (1995) included studies published between 1967 and 1993 with children aged 2–18 and found a mean ES of .71; the average treated child showed better outcomes than 76% of children in control groups.

These four meta-analyses reported relatively similar effect sizes hovering around Cohen's (1988) threshold for a 'large' effect. Note that when weighted for sample size and heterogeneity of variance, ES means may actually be closer to 'medium' than 'large' (Weisz et al., 1995). In addition, McLeod and Weisz (2004) found that mean effects in doctoral dissertation trials, though positive, were markedly smaller than those published in peer-reviewed journals. This is consistent with concerns that the mean effects found in published studies may be inflated, to some degree, because of a bias toward publishing studies with the most favourable findings. It is also possible that studies lacking ample

funding (like most dissertations) are not able to include the 'bells and whistles' (e.g. extensive therapist training, monitoring of quality control and fidelity) that may enhance treatment effects.

Two additional points concern the specificity and durability of treatment effects. We found (Weisz et al., 1995) that effect sizes for the specific problem domains targeted by treatment were about twice as large as effect sizes for problems that were not targeted in treatment (e.g. anxiety treatment produced larger effects specifically on anxiety than on conduct problems). This suggests that tested treatments may be rather specific in their effects, not just general 'feeling better all over' phenomena. Second, the benefits of child therapy appear to be relatively stable over time, at least for several months after treatment. Our findings (Weisz et al., 1987, 1995) revealed that effect sizes at follow-up assessments, which have averaged 5–6 months after the end of treatment in the studies reviewed thus far, were similar to the effect sizes found immediately after treatment, suggesting rather durable treatment benefits. Holding power over longer periods of time remains to be determined. A recent meta-analysis focused specifically on the treatment of youth depression (Weisz, McCarty, & Valeri, 2006) found some evidence that beneficial effects tended to fade after about a year post-treatment, suggesting that treatment benefits do eventually diminish. This could indicate a need for booster sessions to restore the skills learned in treatment; alternatively, it could reflect the episodic nature of depression, which tends to remit and recur at intervals over time.

Specific treatments in the RCT literature

The overall picture of treatment benefit, conveyed through findings of broad-based meta-analyses, can be complemented by examples of specific forms of intervention that have shown replicated success in RCTs. For this purpose, we look to a recent review by Weisz, Hawley and Doss (2004), which focused on four problem clusters that account for a particularly large proportion of youth referrals for clinical care in most western countries, and for which a particularly extensive clinical trials literature exists: (1) anxiety-related problems and disorders; (2) depression-related problems and disorders; (3) attentional problems, impulsivity, and ADHD; and (4) conduct-related problems and disorders. With a focus on these four clusters, Weisz et al. (2004) searched the treatment outcome evidence base to identify those psychotherapies for which success has been demonstrated in multiple RCTs. The review included studies with children aged 3–18, published between 1965 and 2002. Studies were required to: (a) include a comparison of psychotherapy to a control group (waitlist, no treatment, placebo) or an alternative treatment; (b) involve a prospective design and random assignment of subjects to treatment and comparison conditions; (c) use participants selected for having psychological problems or maladaptive behaviour; and (d) include a post-treatment assessment of the psychological problem(s) or maladaptive behaviour for which participants were selected and treated. To ensure a focus on psychotherapy, studies that included medication conditions were excluded. A total of 298 studies

met our criteria for inclusion. Details of the search procedure and studies, and references to the specific treatment programs, can be found in Weisz et al. (2004).

Illustrative evidence-based treatments for anxiety and related conditions

Three forms of treatment for anxiety and related conditions have shown especially positive outcomes and particularly well-replicated success.

Modelling. Modelling is an approach that can take multiple forms. In 'live modelling', the anxious youngster observes a peer, the therapist, or some other model engaging in the behaviour that the child fears and suffering no adverse consequences. In 'symbolic modelling', a video or some other media product is used to present a model to anxious youngsters. In 'participant modelling', the fearful youth performs the feared activities in concert with the model, who guides and encourages the child along the way. Across the different forms of modelling, a frequent practice is to start at low intensity (e.g. sitting at some distance from the feared object) and gradually add increments (e.g. moving closer, touching the object, picking it up).

Reinforced exposure. Another approach to anxiety is called 'reinforced exposure'. It entails graduated steps of exposure to the feared object or situation, with the youngster rewarded for accomplishing each step. A challenge for the therapist is creating graduated steps of exposure that will be appropriately challenging but not so overwhelming that the youth refuses outright. For some fears of situations — darkness, for example — the graduating procedure may involve increments in time (e.g. the amount of time the child remains in a darkened room) or degree (e.g. degree of darkness tolerated). For other fears, the gradations may involve a series of increasingly lifelike and direct exposures. For example, fear of spiders might be treated with exposures that progress from looking at spiders from a distance to letting the spiders walk on the child's arm. In the various procedures, praise or rewards follow successful exposures.

Cognitive-behavioural therapy (CBT). Cognitive-behavioural therapy (CBT) has been used extensively to address youth anxiety. As the name implies, CBT entails efforts to identify and alter cognitions that contribute to the anxiety, and also to identify and alter maladaptive behaviour (such as avoidance of feared situations) that sustains the condition. Some forms of CBT have been used in treatment focused on individual youths, others with youths and their family members in various combinations. Applications of CBT at the individual youth level have ranged from procedurally simple approaches using self-talk, typically addressing specific fears, to much more complex multi-session programs, typically used to address multi-symptom anxiety disorders.

These more complex programs often include psychoeducation about the nature and course of anxiety, plus coping strategies for managing anxious thoughts (e.g. identifying and modifying thoughts that exaggerate danger) and bodily tension (e.g. deep breathing, relaxation exercises).

Illustrative evidence-based treatments for depressive symptoms and disorders

Given the close association between anxiety and depression, perhaps it is not surprising that some of the procedures used to good effect with anxiety reappear in the list for depression.

Cognitive-behavioural therapy. The most extensively tested and supported approach to youth depression treatment is CBT, which addresses cognitive, behavioural, and affective skill deficits. Youngsters typically learn to identify thoughts that are unrealistically negative (e.g. 'No one likes me.') and find more realistic, positive counter-thoughts (e.g. 'Not everyone likes me, but I have some really good friends.'). In addition, youngsters learn to identify activities that boost their mood, and then schedule those activities when mood enhancement is needed. Other skills are likely to be included as well, such as coping with stress by using relaxation techniques, and interrupting negative thoughts to prevent getting stuck in unproductive rumination.

Interpersonal therapy. Relatively recent evidence points to beneficial effects of Interpersonal Therapy for Adolescents (IPT-A). The approach builds on the premise that depression occurs in an interpersonal context, and that an improved understanding and renegotiation of this context can lead to recovery. Therapy focuses on one or two problem areas, which may include grief and loss, interpersonal role disputes, role transitions, or interpersonal deficits. The therapist and adolescent work directly on the identified problem areas, while the adolescent monitors depressive symptoms and uses various strategies (e.g. role play) to try out various forms of behaviour change.

Illustrative evidence-based treatments for ADHD and related problems

Turning to ADHD and related problems, we should note that the largest treatment effect sizes are associated with stimulant medication, but that some psychological interventions have shown measurable benefit.

Cognitive behavioural therapy. One such intervention is CBT, which sometimes involves teaching children self-talk to help them focus and reduce their distractibility and impulsivity, thus improving their performance on tasks requiring concentration. Children may also learn to use coping self-talk when they make mistakes (e.g. 'Oops, I made a mistake. That's OK, everybody does. I'll just get back to work.'). In an added twist, children are sometimes taught to monitor and evaluate their own performance and give themselves appropriate praise and rewards when they do well.

Relaxation and biofeedback training. Our review (Weisz et al., 2004) found surprisingly strong effects for ADHD interventions that were focused on calming or relaxation. In one simple variant of this approach, children simply listened to audiotapes with relaxation-meditation instructions across five half-hour sessions. Some programs have mixed group relaxation training with large muscle exercise followed by rest and deep breathing. Others have combined relaxation training

with biofeedback; for example, one program had hyperactive children listen to parts of a relaxation-induction audiotope while they received visual and auditory feedback (e.g. clicks) from an electromyometer that measured muscle activity in the frontalis area. Children were told that their job was to keep their bodies relaxed enough that the feedback would show a sustained pattern of low activity. Just 3–4 sessions, less than half an hour each, were associated with significantly reduced impulsivity and gains in attention on a post-intervention performance task.

Behavioural parent and teacher training. A widely-practised approach to treatment of problems related to ADHD involves teaching parents to create and maintain environments in which desirable child behaviour (e.g. obeying adults, thinking before acting) is rewarded and undesirable behaviour (e.g. disobedience, impulsive acts) is ignored or met with some aversive consequence (e.g. loss of a privilege). This basic concept of contingency management is the centrepiece of a number of different behavioural parent training programs. Because so many ADHD-related problems appear in the school setting, an important complement to parent training programs is teacher involvement, sometimes in the context of classroom-focused contingencies that will reinforce self-control, attention to school work, and appropriate social behaviour with teachers and peers. Because these programs are used for children with ADHD, when it co-occurs with other conduct-related problems and disorders, we discuss the programs in greater detail in the next section.

Illustrative evidence-based treatments for conduct-related problems and disorders

Treatment of conduct-related problems and disorders has received a great deal of attention in the research literature, with a number of treatments showing substantial positive effects that have been replicated in multiple RCTs. Here we give a few examples.

Youth-focused operant treatment. Treatments in the operant category are those emphasizing reinforcement contingencies. These have included strategic ignoring of misbehaviour (depriving children of attention for such behaviour, and attending closely when children behave well) as well as strategic rewarding (using praise, token, or tangible rewards to reinforce appropriate behaviour). In some settings operant microeconomies have been created. For example, in one youth correctional institution, appropriate behaviour earned institutional ‘dollars’ and points, with dollars used to purchase objects, services, and recreational activities, and points accumulated to earn a recommendation for parole and release. Charts showing points earned by each boy were placed in conspicuous locations as an additional form of reinforcement. The program had significant effects on both behaviour ratings in the institution and recidivism after discharge.

Cognitive-behavioural therapy (e.g. problem-solving skills). CBT for youth conduct problems has taken a variety of forms. One particularly successful example involves training aggressive youngsters to do problem-

solving in situations that might typically lead to aggressive interactions, including fights. Youngsters learn, through a series of game-like sessions with their therapist, to identify the problem, brainstorm multiple solutions, evaluate the proposed solutions in terms of their likely consequences, make a choice, and then evaluate how successful they were in solving the problem. In a related treatment approach using CBT principles, children learn anger management that entails self-calming followed by steps of constructive problem-solving. In both sets of CBT approaches, youngsters learn to apply their new skills to interpersonal situations, including the kinds of situations that have previously led to aggression.

Behavioural parent training. Behavioural parent training is the most extensively tested and supported form of treatment for youth conduct problems and disorders. Although each program has unique characteristics, some common features cut across most of the programs that have shown beneficial effects. In most of the programs, for example, parents learn basic behavioural principles relevant to child rearing, such as the role of attention in the maintenance of both positive and negative behaviours. Parents learn to structure close, highly attentive interactions in ways that strengthen their relationship with their child, and thus increase the impact of parental attention and praise. Parents also learn to use rewarding consequences (attention, praise, points, or privileges) to increase the frequency of appropriate behaviours. To reduce the frequency of unwanted behaviours, parents learn ways of making those behaviours less rewarding; in some cases, ignoring (turning away or looking away) may be sufficient; in other cases, use of a response cost (e.g. deduction of points or loss of privileges) may be required, and in other cases ‘time out’, in which the child is withdrawn from opportunities for reward.

Overview of evidence-based treatments

The illustrative treatments just described are part of a much larger array of interventions [See Weisz et al. (2004) for a fuller description.] that have emerged from more than four decades of research. Although we focused only on problems and disorders in four broad clusters - anxiety, depression, ADHD, and conduct - we found 298 acceptable RCTs, encompassing more than 300 treatment programs that showed at least some significant benefit relative to a control group or alternate treatment group. The evidence base supporting these treatments is certainly not without limitations, and we have highlighted a number of these in previous work (Weisz, Doss, & Hawley, 2005). However, the treatments emerging from these decades of research do warrant close attention, and many of these appear to justify implementation in everyday practice. As we will see in the next section, such implementation does not appear to be happening to a very significant degree.

Despite the evidence favouring EBTs, dissemination is not moving very fast

Despite the replicated beneficial effects of several treatments for youth psychopathology, most of the

tested treatment programs noted above have not made their way into standard everyday clinical practice. The evidence available to us (Ho et al., 2007; Martin et al., 2007; McLeod, 2001; McLeod & Weisz, 2004) suggests that most everyday clinical practice continues to be characterised by interventions that do not rely on behavioural or cognitive-behavioural principles and are not derived from the clinical trials literature. It also appears that evidence-based practices are not showing up increasingly in clinical training, at least within North American doctoral training programs and internship sites in clinical psychology. In a recent survey of directors of clinical training for doctoral training programs as well as directors of pre-doctoral internship sites in the United States and Canada, Woody, Weisz and McLean (2005) assessed changes from 1993 to 2003 in what is being taught to future practitioners in the field. The findings showed that most U.S. and Canadian doctoral programs offer supervised training for cognitive behavioural treatment of depression and generalised anxiety disorder, but treatments that clearly have significant benefits for children with oppositional behaviour and conduct disorder (e.g. behavioural parent training) are actually being taught and supervised less than was true 10 years ago. In fact, the survey revealed that *most* evidence-based approaches, with the exception of CBT and IPT-A for depressed adolescents, are being taught and supervised at doctoral training sites less than they were a decade ago. In 1993, most doctoral programs provided supervision for 11 out of 22 target EBTs; in 2003, the figure dropped to 5 of the EBTs.

Complementing these data from North America, a recent UK survey (Stallard et al., 2007) found that CBT was the dominant approach of only 20% of practitioner respondents. While the data available to us from North America and the UK do not cover all practice or training sites, or all evidence-based treatment models, the partial picture they present is that the evidence-based treatments emerging from four decades of research are not making their way into training or everyday practice at a very rapid pace, and in fact may be losing ground, to some extent. How can this be? Why wouldn't those who chose careers to help children, and those who are training the next generation of clinicians, be eager to adopt practices that research indicates will be effective with children?

A number of possible reasons have been identified previously

Our research with practitioner partners and in practice settings has suggested a number of possible reasons (Weisz, 2004; Weisz, Sandler et al., 2005). They include the fact that most evidence-based treatments are designed for specific disorders and target problems that may only be accurately identified through evidence-based assessment. The challenges of doing such assessment in practice settings may mean that the precise match between treatment and treated problem - required for appropriate and effective use of EBTs - is difficult to achieve in everyday practice. A second possible reason noted previously is that most EBTs have been designed for single problems or disorders; the fact that most children referred to everyday clinical care

settings present with multiple co-occurring problems and disorders may lead many practitioners to doubt the appropriateness of the EBTs for their clientele. A third possible reason that we have noted previously is that EBTs often ask a lot of therapists who take them on - learning a manual (often a very lengthy one), preparing for each session in a way that would be rare in most everyday practice, working from an agenda rather than letting the session take its own course, and finding the creativity and charisma needed to make the agenda-guided content engaging to the child. Another possible reason, and a related one, is that EBTs often ask a lot of the young clients in therapy, requiring that they learn multiple skills, do homework to practice those skills, and figure out when to use which skills when complex problems arise in everyday life. None of these challenges, nor any of the others we have noted in previous writings (Weisz, 2004; Weisz, Doss, & Hawley, 2005) is insurmountable, and we have offered a number of suggestions for addressing each one.

What we would like to do in the remainder of this article is focus attention on one general theme that cuts across many of the specific reasons. This theme is the notion that how we go about developing and testing the treatments known as 'evidence-based' can have a huge impact on their likelihood of being adopted by practitioners and trainers, their readiness for use in the world of everyday practice, and their effectiveness when used and tested in that world. To develop this theme, let us begin with a general concern about the EBT research base that our research group has emphasised frequently over the past decade...

EBT conditions differ from those of everyday clinical care

As we have noted previously, the conditions in which most EBTs have been developed and tested tend to differ markedly from the conditions of actual clinical practice. This limits our ability to extrapolate from the research findings to everyday clinical care of children. It also fuels practitioner concerns that the treatments that look so promising in research trials may not actually be so ready for prime time in real-world clinical practice. Children referred to clinical care settings are apt to differ from children recruited for efficacy trials in diverse ways, including severity and family adversity (Hammen et al., 1999; Southam-Gerow, Weisz, & Kendall, 2003). Therapists employed as staff clinicians in clinics tend to differ in professional goals, work pressures, and numerous other ways from therapists employed as research clinicians in clinical trials (see Palinkas et al., in press; Weisz & Addis, 2006). As one example, the staff clinicians who work in practice settings typically treat a broad array of problems and disorders, often within the same day, so complicating any effort to develop or use concentrated expertise in a single treatment for a single disorder. In addition, clinical practice and research settings tend to differ in a broad range of ways that could affect child outcomes. In practice settings, for example, a variety of time, productivity, financial, and other work pressures prevail that can undermine efforts to optimise the treatment environment for children and families (Weisz & Addis, 2006). Such differences between the children, therapists, and settings of research and those of everyday

clinical care highlight the need to test EBTs under conditions that match those of clinical practice. Only such tests, it could be argued, can tell us whether EBTs are ready for prime time in everyday clinical care.

We have found (Weisz, Doss, & Hawley, 2005) that the research literature on evidence-based treatments does a poor job of exploring how evidence-based treatments perform in real-world practice conditions. Most of the studies in the evidence base were conducted in settings constructed for research (e.g. university laboratories, university-based clinics, rooms in a school set aside for the study), with treatment provided to youths who were recruited or screened specifically for the study, and with treatment not provided by practising clinicians but rather by graduate students or other researcher-trained and researcher-employed therapists. Across the studies, only 13% of the study samples were actually clinically-referred, treatment-seeking youth; only 19% of the studies employed even one practising clinician as a treatment provider; and in only 4% of the studies was treatment provided in an actual clinical service setting separate from the research program. Summing across dimensions, as shown in Table 1, we found that only 1% of the studies reviewed included some clinically referred children, at least one practising clinician, and some treatment carried out in a clinical service setting. Thus, we found a literature that provided relatively little information about how EBTs fare under conditions like those of everyday clinical care.

Most EBTs have never been tested in comparison to usual clinical care

A closely related concern is that the extant literature provides only limited information on a key question: *Do EBTs generate more favourable outcomes than the treatments children would otherwise receive in their usual clinical care?* In some respects, this is the most basic question many in clinical practice might ask. Building competence in a typical manual-guided treatment protocol requires a major investment of time and

funds, and is likely to require considerable sacrifice on the part of most clinicians who are managing the time demands of a practice career. Given this state of affairs, it seems quite reasonable for practitioners to ask whether acquiring skill in a particular evidence-based treatment will lead to better clinical outcomes for children than current practices do. The RCT evidence base provides relatively few answers to this question, but it is not completely silent, as we shall see next.

Results of studies that have compared EBTs to usual care raise important questions

Our research group (Weisz, Jensen-Doss, & Hawley, 2006) judged this question to be so important that we conducted a meta-analysis devoted solely to answering it. From the outset, it was not clear what we would find. Although proponents of evidence-based treatments have suggested that they should be used in preference to usual clinical care, a number of authors have expressed concerns about EBTs and the manuals used to guide them, concerns that raise questions about whether these structured approaches to treatment could be as effective as everyday clinical care. Various writers have suggested that EBTs (a) have been developed and tested with relatively simple, often subclinical cases, and thus may not work well with the complex and severe cases often seen in usual clinical care; (b) have been designed for single problems or diagnoses and thus may not fare so well when they confront co-occurring problems and comorbid diagnoses, which are common in usual clinical care; (c) are so strictly and uniformly structured that they make it hard to individualise treatment to meet distinctive client needs; (d) are so formulaic that they constrain therapist creativity in addressing unusual or unexpected events in client's lives; and (e) are so lacking in spontaneity and flexibility that they interfere with rapport-building and development of a good therapeutic relationship. Several of the concerns reflect the view that evidence-based treatments may not be well-suited to the challenge of treating clinically-referred individuals in the context of

Table 1. Clinical (un)representativeness of the youth treatment evidence base: youths, therapists, and treatment settings employed in outcome studies

	Anxiety	Depression	ADHD	Conduct	All Studies
How youths enrolled in the study:					
% Recruited, not treatment-seeking	90.24	77.78	87.50	60.42	76.69
% Treatment-seeking, clinic-referred	3.66	16.67	12.50	19.79	12.71
% Required via court/justice system	1.22	0.00	0.00	17.71	7.63
% Studies not reporting	4.88	5.56	0.00	2.08	2.97
Who provided the treatment:					
% With any researchers/grads	57.32	47.06	45.00	38.54	47.21
% With any paraprofessionals	20.73	11.11	12.50	22.92	19.49
% With any practising clinicians	1.22	55.56	10.00	30.21	18.64
% Studies not reporting	28.05	11.11	40.00	19.79	25.42
Setting where treatment took place:					
% Research settings	50.00	44.44	42.50	48.96	47.88
% Clinical service settings	2.44	5.56	0.00	7.29	4.24
% Correctional settings	1.22	0.00	0.00	7.29	3.39
% Studies not reporting	46.34	50.00	55.00	37.50	44.49
Representativeness sum:					
% Reporting no rep factors	92.68	38.89	77.50	55.21	70.76
% Reporting one rep factor	7.32	50.00	22.50	34.38	24.15
% Reporting two rep factors	0.00	5.56	0.00	8.33	3.81
% Reporting all three rep factors	0.00	5.56	0.00	2.08	1.27

usual clinical care (Addis & Krasnow, 2000; Addis & Waltz, 2002; Garfield, 1996; Havik & VandenBos, 1996; Strupp & Anderson, 1997; Westen, Novotny, & Thompson-Brenner, 2004a, 2004b). Further, experts on culture and ethnicity have voiced concerns that EBTs may not have been adequately adapted for, and may not so work well with, members of ethnic minority groups (Bernal & Scharron-Del-Rio, 2001; Gray-Little & Kaplan, 2000; Hall, 2001; Sue, 2003).

When we searched for studies that could shed some light on whether evidence-based interventions generally outperformed usual care, we were able to find 32 methodologically acceptable randomised trials (23 published articles and 9 dissertations) that directly compared EBTs to usual care. The studies spanned the years 1973–2004, with mean sample age ranging from 6–17; about three-quarters of the evidence-based treatments were behavioural; most of the remainder were systems- oriented interventions such as multi-systemic therapy. In general, the EBTs did outperform usual care. The mean effect size was .30 at immediate post-treatment, falling between conventional cutoffs for ‘small’ and ‘medium’, with a slight increase to .38 at follow-up (but not all studies included a follow-up

assessment). The superiority of the EBTs was not reduced by high levels of youth severity or by inclusion of minority youths, thus providing no support for at least two of the concerns critics of evidence-based treatments have raised.

On the other hand, the findings did not show dramatic superiority of EBTs over usual care. In fact, in a number of the studies, the group difference showed outcomes for usual care that were comparable or superior to outcomes of the evidence-based practice to which they were compared. This variability in outcomes, shown in Figure 1, suggests some important take-home messages for our field.

Messages of the EBT vs. usual care comparison

Among the most important messages of these findings on usual care, three warrant special attention here.

An expanded EBT vs. usual care database may inform consumer choices

The first message is that the EBT vs. usual care research genre has the potential to support wise

EBT versus usual care effect sizes

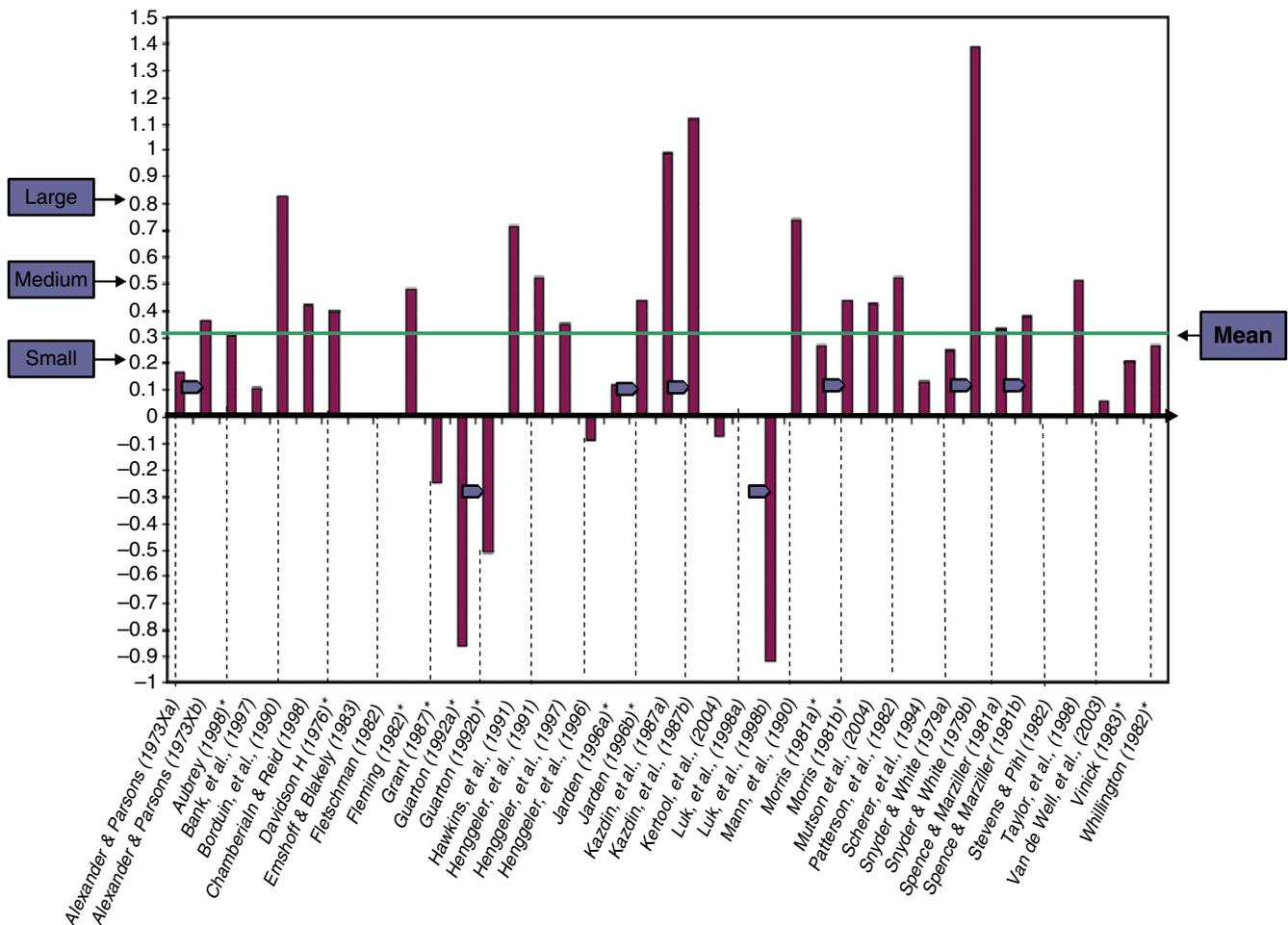


Figure 1. Effect sizes of individual studies included in the Weisz, Jensen-Doss and Hawley (2006) meta-analysis of usual care versus evidence-based treatment comparisons. Horizontal bar at .30 shows mean effect size across the full study set. [Figure provided by John R. Weisz]

consumerism by policy-makers, provider organisations, and individual clinicians who are considering adopting evidence-based practices. The skill-building needed for competence in such practices can be substantial, time-consuming, and sometimes expensive. This argues for judicious selection of the specific EBTs in which so much will be invested by those who set out to learn them.

The highly variable nature of the findings shown in Figure 1 - with most findings favouring EBTs but other findings showing instances in which usual care outperformed or approximated the impact of EBTs - suggests that we cannot assume that all evidence-based treatments are superior to what all clinicians are doing routinely in usual care. Even studies testing EBTs of the *same name* (e.g. anger management, and problem solving skills) showed different effects depending on (a) the specific treatment procedures used, and (b) to what form of usual care the EBTs were compared. What this suggests, from a clinical policy perspective, is that whether a specific form of usual care should be replaced with a particular evidence-based treatment may need to depend on (a) not just the generic form of the EBT being considered (e.g. problem-solving skills) but the specific protocol or manual, and (b) what the current prevailing forms of usual care consist of in the setting in question. Ideally, priority would be given to specific EBT protocols that have been shown to outperform the forms of usual care that prevail in the setting where the change is contemplated.

Of course, we are a long way from the kind of evidence base that will be needed to fully inform such decision-making. With just 32 direct randomised comparisons of specific EBTs with particular forms of usual care, we have only selective glimpses of the EBT vs. usual care terrain, and certainly not the kind of comprehensive and robust evidence base needed by policy makers and clinical administrators to make strategic decisions about changes in practice patterns. However, the sketchy picture provided by Figure 1 does forecast what a more fully populated picture might ultimately consist of: a large number of EBTs, with different specific protocols, each compared to multiple different forms of usual clinical care. Such a variegated display of evidence, across a broad array of trials, could be an invaluable guide to those who are considering changes in their current practices.

However, one additional ingredient will be needed, one that is now sadly lacking in the current evidence base: clear documentation of the usual care procedures to which EBTs have been compared. This brings us to a second take-home message of the EBT vs usual care comparison shown in Figure 1.

Studying usual care with care may lead us to new and effective treatments

A second message is that research like that represented in Figure 1 has the potential to lead us to effective treatments that are not currently a part of the 'evidence-based treatments' armamentarium. Those forms of usual care that markedly outperform a current EBT may well be approaches to intervention that warrant documenting in written protocols and testing in their own right. Some of these protocols might eventually make their way into expanded lists of EBTs. Herein lies

a potentially exciting strategy for discovering real-world interventions that work.

As exciting as this strategy is in principle, our review of the evidence thus far shows that a major shift will be needed to realise the potential. That shift would involve treating usual care not just as a 'control condition' but rather as a set of interventions that warrant documentation and study in their own right. In general, the studies represented in Figure 1 did a good job of describing the EBTs and a very poor job of describing the contents and procedures of the usual care conditions to which the EBTs were compared. It was as if the usual care conditions were regarded as relatively uninteresting and unimportant except as 'control' conditions for comparison to EBTs. Indeed, in most cases we were unable to determine what 'usual care' actually consisted of, even in the most generic sense. This is a problem that needs to be rectified if the 'discovery' potential of research on usual care is to be realised. Knowing that some forms of usual care performed well in comparison to treatments identified as empirically supported makes it especially unfortunate that we have no description of (much less written protocols for) those forms of usual care. In the next generation of empirical work, we need to take seriously the challenge of pairing evidence-based practice with practice-based evidence, thus to open up a potentially rich source of information on promising and perhaps even potent treatments.

The challenge of implementation may require a new model of treatment development and testing

The ups and downs of EBTs as reflected in Figure 1 should also be interpreted in light of the decades of research on implementation, particularly implementation of tested practices in settings other than those in which they were originally tested (Fixsen et al., 2005). That body of research strongly suggests that any practices - including EBTs - risk a loss of potency and impact when moved into contexts that are new and different from those where they were previously used and tested. Indeed, successful implementation typically requires much more than simply relocating the intervention; instead, multiple steps of intervention adaptation, intervener selection and coaching, and perhaps even recipient and organisational preparation may be required (Fixsen et al., 2005). The modest mean effect size shown in Figure 1 and the fact that a number of EBTs did not outperform usual care might be seen as illustrating the general challenge of implementation. A case can be made that this challenge is particularly serious in cases where the implementation context differs most markedly from the original development and testing context.

This brings us back to a concern noted earlier in this paper, one that our research group has stressed for many years - i.e. that the contexts and conditions under which most EBTs have been developed and tested tend to differ rather markedly from the conditions of everyday clinical care to which treatment developers wish to see their protocols deployed. In other words, the ways most evidence-based psychotherapies for children have been developed and tested do not expose these therapies to the full array of factors present in everyday, real-world clinical care. As a consequence, we have argued, the therapies emerging from such development and

testing procedures may not be very robust in the face of real-world treatment conditions, and in fact may not look so strong in relation to the everyday clinical care that takes place under those conditions. This, of course, is quite consistent with the findings of our EBT vs. usual care meta-analysis (Weisz, Jenson-Doss, & Hawley, 2006), as depicted in Figure 1.

Taken together, these concerns and findings raise the question of whether our field needs a new model of treatment development and testing. We certainly believe that empirical testing is critical to the development of beneficial treatments. However, a critical question for the field is what *approach* to empirical testing will give us the strongest treatments that are most robust in actual clinical practice. Our recent findings lead us to wonder whether the primary model that has guided the development and testing of psychotherapies for decades is well-suited to the task.

The successive efficacy trial model, derived from medical-pharmaceutical research. Most of the child psychotherapies now identified as evidence-based have been developed and tested via a model quite similar to the one that has guided medical and pharmaceutical research for years. That model - we will call it the 'successive efficacy trial (SET) model' - may work reasonably well for biological interventions, but it may not be quite as well-suited to the production of clinic-ready psychotherapies. The SET model involves a stepwise sequence in which experimental treatments and their protocols are first developed in the laboratory, and then tested via an extensive array of 'efficacy studies'. Efficacy studies use experimental control to test treatment impact under carefully arranged idealised conditions - e.g. with just the right kind of clients selected, often with the most troubling comorbidities excluded, with treatment done by particularly skilled therapists who are selected and paid by the researcher and trained to deliver the target treatment (and that treatment alone) as faithfully as possible, and with arrangements designed to keep therapists functioning at their best and treated clients engaged and attending. Efficacy studies are contrasted with 'effectiveness studies', in which intervention effects are assessed under ordinary clinical conditions, with treatment delivered to 'average' or representative patients or clients, by 'average' or representative practitioners, working under conditions that reflect typical practice realities (e.g. large caseloads, clinic productivity pressures, frequent appointment no-shows). Given the nature of the research conducted to date, we know a good deal about the *efficacy* of our tested treatments; we know far less about their *effectiveness*. In the SET model, a sometimes lengthy series of efficacy trials may be conducted, perhaps testing various ways of structuring and delivering an intervention, perhaps dismantling the intervention to test the relative impact of different components, and perhaps testing such variations as post-treatment booster sessions or the inclusion of additional family members. Typically, it is only after a sometimes extended series of such efficacy trials that the intervention is brought into community settings 'to measure the public health impact' (Greenwald & Cullen, 1984; National Institutes of Health, 1994).

From an implementation perspective, this SET approach may work reasonably well for interventions that operate directly on the biological system - e.g. psychoactive drugs and medical procedures for cancer treatment - or other treatment targets for which differences between research and clinic conditions may not greatly alter the intervention effect. Under such conditions, placing effectiveness tests at the very end of a series of efficacy trials may be reasonable, because the intervention may require relatively little modification to be rendered effective in real-world clinical care contexts. [Evidence on stimulant treatment of ADHD in community settings does suggest, though, that the 'bringing to scale' process can be quite necessary, even if the treatment involved is a medication (MTA Cooperative Group, 1999).]

However, the gap between research and practice is often quite wide in the case of psychotherapies, arguably much wider than with biologically-focused treatments. For psychotherapies the gap includes (1) psychological and social characteristics of the treated individuals (e.g. clinic-referred youth tend to be more severely disturbed, more likely to meet criteria for a diagnosis, more likely to have comorbidities, and more likely to miss appointments or drop out of treatment); (2) characteristics of their families (e.g. more parental psychopathology, family life event stressors, and perhaps even child maltreatment); (3) reasons for seeking treatment (e.g. not recruited through ads or screening, but referred by caregivers because of serious problems or family crisis, perhaps even court-ordered); (4) the settings in which treatment is carried out (e.g. more financial forms to complete, more bureaucracy, and sometimes a less welcoming approach in the clinic); (5) the therapists who provide treatment (e.g. not graduate students or research assistants hired by and loyal to the advisor/employer and committed to her/his treatment program, but rather staff therapists who barely know the treatment developer or the specific treatment, and who may prefer different treatment methods); (6) the incentive system (e.g. not paid by the treatment developer to deliver the treatment with close adherence to the manual, but paid by the clinic to see many cases and with no method prescribed); and (7) the conditions under which therapists deliver the treatment (not graduate students' flexible time, but strict productivity requirements, paperwork to complete, insurance requirements to satisfy, and little time to learn a manual or adhere closely to it).

Perhaps the numerous differences between psychotherapy in efficacy research and psychotherapy in actual clinical practice are too pronounced to be bridged as simply the final step after a series of efficacy experiments. [And, truth be told, this final step has not actually been taken yet for most protocol-guided EBTs.] Perhaps the number of dimensions along which treatment would need to be adjusted to span the lab-to-clinic gap makes the task of moving from efficacy trials to effectiveness tests so complex that the task needs to be made an integral part of the treatment development process. In fact, the very real-world factors that efficacy trial investigators might view as a nuisance (e.g. child comorbidity, parent pathology, life stresses that produce no-shows and dropouts, therapists with heavy caseloads) and thus attempt to avoid (e.g. by recruiting

and screening cases, applying exclusion criteria, hiring their own therapists), may in fact be precisely the kinds of factors that need to be included, understood, and addressed, if psychotherapy treatment protocols are to be created that fit well into everyday clinical practice. Treatments that cannot cope with these real-world factors may not fare so well in practice, no matter how beneficial they are in efficacy trials.

A related point is that implementation of EBTs in real-world practice settings may require certain interventions in the settings themselves, to eliminate obstacles to effective use of the EBTs. As an example, treatments that are built on weekly instalments of skill-building with children or parents may require new family engagement procedures to generate faithful attendance in a setting where appointments are often missed in usual care. As another example, new interventions that do not fit neatly within the organisational structure or standard procedures of a clinic (e.g. because they require approaches to assessment or supervision that differ from the clinic routine) may only be workable and effective if paired with procedures for organisational problem-solving. Development of effective versions of such setting-focused interventions would seem to require situations in which investigators take EBTs into the real-world settings for which those treatments are ultimately intended and identify the setting-focused interventions that are needed to make the target treatments succeed.

A deployment-focused model of treatment development and testing

For those who seek to create treatments that are robust in clinical practice, there may be value in considering a shift from the traditional SET approach to a different model that we have called the 'deployment-focused model (DFM)' of treatment development and testing (Weisz, 2004). This model is designed to (a) bring treatments into the crucible of clinical practice early in their development, and (b) construe testing in the practice setting as a sequential process, not as a single final phase. A primary goal of the DFM is to create a process by which treatments that show beneficial effects in efficacy trials can be adapted for testing and use in the practice contexts for which they are ultimately intended. A testable premise underlying the model is that an evidence-based treatment that has fared quite well in efficacy trials may have potential to be beneficial in a practice context, but that the potential is most likely to be realised if the treatment has undergone subsequent adaptation to practice conditions.

The steps of the DFM have been detailed elsewhere (Weisz, 2004), so they can be briefly summarised here. In *Step 1*, theory and evidence on the nature and treatment of the target condition, the clinical literature (e.g. published case studies), and input from clinicians who have treated youngsters with that condition, are used to guide the design of treatment components and to plan clinically sensitive ways of presenting those components within a written protocol. In *Step 2*, an initial group-design *efficacy* trial is used to assess whether the treatment (compared to a control group) can produce beneficial effects with recruited symptomatic youth who are treated under controlled labo-

ratory conditions. The goal is to determine whether the program is sufficiently promising when delivered under optimal experimental conditions to warrant further development and testing in clinical practice settings with referred youth. *Step 3* is a series of single-case pilot tests with clinic-referred individuals, treated in clinical settings, with successive modifications made in the specifics of the treatment protocol and procedures, to fit the protocol to the needs of clinical care while adhering to the principles that guide the treatment protocol. *Step 4* is a series of group-design effectiveness tests in which different effectiveness elements - referred children, practitioners providing the treatment, treatment taking place in a practice setting - are put into place within the testing process. *Step 5* is full-scale effectiveness testing, with the target treatment provided to clinically-referred youngsters, in the clinical care settings, by staff practitioners who have been trained to proficiency in the protocol. Subsequent steps include studies of the relation between the treatment program and the practice contexts in which it is employed - e.g. to identify factors that predict (a) the likelihood that practitioners will use the protocol, (b) the degree to which those who do use it will adhere to the manual, and (c) the extent to which use of the protocol is sustained in the setting over time.

The deployment-focused model and implementation science

The DFM connects nicely to the emerging field of implementation research. The recent review of implementation studies noted previously (Fixsen et al., 2005) shows that when interventions that succeed in one setting are moved to a very different setting, the transitions are often quite challenging. Indeed, it is common for such efforts to fail initially, and for considerable work to be required before a successful result is achieved. This pattern underscores an important fact relevant to evidence-based treatments: When these treatments are not successful in new contexts - such as everyday clinical care settings - that outcome does not necessarily mean the treatments lack potential for real-world impact. Instead, it may simply underscore the fact that implementation is challenging, and that multiple steps of intervention adaptation, intervener selection and coaching, and recipient and organisational preparation may be needed before the goal of effectiveness in the new context is achieved (see Fixsen et al., 2005). If that new context is everyday clinical care of children, surely the goal is a worthy one.

Acknowledgements

The work described in this article was supported by grants from the John D. and Catherine T. MacArthur Foundation (*Child STEPs, the Research Network on Youth Mental Health*) and the US National Institute of Mental Health (R01-MH57347, R21-MH63302, and R01-MH068806). The paper was originally presented by John Weisz as the ACAMH Jack Tizard lecture, and is reprinted in ACAMH Occasional Papers Number 27, *Evidence-based psychotherapies in child and adolescent mental health practice*.

References

- Addis, M. E., & Krasnow, A. D. (2000). A national survey of practicing psychologists' attitudes toward psychotherapy treatment manuals. *Journal of Consulting and Clinical Psychology, 68*, 331–339.
- Addis, M. E., & Waltz, J. (2002). Implicit and untested assumptions about the role of psychotherapy treatment manuals in evidence-based mental health practice. *Clinical Psychology: Science and Practice, 9*, 421–424.
- Beck, A. T. (1964). Thinking and depression: Theory and therapy. *Archives of General Psychiatry, 10*, 561–571.
- Bernal, G., & Scharron-Del-Rio, M. R. (2001). Are empirically supported treatments valid for ethnic minorities? Toward an alternative approach for treatment research. *Cultural Diversity and Ethnic Minority Psychology, 7*, 328–342.
- Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin, 98*, 388–400.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd edition. Hillsdale, NJ: Erlbaum.
- Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology, 16*, 319–324.
- Fixsen, D. L., Naoom, S. F., Blasé, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tamh34 FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).
- Garfield, S. L. (1996). Some problems associated with "validated" forms of psychotherapy. *Clinical Psychology: Science and Practice, 3*, 218–229.
- Gray-Little, B., & Kaplan, D. (2000). Race and ethnicity in psychotherapy research. In C. R. Snyder & R. E. Ingram (Eds.), *Handbook of psychological change: Psychotherapy processes and practices for the 21st century* (pp. 591–613). New York: Wiley.
- Greenwald, P., & Cullen, J. W. (1984). The scientific approach to cancer control. *CA: A Cancer Journal for Clinicians, 34*, 328–332.
- Hall, G. C. N. (2001). Psychotherapy research with ethnic minorities: Empirical, ethical, and conceptual issues. *Journal of Consulting and Clinical Psychology, 69*, 502–510.
- Hammen, C., Rudolph, K., Weisz, J. R., Burge, D., & Rao, U. (1999). The context of depression in clinic-referred youth: Neglected areas in treatment. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 64–71.
- Havik, O. E., & VandenBos, G. R. (1996). Limitations of manualized psychotherapy for everyday clinical practice. *Clinical Psychology: Science and Practice, 3*, 264–267.
- Ho, A. C., Weisz, J. R., Austin, A. A., Chorpita, B. F., Wells, K. C., Southam-Gerow, M. A., & The Research Network on Youth Mental Health (2007). Bridging science and community practice: Clinician and organizational engagement in community clinics in the clinic treatment project. *Report on Emotional and Behavioral Disorders in Youth, 7*, 13–16.
- Jones, M. C. (1924). A laboratory study of fear: The case of Peter. *Pedagogical Seminary, 31*, 308–315.
- Kazdin, A. E. (2002). The state of child and adolescent psychotherapy research. *Child and Adolescent Mental Health, 7*, 53–59.
- Kazdin, A. E., Bass, D., Ayers, W. A., & Rodgers, A. (1990). Empirical and clinical focus of child and adolescent psychotherapy research. *Journal of Consulting & Clinical Psychology, 58*, 729–740.
- Levitt, E. E. (1957). The results of psychotherapy with children: An evaluation. *Journal of Consulting Psychology, 21*, 189–196.
- Levitt, E. E. (1963). Psychotherapy with children: A further evaluation. *Behaviour Research and Therapy, 60*, 326–329.
- Martin, J. L., Weisz, J. R., Chorpita, B. F., Higa, C. K., Wells, K. C., Southam-Gerow, M. A., & The Research Network on Youth Mental Health (2007). Moving evidence-based practices into everyday clinical care settings: Addressing challenges associated with pathways to treatment, child characteristics, and structure of treatment. *Report on Emotional and Behavioral Disorders in Youth, 7*, 5–8.
- McLeod, B. D. (2001). *The therapy process observational coding system for child psychotherapy*. Los Angeles: University of California.
- McLeod, B. D., & Weisz, J. R. (2004). Using dissertations to examine potential bias in child and adolescent clinical trials. *Journal of Consulting and Clinical Psychology, 72*, 235–251.
- MTA Cooperative Group (1999). A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. *Archives of General Psychiatry, 56*, 1073–1086.
- Meichenbaum, D., & Goodman, S. (1971). Training impulsive children to talk to themselves: A means of developing self-control. *Journal of Abnormal Psychology, 77*, 115–126.
- National Institutes of Health (1994). Behavioral therapies development program. *NIH Guide, 22*, No. 26.
- Palinkas, L. A., Hogwood, K., Schoenwald, S., Weisz, J., Chorpita, B., & the Child STEPS Research Network (in press). Implementation of evidence-based practice in child mental health: First steps. *Administration and Policy in Mental Health and Mental Health Services Research*.
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century.
- Skinner, B. F. (1974). *About behaviorism*. New York: Knopf.
- Southam-Gerow, M. A., Weisz, J. R., & Kendall, P. C. (2003). Youth with anxiety disorders in research and service clinics: Examining client differences and similarities. *Journal of Clinical Child and Adolescent Psychology, 32*, 375–385.
- Stallard, P., Udwin, O., Goddard, M., & Hibbert, S. (2007). The availability of cognitive behaviour therapy within specialist child and adolescent mental health services (CAMHS): A national survey. *Behaviour and Cognitive Psychotherapy, 35*, 501–506.
- Strupp, H. H., & Anderson, T. (1997). On the limitations of therapy manuals. *Clinical Psychology: Science and Practice, 4*, 76–82.
- Sue, S. (2003). In defense of cultural competency in psychotherapy and treatment. *American Psychologist, 58*, 964–970.
- Weisz, J. R. (2004). *Psychotherapy for children and adolescents: Evidence-based treatments and case examples*. New York: Cambridge University Press.
- Weisz, J. R., & Addis, M. E. (2006). The research-practice tango and other choreographic challenges: Using and testing evidence-based psychotherapies in clinical care settings. In C. D. Goodheart, A. E. Kazdin & R. J. Sternberg (Eds.), *Evidence-based psychotherapy: Where practice and research meet* (pp. 179–206). Washington, DC: American Psychological Association.
- Weisz, J. R., Doss, A. J., & Hawley, K. M. (2005). Youth psychotherapy outcome research: A review and critique of the evidence base. *Annual Review of Psychology, 56*, 337–363.
- Weisz, J. R., Hawley, K. M., & Doss, A. J. (2004). Empirically test psychotherapies youth internalizing and externalizing problems and disorders. *Child and Adolescent Psychiatric Clinics of North America, 13*, 729–815.
- Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus usual clinical care: A meta-analysis of direct comparisons. *American Psychologist, 61*, 671–689.
- Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: A meta-analysis. *Psychological Bulletin, 132*, 132–149.
- Weisz, J. R., Sandler, I. N., Durlak, J. A., & Anton, B. S. (2005). Promoting and protecting youth mental health through evidence-based prevention and treatment. *American Psychologist, 60*, 628–648.

- Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology, 55*, 542–549.
- Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A., & Morton, T. (1995). Effects of Psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin, 117*, 450–468.
- Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004a). The empirical status of empirically supported therapies: Assumptions, findings, and reporting in controlled trials. *Psychological Bulletin, 130*, 631–663.
- Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004b). The next generation of psychotherapy research: Reply to Ablon and Marci (2004), Goldfried and Eubanks-Carter (2004), and Haaga (2004). *Psychological Bulletin, 130*, 677–683.
- Woody, S. R., Weisz, J. R., & McLean, C. (2005). Empirically supported treatments: Ten years later. *The Clinical Psychologist, 58*, 5–11.

Copyright of *Child & Adolescent Mental Health* is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.