Shrinking the Gap Between Research and Practice: Tailoring and Testing Youth Psychotherapies in Clinical Care Contexts

John R. Weisz, Lauren S. Krumholz, Lauren Santucci, Kristel Thomassin, and Mei Yi Ng

Department of Psychology, Harvard University, Cambridge, Massachusetts 02138; email: john_weisz@harvard.edu

Keywords
youth psychotherapy, treatment research, children, adolescents

Abstract
Most youth psychotherapy research involves conditions quite unlike the clinical practice it is designed to strengthen. Most studies have not tested interventions with clinically referred youths and practicing clinicians in clinical care settings, nor have they tested whether new treatments produce better outcomes than usual practice. Limited exposure to real-world conditions and questions may partially explain why empirically supported treatments show such modest effects when tested under more representative conditions, against usual care. Our deployment-focused model calls for intervention development and testing with the kinds of participants (e.g., clients and clinicians) and in the contexts (e.g., clinics) for which the interventions are ultimately intended, and for randomized comparisons to usual clinical care. Research with the Child STEPs (system and treatment enhancement projects) treatment approach illustrates the methods and potential benefits of the deployment-focused model. Findings supporting Child STEPs are but one part of a rich research matrix needed to shrink the gap between intervention research and clinical practice.
INTRODUCTION

Evan, age 13, is referred for treatment in a clinic that specializes in youth depression. Evan meets the criteria for major depressive disorder, and he has no other psychiatric disorder and no other significant problems that could interfere with depression treatment. Evan’s therapist, an expert who treats only depression, uses a cognitive behavioral manual that includes 17 structured sessions, and she devotes 45–60 minutes before each session to planning and preparing the session, as the manual suggests. The sessions begin with psychoeducation, then focus on a series of depression-coping skills in a prescribed order (e.g., positive activity scheduling, mood management through relaxation, problem solving, building social skills, identifying and changing unrealistic depressogenic cognitions), then conclude with a planned termination that includes summarizing the skills learned and discussing when those skills will be needed in the future. The treatment goes smoothly. Evan attends every session and learns the skills in the prescribed order, with no unexpected problems arising that might change the focus on depression and no unexpected crises that might need to be addressed outside the manual. At termination, Evan no longer meets criteria for any depressive disorder, and he is functioning well at home and in school.

Clinical practitioners who work with children and teens will correctly conclude that Evan’s case report is fiction. It would be very difficult to find clinics or clinicians who have the luxury of treating only youth depression, or clinics in which clinicians are allowed an hour of prep time before each session; most clinics and clinicians deal with a broad range of psychopathologies and confront significant financial and time pressures. It would also be very difficult to find many
referred youngsters who have only one psychiatric disorder and no other significant problems; among referred youths, comorbidity is the rule, in addition to multiple significant problems that can affect the course or outcome of treatment though they may not meet diagnostic criteria. As a consequence, when the structured, manual-guided treatments that so much of our clinical research has produced are used in everyday clinical practice, they may encounter challenges that disrupt the prescribed sequence of sessions, crises that make the planned focus on one disorder difficult to sustain, and life events that wipe out plans for a predictable course of treatment and an orderly termination. Stated briefly, everyday treatment is often not as linear as most treatment manuals are. This article discusses the gap between actual clinical practice and the products of scientific research designed to improve practice. We focus on the nature of that gap, how it came to exist, and some steps that can be taken to bridge it, with the goal of bringing the science and practice of youth mental health care closer together.

HISTORICAL ORIGINS, EARLY EVALUATIONS, AND EVOLUTION OF YOUTH PSYCHOTHERAPIES

Psychotherapy for children and adolescents has deep roots, extending through thousands of years of history and tradition within two broad streams. One stream entails parenting guidance, epitomized by biblical injunctions (e.g., “The rod and reproof give wisdom, but a child left to himself brings shame to his mother”; Proverbs 29:15). The other stream involves guidance on how to help people change, as epitomized by the classical Greek philosophers who used human discourse to provoke new ways of thinking and behaving. For example, Socrates’ (469–399 BCE) philosophical dialectic, later labeled the Socratic method, involved questioning others to prompt them to reexamine their beliefs and the actions stimulated by those beliefs. A confluence of guidance on parenting and strategies for human change—together with many other influences—contributed, over time, to what is now known as youth psychotherapy. The notion of psychotherapy as a profession can be traced back about a century (Freedheim et al. 1992), arguably to the work of Sigmund Freud (1856–1939), including his consultation with the father of Little Hans and the psychoanalysis of his own daughter, Anna (1895–1982), who became a prominent child analyst in her own right. The psychoanalytic model eventually competed with treatment models derived from the grand theories of personality, from humanistic alternatives, and from early forms of behaviorism.

As multiple forms of therapy sprang up, optimism about their supposed benefits spread. So it was a shock to many when Eysenck (1952) published a review of research suggesting that psychotherapy might not actually work, and Levitt (1957, 1963) reached the same conclusion about youth therapy in particular. Levitt’s reviews concluded that rates of improvement in troubled youths were about the same with or without treatment. These mid-century reviews prompted some serious rethinking regarding what therapy should look like. In their wake, therapy models that had been rather vaguely characterized gave way over later decades to more structured approaches that were increasingly well documented in treatment manuals. The targets of therapy were also better defined, with the manuals identifying which kinds of problems the treatments were designed to address. In addition, research methods grew more rigorous, with randomized controlled trials (RCTs) becoming a gold standard. Finally, an increasingly detailed diagnostic system, with successive iterations of the Diagnostic and Statistical Manual of Mental Disorders, helped tighten the linkage between treatment manuals and specific disorders. This linkage was bolstered by funding agencies whose grant review priorities favored diagnosis-specific treatments. Now, after this long sequence of steps, we have an array of structured treatment manuals, most of which are designed for one target disorder or problem or a homogeneous cluster (e.g., depressive disorders or conduct-related disorders or problems).
Figure 1

Mean effect size values in meta-analyses of adult psychotherapy randomized controlled trials (RCTs) (left two bars), youth psychotherapy RCTs (middle four bars), and (c) youth RCTs pitting empirically supported treatments (ESTs) against usual care (right two bars), analyzed by Weisz et al. (2006) (32 studies) and Weisz et al. (2013a) (52 studies). The two meta-analyses comparing ESTs to usual care revealed a probability of only 0.58 (versus chance at 0.50) that a randomly selected youth from the EST condition would show a better outcome after treatment than a randomly selected youth from the usual care condition. The full bar for Kazdin et al. (1990) shows the mean effect size for treatment versus inert control group comparisons; the dashed line shows the mean effect size for treatment versus active control group comparisons. The full bar for Weisz et al. (1995) shows the mean effect size when unweighted least squares analyses were used; the dashed line shows the mean when weighted least squares analyses were used. Figure adapted with permission from John R. Weisz.

META-ANALYTIC FINDINGS ON THE BENEFITS OF YOUTH PSYCHOTHERAPIES

The level of benefit derived from research-tested treatments can be summarized through research syntheses called meta-analyses, which involve pooling the findings of multiple RCTs. Among these meta-analyses, four have synthesized findings from especially broad arrays of youth treatments and forms of dysfunction. In the first of these, Casey & Berman (1985) focused on studies with children aged 12 and younger. In a subsequent meta-analysis, Weisz et al. (1987) included studies with 4- to 18-year-olds. Kazdin et al. (1990) also pooled findings of studies with 4- to 18-year-olds, and Weisz et al. (1995) included studies spanning ages 2–18. Mean effect sizes found in these four meta-analyses are shown in Figure 1; a comparison to two widely cited meta-analyses of predominantly adult psychotherapy is shown to the left (Shapiro & Shapiro 1982, Smith & Glass 1977). As the figure shows, the youth psychotherapy effect sizes fall roughly within the range of what has been found for adult therapy, and on average within the range of what Cohen (1988) suggests as benchmarks for medium (i.e., 0.5) to large (i.e., 0.8) effects.

IDENTIFYING EMPIRICALLY SUPPORTED TREATMENTS

As accumulating research evidence pointed to the beneficial effects of well-documented (typically manual-guided) psychotherapies for youths, adults, couples, and families, efforts were launched to
identify empirically supported treatments (ESTs). Task forces and review teams were formed— notably the American Psychological Association’s Division 12 Task Force on Promotion and Dissemination of Psychological Procedures (e.g., Chambless et al. 1998)—to distill the evidence from outcome studies and identify therapies that reached threshold for different levels of empirical support. The criteria for EST status at various levels vary somewhat across review groups, but most of them require empirical support in more than one study (e.g., with the target treatment showing outcomes superior to those of a viable control or comparison group) and that the studies meet certain standards of methodological rigor. Building on the work of the Division 12 Task Force, experts in youth psychotherapy carried out systematic reviews, sometimes including meta-analytic findings, to compile reports on ESTs for children and adolescents (see Lonigan et al. 1998, Silverman & Hinshaw 2008b). In the most recent report, edited by Silverman & Hinshaw (2008a), reviewers identified psychotherapies that met criteria for four levels: well established (e.g., two good group-design experiments by different research teams in two different settings, showing the treatment to be “superior to pill or psychological placebo or to another treatment”) (Silverman & Hinshaw 2008b, p. 5), probably efficacious (e.g., “at least two good experiments showing the treatment is superior . . . to a wait-list control group”) (Silverman & Hinshaw 2008b, p. 5), possibly efficacious (e.g., “At least one ‘good’ study showing the treatment to be efficacious in the absence of conflicting evidence”) (Silverman & Hinshaw 2008b, p. 5), or experimental (e.g., “not yet tested in trials meeting task force criteria . . .”) (Silverman & Hinshaw 2008b, p. 5). Table 1 shows the kinds of treatments classified at the two highest levels—i.e., well established and probably efficacious—in the Silverman & Hinshaw (2008a) special issue. A different system, the National Registry of Evidence-Based Programs and Practices, sponsored by the US Substance Abuse and Mental Health Services Administration, as of January 17, 2015, listed 211 interventions for youths aged 17 and under that met its criteria (see http://www.nrepp.samhsa.gov/ViewAll.aspx).

YOUTH PSYCHOTHERAPY RESEARCH AND THE NATURE OF CLINICAL CARE

The work of treatment developers over the past five decades has been remarkable, the products of that work have dramatically altered the face of youth psychotherapy, and the identification of ESTs has been a major advance. However, concerns have been raised about how well these therapies fit into the context of the everyday clinical care that they are designed to improve. Some critics have argued that these structured, manual-guided treatments have significant limitations that undermine their usefulness in usual clinical practice. They have suggested that ESTs (a) have often been developed and tested with relatively simple, sometimes subclinical cases and thus may not work well with the complex and severe cases seen in usual clinical care; (b) have been designed for single problems or diagnoses and thus may not work as well with the comorbid cases so often seen in usual clinical care; (c) are so prescriptive and linear in design that they make it hard to individualize treatment to meet distinctive client needs; (d) have so many procedural proscriptions that they constrain therapist creativity in addressing unusual or unexpected events in clients’ lives; and (e) inhibit the spontaneity and flexibility needed to build rapport and develop a good therapeutic relationship. Several of these concerns reflect the view that ESTs may not be well suited to the challenge of treating clinically referred individuals in the context of usual clinical care (see examples of such concerns discussed in Addis & Krasnow 2000; Addis & Waltz 2002; Garfield 1996; Havik & Vandenbos 1996; Strupp & Anderson 1997; Weisz & Addis 2006; Weisz & Gray 2008; Weisz et al. 2011b; Westen et al. 2004a,b).

Our review of the scientific literature on youth treatment outcomes and our testing of standard manualized youth treatments in everyday clinical practice contexts have highlighted some of these
Table 1  Youth psychotherapies identified as well-established or probably efficacious<sup>a,b</sup>

<table>
<thead>
<tr>
<th>Problem/disorder category</th>
<th>Well-established therapies</th>
<th>Probably efficacious therapies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early autism (Rogers &amp; Vismara 2008)</td>
<td>Lovaas model: intensive behavioral intervention</td>
<td>Pivotal response treatment</td>
</tr>
<tr>
<td>Eating disorders in adolescence (Keel &amp; Haedt 2008)</td>
<td>Family therapy for anorexia nervosa</td>
<td>None</td>
</tr>
<tr>
<td>Depression (David-Ferdon &amp; Kaslow 2008)</td>
<td>Cognitive-behavioral therapy (CBT) for children</td>
<td>Behavior therapy for children</td>
</tr>
<tr>
<td></td>
<td>CBT for adolescents</td>
<td></td>
</tr>
<tr>
<td>Interpersonal psychotherapy for adolescents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phobic and anxiety disorders (Silverman et al. 2008b)</td>
<td>None</td>
<td>Group CBT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Group CBT for social phobia</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Individual CBT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Social effectiveness training for social phobia</td>
</tr>
<tr>
<td>Obsessive-compulsive disorder (Barrett et al. 2008)</td>
<td>None</td>
<td>Individual exposure-based CBT</td>
</tr>
<tr>
<td>Youths exposed to traumatic events (Silverman et al. 2008a)</td>
<td>Trauma-focused CBT</td>
<td>School-based group CBT</td>
</tr>
<tr>
<td>Attention-deficit/hyperactivity disorder (Pelham &amp; Fabiano 2008)</td>
<td>Behavioral classroom management</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Behavioral parent training</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intensive peer-focused behavioral interventions in recreational settings</td>
<td></td>
</tr>
<tr>
<td>Disruptive behavior (Eyberg et al. 2008)</td>
<td>Parent management training—the Oregon model</td>
<td>Anger control training</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Group assertive training</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Helping the noncompliant child</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Incredible years child training</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Incredible years parent training</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multidimensional treatment foster care</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multisystemic therapy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Parent–child interaction therapy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Positive parenting program—enhanced</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Positive parenting program—standard</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Problem-solving skills training (PSST)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PSST + parent management training</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PSST + practice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rational-emotive mental health program</td>
</tr>
<tr>
<td>Adolescent substance abuse (Waldron &amp; Turner 2008)</td>
<td>Functional family therapy</td>
<td>Behavioral family therapy</td>
</tr>
<tr>
<td></td>
<td>Group CBT</td>
<td>Brief strategic family therapy</td>
</tr>
<tr>
<td></td>
<td>Individual CBT</td>
<td>Multisystemic therapy</td>
</tr>
<tr>
<td></td>
<td>Multidimensional family therapy</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Table based on reviews in Silverman & Hinshaw (2008a).

<sup>b</sup>This table shows classifications for broad forms of psychotherapy (e.g., CBT); some reports in the special issue of the journal (Silverman & Hinshaw 2008a) also classified specific treatment subtypes (e.g., group CBT for children, individual adolescent CBT plus parent/family component), which are not included in the table due to space limitations.
concerns. In our review of youth RCTs (Weisz et al. 2005b), we found that most studies took place in settings created for research (e.g., university labs and lab clinics), with treatment provided to youths who were recruited (e.g., through ads) and delivered not by practicing clinicians but rather by graduate students or other individuals dependent on the researcher for their employment, income, or evaluation. Recently, we examined 461 youth psychotherapy RCTs from the 1960s through the most recent decade, including 1,160 treatment and control groups (Weisz et al. 2014). We found that the great majority of the 1,160 treatment and control groups involved youths who had not been referred for clinical care, therapists who were not practitioners, and therapy that was not delivered in clinical practice settings (see Weisz et al. 2014). Across the five decades, only 2.1% of all study groups were described by authors as involving clinically referred clients treated by practitioners in practice settings; even for studies in the most recent decade, the figure was only 4.5% (see Table 2). Conducting research outside the clinical practice context has very clear advantages experimentally. Structuring treatment trials in ways that provide control over the nature of the study sample, the behavior and adherence of the therapists, and conditions in the treatment setting reduces experimental noise and—in principle—adds precision to the test of the intervention. Precision and control are often amplified by using participants who are good treatment candidates and who have the target disorder without potentially interfering comorbidities; therapists who are screened, tested with practice cases, and paid by the researcher; and settings (e.g., university lab, lab clinic, or designated rooms in a school) over which the researcher has significant design authority and day-to-day control. Although such provisions may well enhance experimental control, there may also be a cost when the topic under study is clinical intervention. In these cases, controlling complicating factors that are so often present in everyday clinical care may reduce opportunities to learn how to make treatments effective in the crucible of real life, with all the noisy variables that constitute what we have called “the mental health ecosystem” (Weisz et al. 2013b, p. 274). That ecosystem includes child characteristics (e.g., comorbid disorders and co-occurring problems, which are characteristic of most referred youths), family characteristics (e.g., parental mental illness, stress, and time constraints, which can lead to missed appointments and treatment dropout), practitioner factors (e.g., full caseloads, which can compete with the session preparation needed for manualized treatments), clinic factors (e.g., limited resources for training and supervision or productivity requirements generated by insurance rules that eliminate most nonreimbursable activities), and a broad variety of other real-life factors (e.g., parental job loss, dangerous neighborhoods, intervention of child protective services). Such ecosystem factors, detailed further in Table 3, might well prevent youths from being enrolled in an RCT populated through ads, but any of these factors could well be in play within everyday clinical practice.

PUTTING EMPIRICALLY SUPPORTED TREATMENTS INTO CLINICAL CARE CONTEXTS

The emphasis on experimental control may therefore have been a double-edged sword. On the one hand, that emphasis may support scientific rigor, and it is quite understandable for all of us who are trained to do well-controlled studies that minimize noise; moreover, the practice has been incentivized by the guidelines and practices of some of the funding agencies without whose support few treatment studies would be possible. On the other hand, these research practices may have reduced opportunities for us as researchers to learn what we need to know to navigate the real-life conditions that arise in actual clinical practice but are so often minimized in clinical research. It is an interesting paradox that funding designed to improve clinical care through research may have produced treatments that do not fit the very clinical care they were designed to improve. One risk is that manual-guided psychotherapies that look relatively strong in traditional treatment outcome...
Table 2  Percentage of treatment/control groups in youth psychotherapy outcome studies that employed clinically representative youths, therapists, and treatment settings

<table>
<thead>
<tr>
<th>Decade</th>
<th>1960s</th>
<th>1970s</th>
<th>1980s</th>
<th>1990s</th>
<th>2000s</th>
<th>All decades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of studies</td>
<td>13</td>
<td>62</td>
<td>99</td>
<td>100</td>
<td>187</td>
<td>461</td>
</tr>
<tr>
<td>Number of groups</td>
<td>35</td>
<td>183</td>
<td>273</td>
<td>244</td>
<td>425</td>
<td>1,160</td>
</tr>
</tbody>
</table>

How youths were enrolled in the study

<table>
<thead>
<tr>
<th>Method</th>
<th>1960s</th>
<th>1970s</th>
<th>1980s</th>
<th>1990s</th>
<th>2000s</th>
<th>All decades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recruited, nontreatment-seeking</td>
<td>62.9</td>
<td>85.8</td>
<td>65.9</td>
<td>57.8</td>
<td>62.8</td>
<td>66.1</td>
</tr>
<tr>
<td>Clinic-referred, treatment-seeking</td>
<td>5.7</td>
<td>4.9</td>
<td>24.2</td>
<td>26.2</td>
<td>24.5</td>
<td>21.1</td>
</tr>
<tr>
<td>Required via court or justice system</td>
<td>17.1</td>
<td>8.2</td>
<td>8.8</td>
<td>11.5</td>
<td>7.8</td>
<td>9.1</td>
</tr>
<tr>
<td>Enrollment method not reported</td>
<td>14.3</td>
<td>1.1</td>
<td>1.1</td>
<td>4.5</td>
<td>4.9</td>
<td>3.6</td>
</tr>
</tbody>
</table>

Who provided the treatment

<table>
<thead>
<tr>
<th>Percentage</th>
<th>1960s</th>
<th>1970s</th>
<th>1980s</th>
<th>1990s</th>
<th>2000s</th>
<th>All decades</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 50% therapists are practitioners</td>
<td>65.7</td>
<td>42.1</td>
<td>55.3</td>
<td>40.2</td>
<td>35.5</td>
<td>43.1</td>
</tr>
<tr>
<td>&gt; 50% therapists are practitioners</td>
<td>2.9</td>
<td>9.8</td>
<td>7.7</td>
<td>10.7</td>
<td>19.1</td>
<td>12.7</td>
</tr>
<tr>
<td>Therapist vocation not reported</td>
<td>14.3</td>
<td>27.9</td>
<td>23.4</td>
<td>32.0</td>
<td>25.4</td>
<td>26.4</td>
</tr>
<tr>
<td>No treatment or waitlist</td>
<td>17.1</td>
<td>20.2</td>
<td>13.6</td>
<td>17.2</td>
<td>20.0</td>
<td>17.8</td>
</tr>
</tbody>
</table>

Where the treatment took place

<table>
<thead>
<tr>
<th>Setting</th>
<th>1960s</th>
<th>1970s</th>
<th>1980s</th>
<th>1990s</th>
<th>2000s</th>
<th>All decades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research setting</td>
<td>11.4</td>
<td>6.6</td>
<td>9.9</td>
<td>18.4</td>
<td>19.5</td>
<td>14.7</td>
</tr>
<tr>
<td>Custodial, school, or supervised setting</td>
<td>8.6</td>
<td>29.5</td>
<td>32.6</td>
<td>20.5</td>
<td>22.6</td>
<td>25.2</td>
</tr>
<tr>
<td>Clinical service setting</td>
<td>0</td>
<td>2.7</td>
<td>5.9</td>
<td>11.5</td>
<td>14.8</td>
<td>9.7</td>
</tr>
<tr>
<td>Correctional setting</td>
<td>14.3</td>
<td>1.1</td>
<td>5.1</td>
<td>4.1</td>
<td>2.1</td>
<td>3.4</td>
</tr>
<tr>
<td>Treatment setting not reported</td>
<td>48.6</td>
<td>39.9</td>
<td>33.0</td>
<td>28.3</td>
<td>20.9</td>
<td>29.1</td>
</tr>
<tr>
<td>Waitlist or no treatment</td>
<td>17.1</td>
<td>20.2</td>
<td>13.6</td>
<td>17.2</td>
<td>20.0</td>
<td>17.8</td>
</tr>
</tbody>
</table>

Sum of representativeness factors

<table>
<thead>
<tr>
<th>Factors reported</th>
<th>1960s</th>
<th>1970s</th>
<th>1980s</th>
<th>1990s</th>
<th>2000s</th>
<th>All decades</th>
</tr>
</thead>
<tbody>
<tr>
<td>No factors reported</td>
<td>91.4</td>
<td>83.6</td>
<td>70.0</td>
<td>63.9</td>
<td>61.2</td>
<td>68.3</td>
</tr>
<tr>
<td>One factor reported</td>
<td>8.6</td>
<td>15.3</td>
<td>22.7</td>
<td>25.4</td>
<td>23.8</td>
<td>22.1</td>
</tr>
<tr>
<td>Two factors reported</td>
<td>0</td>
<td>1.1</td>
<td>7.0</td>
<td>9.0</td>
<td>10.6</td>
<td>7.6</td>
</tr>
<tr>
<td>Three factors reported</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
<td>1.6</td>
<td>4.5</td>
<td>2.1</td>
</tr>
</tbody>
</table>

*Because treatment provider and setting are group-level variables, percentages of treatment/control groups rather than percentages of studies are reported for all three variables. Three studies that employed a combination of nontreatment-seeking and treatment-seeking youths, and one study with a treatment condition that employed a combination of clinical and research settings, were excluded from the analysis. Table reproduced with permission from John R. Weisz.

Studies may not look as strong when tested under more representative clinical practice conditions. That is, when ESTs are introduced into everyday practice contexts, they may be hampered by the kinds of ecosystem factors represented in Figure 2 to which their research was not exposed. The treatment outcomes they generate may suffer as a consequence. Because limited research exposure to such ecosystem factors is characteristic of many tested treatments (as shown in Table 2), some of the treatments that enjoy considerable research support may not look so strong when tested in more clinically representative contexts.

We have tested this notion in two different ways: through randomized effectiveness trials and through meta-analyses pitting ESTs head-to-head against usual care. The effectiveness trials took place in Los Angeles County. One trial tested a cognitive behavioral therapy (CBT) program for youth depression (Weisz et al. 2009); the other tested a CBT program for youth anxiety disorders (Southam-Gerow et al. 2010). Both effectiveness trials were conducted at multiple community mental health clinics. Both trials focused on children and adolescents who had been referred to the clinics through normal community channels; there were no ads and no recruitment. In both studies,
Table 3  Components and characteristics of the youth public mental health ecosystem that can affect the use of empirically supported psychotherapies

<table>
<thead>
<tr>
<th>Participants</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinically referred youths</td>
<td>Comorbidity and co-occurring problems; high rates of externalizing problems; frequent crises and shifts in most pressing needs during treatment</td>
</tr>
<tr>
<td>Families and caregivers</td>
<td>Relatively low income; high stress; caregiver and sibling psychopathology; complex family systems and single parenthood; ethnocultural diversity; seeking help for youth problems of daily functioning, not diagnoses</td>
</tr>
<tr>
<td>Practitioners</td>
<td>Differing theoretical orientations and educational backgrounds with limited exposure to empirically supported treatments (ESTs); large caseloads; diverse caseload with broad array of problems; minimal to no time for treatment preparation, supervision, and additional training; fee for service or salaried with high productivity requirements</td>
</tr>
<tr>
<td>Provider organizations</td>
<td>Extreme financial pressures resulting in staff layoffs; shrinkage in the percent of salaried employees and increases in the percent of fee-for-service employees; escalating productivity requirements; significant staff turnover; minimal incentives and potential financial risk for investment in EST trainings</td>
</tr>
<tr>
<td>Networks of youth service systems</td>
<td>Rules, regulations, and procedures of the systems make it difficult for them to work together; systems may work against each other based on tradition and policies; difficult to implement ESTs across various systems</td>
</tr>
<tr>
<td>Policy context</td>
<td>Reimbursement is based largely on categories of care provided and amount of time provided, not on the nature of the intervention or whether it is supported by scientific evidence; no real policy or fiscal incentives to using ESTs; changes in political leadership influence mental health care system</td>
</tr>
</tbody>
</table>


As a complement to our own effectiveness trials, we have drawn on a larger database to conduct two relevant meta-analyses. For both of these, we identified RCTs in which youth treatments classified as empirically supported (also known as evidence-based psychotherapies) were tested directly against usual clinical care. For the first of these meta-analyses (Weisz et al. 2006), our search identified 32 such studies; for the second meta-analysis (Weisz et al. 2013a) we found 52 studies. Not surprisingly, these studies involved more clinically representative conditions (i.e., more referred youths, clinical practitioners, real-world treatment settings) than the overall youth psychotherapy RCT database. In both meta-analyses, we found, as shown in Figure 1, that treatment benefit dropped markedly when the ESTs were tested in these more representative conditions and compared to usual clinical care. Figure 3 shows effect size values for all 52 studies in our 2013 meta-analysis. This figure shows that there was considerable variability in findings but that a substantial number of studies produced effects either well below zero (indicating superior outcomes...
When empirically supported treatments are introduced into everyday clinical practice, therapies whose research base has excluded the kinds of mental health ecosystem factors shown here may have difficulty addressing those factors, and their outcome may be adversely affected. Figure adapted with permission from John R. Weisz.

Overall, we found markedly lower mean effect sizes than in the main body of RCT research, as reflected in the four middle bars in Figure 1. In fact, the mean effect sizes of 0.30 (Weisz et al. 2006) and 0.29 (Weisz et al. 2013a) in these EST versus usual care meta-analyses reflect a probability of only 0.58 (versus chance at 0.50) that a randomly selected youth from the EST condition in these studies would be better off after treatment than a randomly selected youth from the usual care condition. Especially worrisome was the finding in our 2013 meta-analysis that empirically supported therapies did not significantly outperform usual care among studies with clinically referred youths or youths impaired enough to meet criteria for a formal diagnosis. This is bad news because these are two groups for whom treatment success should have especially high priority.

The challenge of fitting the tested youth treatments into everyday practice and making them work well in that context may help explain why everyday use of most of these treatments has not spread very fast despite the diffusion of training opportunities. Our research team has learned a good deal about practitioner response from our community clinic implementation efforts. For example, 3–5 years after completing the Los Angeles County effectiveness trials described earlier (Southam-Gerow et al. 2010, Weisz et al. 2009), our research team contacted the practitioners we
Figure 3

Effect sizes reported in 52 randomized trials comparing empirically supported treatments (ESTs) to usual care (in Weisz et al. 2013a). Bars above 0 indicate that EST was superior to usual care, bars at 0 indicate no difference in outcome, and bars below 0 indicate that usual care was superior to the EST. The horizontal line at 0.29 shows the mean effect size across the full study set. Note the number of studies for which usual care showed effects similar to or superior to EST. References cited in the figure are available in the Supplemental Material associated with this article, access it by following the Supplemental Materials link from the Annual Reviews home page at http://www.annualreviews.org. Figure adapted with permission from John R. Weisz.
had trained in CBT, asking them how often they still used the CBT treatments they had mastered during the trial (Chu et al. 2014). The practitioners reported using the full CBT protocol for anxiety (i.e., Coping Cat) with only 7.5% of their anxiety cases and the full CBT protocol for depression [i.e., Primary and Secondary Control Enhancement Training (PASCET)] with only 20% of their depression cases. Interestingly, the practitioners did report using selected components of both protocols (e.g., graduated exposure, problem solving, cognitive restructuring) quite often (78.5% for anxiety, 58.6% for depression), presaging an approach to treatment tailoring that we have come to appreciate and adopt (see the Modular Approach to Therapy for Children section below).

We have found that a variety of youth mental health ecosystem factors like those shown in Figure 2 and detailed in Table 3 can make it difficult for practitioners to use fully intact treatment protocols in their everyday practice. A number of the most important challenges relate to the fact that most tested youth treatments are designed for a single disorder or homogeneous cluster (e.g., depressive disorders or conduct problems). This is a problem for at least three reasons:

1. **Practitioner caseloads.** Most youth practitioners carry rather broad caseloads encompassing a diverse array of disorders and problems. Learning a new treatment for one disorder may not be relevant to most of the caseload.

2. **Youth comorbidity and co-occurring problems.** Most clinically referred youths have multiple disorders and problems, often including both internalizing and externalizing forms of dysfunction. An EST for one disorder or problem may not address the others.

3. **Flux in the problems that need attention in treatment.** Youths in everyday clinical care often present a moving target. There is flux in the problems that most need attention during a treatment episode, and of course new information acquired during treatment may change the picture of what the core problems actually are. A clinician who knows only an EST for depression may struggle when a young client begins to show serious conduct problems or intense fear of separation from parents.

Stated simply, most ESTs are more narrowly focused, and more linear in design, than the everyday clinical practice they are designed to enhance. This is quite understandable, given the value clinical scientists have learned to place on clarity, focus, and logical order. These values make it appealing to design protocols that plan treatment steps in a prescribed order—in CBT for depression, for example, an ideal sequence might begin with relationship building and psychoeducation and then proceed to problem-solving skills, to learning to schedule pleasant activities, to identifying and modifying negative cognitions, and ultimately end with a planned termination, always with a focus on depression. The problem is that such linear plans may not fit the reality of everyday treatment for many young people.

**RESTRUCTURING YOUTH THERAPIES (AND ASSESSMENTS) TO FIT CLINICAL CARE: THE CHILD STEP Approach**

What can be done to address the areas of mismatch between most ESTs and the nature of treatment in everyday practice? There may certainly be many useful strategies. The approach we describe here was developed as part of the work of the Research Network on Youth Mental Health (see Schoenwald et al. 2008), which was formed in part to address challenges that can arise in the implementation of ESTs in everyday clinical care. Our focus in the network was on community clinic and school-based clinic settings, the contexts where most youth mental health care is provided in the United States. However, there is plenty of room to broaden that focus in the future, extending efforts to adapt and tailor interventions to the many other contexts in which care is provided (e.g., medical settings, child welfare services, juvenile justice programs). The approach we developed
in the network is called Child STEPs (child system and treatment enhancement projects). It includes (a) a modular, transdiagnostic treatment protocol and (b) a clinical information system that provides ongoing monitoring of youth treatment response and frequent feedback to therapists to guide their decision making throughout episodes of care.

Modular Approach to Therapy for Children
The treatment manual that emerged from our network is essentially a menu of common components of ESTs for particularly common youth problems, concise descriptions of each component in the menu, and decision support in the form of flowcharts to help therapists individualize each treatment episode by selecting from the menu. In this respect, the approach is integrative or transdiagnostic. Our goal (see Chorpita & Weisz 2009, Weisz & Chorpita 2011) was to provide broader and more flexible coverage of youth problems and disorders than most standard treatment protocols do, while retaining the core components of the ESTs that have been developed and tested so carefully over the decades by so many distinguished treatment developers and clinical scientists. The current version of this treatment manual is called Modular Approach to Therapy for Children with Anxiety, Depression, Trauma, or Conduct Problems (MATCH; Chorpita & Weisz 2009).

The Structure of the Modular Approach to Therapy for Children
The MATCH manual includes 33 modules; brief summaries (3–4 pages each) of treatment components that are frequently included in CBT for depression (e.g., behavioral activation); CBT for anxiety, including posttraumatic stress (e.g., graduated exposure); and behavioral parent training for disruptive conduct (e.g., teaching caregivers to give clear instructions and use labeled praise). Examples of the modules are shown in Figure 4. Because MATCH includes so many modules and spans multiple disorders and target problems, considerable decision making is required of clinicians throughout treatment—for example, decisions about which modules to use when and about when the treatment focus should shift from one problem area to another. MATCH includes flowcharts to guide this decision making. Because a treatment episode may begin with any of four problem foci—anxiety, posttraumatic stress, depression, or conduct—there is a flowchart for each of these starting points. For each problem area, the flowchart begins with a default sequence of modules—that is, a suggested order the clinician may use if treatment is routine and does not need to address other problem areas. For example, the default sequence for depression begins with getting acquainted and building rapport, followed by psychoeducation on depression for the youth and caregiver, followed by skill building in problem solving, behavioral activation, relaxation and self-calming, social interaction (presenting a positive self), and identifying and restructuring unrealistic negative cognitions. The flowchart also includes branches or detours the clinician may take if there is interference—for example, if progress in using the depression modules is undermined by youth conduct problems or anxiety. When such interference arises, the flowcharts provide suggestions for the use of modules designed for other problem areas. For example, if a youngster learns depression-coping skills in therapy but resists using them outside therapy sessions (where they are most needed), and thus is not showing reduced depression, the therapist and parents might use the rewards module from the conduct section of MATCH to create a plan for incentivizing frequent use of those skills at home and at school (see the case of Sophia, below, for an example).

Treatment with the Modular Approach to Therapy for Children
Treatment with MATCH begins with an initial assessment designed to determine which of the problem areas is the most appropriate initial focus of treatment. In cases in which treatment can
focus in a linear fashion on a single problem area throughout the episode of care, the default sequences described previously may be sufficient. However, in the much more frequent cases in which treatment interference arises or a youth’s treatment needs shift, therapists can respond by navigating across problem areas and modules—like those shown in Figure 4—to adjust the focus of treatment as needed. By encompassing multiple broad domains of psychopathology, the modular treatment approach can address the concern that practitioners typically carry broad caseloads and that referred youths tend to have multiple co-occurring disorders and problems. The multiproblem focus of MATCH and the decision-support flowcharts also make it possible to address the flux in treatment needs and problems that youths so often show during episodes of care.

**Using Frequent Assessment to Monitor Treatment Response and Guide Therapists**

The decision making required in this process is informed by an assessment approach that provides weekly feedback to the clinician on the youth’s response to treatment. This assessment...
strategy is the second element of the Child STEPs approach to treatment. Navigating across modules—and sometimes between sections of the manual (e.g., the depression and the conduct sections)—requires ongoing feedback on the youth’s current functioning and response to treatment. This need for ongoing, frequent feedback has led us to develop two brief, psychometrically sound measures—one standardized, the other idiographic. The Brief Problem Checklist (BPC; Chorpita et al. 2010) is a 12-item measure used to obtain standardized weekly youth and caregiver reports on the severity of the youth’s internalizing and externalizing problems. The BPC was derived from the Child Behavior Checklist (CBCL) and Youth Self-Report (YSR), which are standardized 118-item parent- and youth-report problem checklists described in Achenbach & Rescorla (2001), using item response theory analyses. (The BPC has now been replaced by a somewhat longer measure, the Brief Problem Monitor, for which a license is required by the ASEBA Corporation; http://www.aseba.org.) The other measure, Top Problems Assessment (TPA; Weisz et al. 2011a), involves an idiographic consumer-driven assessment approach; the youth and the caregiver each identify, at pretreatment, the three most important problems for which help is needed in therapy and then rate the severity of these problems frequently throughout treatment. Weekly ratings on these two measures are displayed within a web-based system that provides quick access for therapists and clinical supervisors and can thus be used to make intervention plans and adjust these plans throughout treatment. This monitoring and feedback system provides the kind of frequent updates on the youth’s response to treatment that can guide decisions as to whether shifts in treatment focus are needed and which modules may be most appropriate for the next steps of treatment. Emerging evidence indicates that such monitoring and feedback systems may enhance the effectiveness of psychotherapy even when ESTs are not being used (Bickman et al. 2011, Shimokawa et al. 2010).

Child STEPs Case Example: Sophia

The Child STEPs model pairs this ongoing monitoring and feedback with the modular MATCH protocol to support therapists’ efforts to personalize treatment for each of their young clients. A simple case example, depicted in Figure 5, illustrates a Child STEPs treatment episode in which MATCH modules from different problem areas were used to address comorbidity and interference, and monitoring and feedback were used to document and guide the treatment process. Thirteen-year-old “Sophia” (a pseudonym) and her father sought treatment to deal with Sophia’s depression and defiant behavior at home; these problems spiked after her parents’ separation and divorce, following her mother’s ongoing drug use and physical abuse of Sophia. A multi-informant assessment indicated that Sophia’s depression was the most prominent area of concern, so the therapist used a standard sequence of modules for depression in MATCH (e.g., systematic problem solving, engaging in mood-boosting activities, and cognitive restructuring). However, the safety planning module was also needed, early on, given the risks posed by Sophia’s mother. Sophia learned most of the depression-coping skills, but one aspect of her defiant behavior was that she resisted using them at home. Therefore, the therapist worked with Sophia and her father—using the rewards module from the conduct problems section of MATCH—to develop a reward system that would incentivize daily home use of the coping skills. For example, when Sophia was down in the dumps about a problem at school or with peers, her father reminded her that she could earn extra video game time if she would work through her problem-solving steps; when Sophia made unrealistic, gloomy comments, her father reminded her that she could earn points for special privileges if she could come up with realistic positive counterthoughts. The bottom half of Figure 5 shows the modules used, week by week; the top half shows Sophia’s response to treatment. As the figure shows, real reductions in Sophia’s internalizing problem
Case example: Sophia’s course of treatment using the modular approach to therapy for children (MATCH), displayed using one version of a monitoring and feedback system. In this version, the upper panel displays changes in Sophia’s problem ratings (here limited to parent and child internalizing problem scores) across 18 weeks of treatment. The lower panel shows which MATCH modules were implemented in each treatment session. BLUE is an acronym representing four different patterns of depressogenic thinking: B, blaming myself; L, looking for bad news; U, unhappy guessing; and E, expecting bad things to happen. Figure reprinted with permission from John R. Weisz.

levels were evident only after the rewards module had been implemented and Sophia had begun to use her depression-coping skills regularly. This case, and the diagram in Figure 5, illustrates the value of regular feedback on youth treatment response. First, Sophia’s persistent high ratings on internalizing problems, despite receiving all the depression modules, showed that some other module was needed. Second, the rapid drop in internalizing scores as soon as the rewards module was implemented showed that this particular module had been a wise choice. Third, the subsequent sustained low levels of internalizing problems, as reported by both Sophia and her father, indicated that treatment gains were in place and that termination planning could begin.

**TESTING CHILD STEPs IN A RANDOMIZED EFFECTIVENESS TRIAL**

We have tested the Child STEPs approach in one published randomized effectiveness trial (Weisz et al. 2012) and a two-year long-term outcomes report on that trial (Chorpita et al. 2013). In the original trial, community practitioners from ten different outpatient service settings were randomly assigned to three conditions: standard manual treatment (therapists used separate, preexisting, linearly designed manuals for CBT for depression, CBT for anxiety, and behavioral parent training for conduct problems), MATCH (the modular manual encompassing common components of
CBT for depression, CBT for anxiety, and behavioral parent training for conduct problems), and usual care. Outcomes were assessed during treatment using the weekly BPC and TPA measures and through standardized diagnostic assessments at pre- and posttreatment.

Our analyses showed that MATCH produced significantly steeper trajectories of improvement than usual care and standard treatment on both the BPC and TPA measures and that youths treated with MATCH also had significantly fewer diagnoses at posttreatment than youths treated with usual care. In contrast, outcomes of standard manual treatment did not differ significantly from usual care in the Weisz et al. (2012) report. The long-term follow-up report (Chorpita et al. 2013) focused on outcomes via the full 118-item CBCL and YSR measures, which were administered at quarterly intervals over a two-year period beginning at pretreatment. The long-term findings showed continued evidence that MATCH outperformed usual care and that standard manual treatment did not (although MATCH did not significantly outperform standard manual treatment in the long-term analyses). Findings from both pre- to posttreatment assessments (Weisz et al. 2012) and the two-year follow-up (Chorpita et al. 2013) suggest that a modular, transdiagnostic redesign—one that integrates core elements of empirically supported treatment for multiple forms of dysfunction—may be an effective approach to adapting ESTs for everyday clinical care. Other randomized effectiveness trials of MATCH, in the years ahead, should help clarify the boundary conditions within which MATCH may be effective and useful.

DESIGNING AND TESTING TREATMENTS WITH THEIR DESTINATION IN MIND: THE DEPLOYMENT-FOCUSED MODEL

Our experience conducting meta-analyses has taught us that one treatment trial is rarely definitive. Further tests will be needed for a full and fair evaluation of MATCH, as for any treatment. Moreover, MATCH, and the modular approach in general, is only one of many possible strategies for bridging the gap between research and practice. Newer and quite likely better approaches will emerge over time; indeed, some current treatment approaches seem to bridge the research–practice gap effectively for certain groups in the child welfare and juvenile justice populations (Henggeler & Schaeffer 2010, Smith & Chamberlain 2010). Treatment methods aside, we do believe that the effectiveness research strategy used in the initial test of MATCH warrants attention and continued use. That strategy includes testing the candidate treatment (a) with individuals who have actually been clinically referred (i.e., not recruited), (b) with the treatment being delivered by clinical practitioners (i.e., not research employees), (c) in clinical care settings (i.e., not settings that have been structured for research), and (d) in comparison to representative usual care. These four design features ensure that a study can answer such critical questions as how the treatment performs with the clientele and providers, and in the settings of actual clinical practice, and in comparison to the treatment that would ordinarily be provided in those settings. If a new treatment is designed to improve on current practice, research should certainly be structured to address this last question. Indeed, answering each of these questions is important for any treatment that is ultimately intended for use in real-world clinical practice.

This research strategy is part of a broader approach that has been identified in prior work (e.g., Weisz 2004, Weisz et al. 2005a) as the deployment-focused model of treatment development and testing. This model includes a series of steps designed to locate the process of building and testing interventions within the real-world contexts that are most relevant to the interventions’ intended applications. The idea is that interventions should be tested with the clientele, with the clinicians, and in the contexts for which they are ultimately intended, and tested for their effectiveness relative to current practice in those contexts to determine whether the new interventions actually represent an improvement. A primary aim of the
deployment-focused model is to create a process through which the treatment characteristics needed for success in real-world clinical application can be identified and built into the intervention as a natural part of the scientific process. The model rests on the testable premise that the potential of an intervention to be beneficial in a practice context is most likely to be realized if the treatment has been adapted to practice conditions as part of its ongoing development.

The model is consistent with evidence from implementation science (Fixsen et al. 2005) that when interventions that have been successful in one setting are moved to a very different setting, they may struggle or fail initially. Some will ultimately be successful, after multiple steps of intervention adaptation. The deployment-focused model is a way of making that adaptation process a natural, ongoing part of treatment development and testing, with the objective of efficiently building treatments that are practice-ready and likely to succeed in real-world clinical care contexts.

The deployment-focused model grew out of a concern that successfully implementing interventions in actual clinical practice may be most difficult when there are big differences between the implementation context and the context in which treatment development and testing took place. As suggested by Table 2, most randomized trials through which the majority of youth ESTs have been shaped and tested have not exposed the therapies to the broad array of factors present in the clientele, clinicians, and clinical care contexts of everyday treatment. That might not be such a big problem for psychotropic medication or other interventions whose mechanisms of action are mainly biological; but for psychosocial interventions, a process of development and testing that bypasses those human and environmental factors that can so profoundly affect treatment process and outcome could leave the resulting therapies poorly prepared for everyday treatment conditions.

A common pattern in youth therapy research has been for investigators to devote many years to carefully controlled efficacy studies and to build an evidence base for their protocols, with the plan that effectiveness tests in representative clinical conditions will come later. However, this last step has been rather rare in youth psychotherapy research to date. Moreover, even when that last step is taken, the differences between treatment within efficacy research and treatment in actual practice may well be too numerous and pronounced to be bridged in one final step of research. The number of dimensions along which treatment must be adjusted to span the lab-to-clinic gap may make the task of moving efficacy-tested treatments into everyday clinical care so complex that the task really needs to be a part of the ongoing treatment development process, from start to finish. Indeed, the very real-world factors that efficacy trial researchers might view as a nuisance or noise (e.g., youth comorbidity, parent pathology, family stressors 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, adding incentives for therapy attendance, hiring their own therapists) may in fact be precisely the kinds of factors that need to be understood and addressed if psychotherapy treatment protocols are to be created that fit well into clinical practice. ESTs that are stymied by these real-world factors may not fare so well in practice, no matter how strong they look in efficacy trials.

A related point is that implementing ESTs in practice settings may require creating adjunct interventions specifically to address obstacles. For example, treatments that call for weekly sessions with youths or parents may require new family engagement, problem anticipation, and problem-solving procedures (e.g., Nock & Kazdin 2005) to generate reliable attendance. And interventions that do not fit easily into an organization’s standard procedures (e.g., assessment or supervision requirements that go beyond clinic routines) may be workable only if paired with organizational assessment and with interventions designed explicitly to modify organizational practices and culture (e.g., Glisson & Schoenwald 2005). Developing and testing treatments, and potentially the
adjunctive interventions needed to support them, within the settings for which they are ultimately intended may be needed to support effective implementation in practice.

**CHALLENGES FOR THE DAYS AHEAD**

The goal of building empirically supported treatments that can fit smoothly into everyday practice and work well with clinically referred youths is worthwhile but challenging. In our pursuit of this goal, we have encountered challenges related to the referred youths, their families, the practitioners who serve them, the organizations within which the practitioners work, and the broader service system within which they all function. There are no villains here; the challenges arise because conditions, constraints, and possibilities are so very different from these different perspectives. Some of the challenges are summarized conceptually in Table 3, but a few more specific examples may clarify the kinds of concerns and issues that will need attention in the near future on the intervention development front.

1. **Growth in intervention scope and complexity can create implementation challenges.**
   Efforts to build treatments that encompass multiple and diverse problems and disorders can bump up against limitations in what busy practitioners have time to learn and master and in the resources available to sponsoring organizations. MATCH, for example, includes so much content that training alone requires 5–6 full days, and clinicians may need a year of subsequent weekly case consultations from MATCH experts before they can work independently. This places significant time demands on busy clinicians, who are likely to face productivity pressures in their workplaces. Our impression is that individual or very small group consultation works best, and that when consultation groups grow large, treatment quality and outcomes are undermined; however, more individualized consultation requires more expert consultants’ time and increases costs. It is possible that the time and cost requirements of complex treatments like MATCH may limit their implementability.

2. **Flexibility calls for clinical decision making, which cannot be entirely evidence based.**
   Protocols like MATCH that offer clinicians increased flexibility also increase clinicians’ responsibility for decisions that are not required in more linear protocols. Practitioners using MATCH must decide, for example, which module to use first, when to switch to another, when to rely on the default order in a flowchart and when to detour out of the default, whether (and if so, when) to shift to a new problem focus (e.g., from conduct to depression), and so forth. The flowcharts described earlier do provide useful guidance and a structure for such decision making, but they do not eliminate—in fact, they require—the use of clinical judgment. To illustrate this point, consider a youth in treatment for depression for whom the activity selection (i.e., behavioral activation) module has been used for three sessions but doesn’t seem to be working. Does this reflect simply a wrong choice of activities such that further tweaking and perseverance are needed, or does this represent interference in the form of misconduct and thus a need to shift to modules from the conduct section of MATCH? The flowchart requires a judgment on such questions but does not tell the clinician how to make that judgment. Moreover, there is no reliable way to assess whether such clinical judgments are correct. In sum, increased flexibility may actually necessitate increased clinical decision making, which cannot be entirely evidence based.

3. **Monitoring and feedback systems can inform, but not replace, clinical judgment.**
   The kinds of clinical decision making just discussed can certainly be informed and assisted by the evidence that monitoring and feedback systems (MFS) can provide (see Figure 5, for example). On one hand, the clinicians we work with place a high value on weekly data conveying how each youth is responding to treatment. On the other hand, these data do
not tell clinicians what is causing good or bad treatment response or what to do when treatment response is poor. If a youth’s problem levels have not changed after four weeks of treatment using one particular module, does that mean the clinician should try different ways of introducing and practicing that module, abandon that module and try another from the same problem area of MATCH, or shift entirely to a different problem focus? MFS data can tell the clinician what the youth’s trajectory of change has been and whether the youth is responding well to treatment, but the data cannot tell the clinician why this is happening or what to do in response. The MFS can substantially enrich the mix of data a clinician uses to inform clinical judgments—and that is a positive development—but the MFS cannot make the judgment for the clinician.

4. **Comprehensive treatment may lack a unified theoretical core.** MATCH may have practical utility in its coverage of both internalizing and externalizing problems, but that very coverage also means that it lacks a single unified theoretical core. A focus on anxiety and depression alone might build on theories and evidence on internalizing problems and even prior research on neurosis. MATCH is too broad to have such a core and thus tends to be more pragmatic than theoretical in its focus. This may be a problem in efforts to explain effects and identify mechanisms of change. It can also mean that the therapist facing a dilemma during treatment cannot resolve the dilemma by drawing guidance from a single core theory. However, a case could be made for the value of a treatment approach that builds on both CBT (for anxiety and depression) and behavioral theory (for conduct problems). It is possible that the most complete and accurate models accounting for therapeutic change may ultimately encompass both theoretical frameworks. If that turns out to be the case, the trans-theoretical approach underlying MATCH may be a virtue.

5. **Efforts to adapt and fit treatments must be structured to address the challenges of implementation and sustainability.** This article has focused mainly on the importance of developing and adapting ESTs to fit clinical care contexts and on how to accomplish that goal using strategies of the deployment-focused model. However, once these ESTs are developed and adapted, a great deal will need to be known about how to implement them effectively and cost-effectively in clinical practice settings. Building that knowledge base will require research on what are the best methods for training clinicians, what approaches to case consultation are most effective in building fidelity and competence, what kinds of organizational supports are needed to make implementation successful, and a variety of related questions. An associated question, a major one for many organizations and funders, is how best to sustain continued use of the ESTs and ensure continued fidelity after the clinician training and skill building are done to maximize ongoing benefit to those receiving treatment. Researchers have only recently begun to study these important topics, and a great deal remains to be learned (see McHugh & Barlow 2010, Stirman et al. 2012, Weisz et al. 2014).

6. **Adaptation and fitting become more complicated when we go beyond traditional treatment settings and outside the conventional US healthcare system.** As this article illustrates, it is a significant challenge to fit ESTs into traditional US mental health clinic contexts and the standard reimbursement systems within which those clinics function. The challenges grow even thornier when efforts reach beyond such standard contexts. Compared to treatment in traditional mental health clinics, for example, treatment in schools tends to involve briefer sessions, more unpredictable locations (there are no therapy rooms in most schools), and clinicians for whom therapy is a much less significant part of the workweek and job expectations (see, e.g., Fox et al. 2014, Stark et al. 2009). Therapy in primary care settings has its own distinctive characteristics (Asarnow et al. 2005). Complexity escalates...
dramatically when one tries to fit treatments tested in the United States into new cultural contexts and nations where the mental health and payment systems are markedly different. At the extreme, efforts to implement ESTs in countries that lack an organized mental health delivery system—and even for youth problems not seen in the United States (e.g., emotional problems of former child soldiers)—may require a totally different way of thinking about how to structure and deliver treatments (see, for example, Betancourt et al. 2014). Clearly, our work and that of our colleagues on adapting and fitting treatments to real-world intervention contexts is but a tiny part of what will be a massive ongoing process engaging many intervention scientists and practitioners and touching a broad array of organizational, cultural, and national contexts.

7. **Efforts to fit treatments to contexts may need to be complemented by changing the contexts.** Although it is useful to adapt treatments to existing contexts, it is important to note that the contexts themselves may also be modified to better reach and serve those who need mental health care. Kazdin & Blase (2011) have advocated for expanded models of treatment delivery that go beyond the traditional model of weekly 50-minute office visits to increase the public health impact of interventions. These authors propose a broad portfolio of delivery methods to reduce the burden of mental illness (see also Rotheram-Borus et al. 2012). Strategies particularly relevant for youths include embedding interventions within everyday settings such as summer camps (e.g., Pelham et al. 2005, Santucci & Ehrenreich-May 2013) and using technology to meet youths on their own turf through interactive, computer-based treatments such as Camp Cope-A-Lot (Khanna & Kendall 2010) the BRAVE Program (Spence et al. 2011) for youth anxiety, and SPARX (Merry et al. 2012) for youth depression. Efforts like these, restructuring where and how mental health services are received, will be a valuable complement to the kind of work described in this article.

**SUMMARY AND CONCLUSIONS**

Youth psychotherapy and psychotherapy research have deep historical roots and a shared objective of improving mental health and adaptive functioning in boys and girls. Despite this common ground, clinical care and treatment research have diverged significantly over the decades, such that most treatment research has tested interventions under conditions that do not look very much like the clinical practice that research is designed to strengthen. The research has generally not tested interventions with clinically referred youths treated by practicing clinicians in actual clinical care settings, and it has not addressed the critical question of whether interventions created through research actually produce better clinical outcomes than treatment as usual in practice settings. One consequence may be that the interventions identified as empirically supported, based on research that is not very clinically representative, may have difficulty coping with those real-world factors in everyday clinical care that they have not had to confront during development and testing. This may partially explain why the mean effects of ESTs look quite modest when tested under more representative conditions and compared in randomized trials to usual clinical care.

To address this problem, we have proposed a deployment-focused model: Intervention development and testing within this model are conducted from a very early stage with the kinds of participants (e.g., clients and clinicians) and in the kinds of contexts (e.g., community clinics) for which the intervention is ultimately intended, and new treatments are tested against usual care. The development and testing of one particular treatment protocol—a modular intervention called MATCH—illustrates key aspects of the deployment-focused model. The findings supporting the efficacy of MATCH are but one part of a rich research matrix needed in the days ahead. This should include MATCH replication trials, studies of optimum methods for implementation and
dissemination in treatment settings, research on ways to ensure sustainability once training and clinician skill building in a new treatment have ended, studies of treatment adaptation to fit contexts other than mainstream US clinics (e.g., schools, primary care, treatment settings in diverse cultures, and implementation in countries that lack any mental health care system), and research on adapting treatment delivery methods and contexts to better reach and serve those who need effective mental health care.

**DISCLOSURE STATEMENT**

John R. Weisz receives royalties on some of the works cited in this article.

**ACKNOWLEDGMENTS**

Some of the research reported here was supported by grants from the National Institute of Mental Health (MH57347, MH068806, MH085963), the Norlien Foundation, the MacArthur Foundation, Casey Family Programs, and the Annie E. Casey Foundation. We are grateful to these funders and to the children, families, clinicians, clinic administrators, and government program and policy leaders who have participated in and supported our research and enriched our thinking. We thank Alaina Baker, Samantha Cootner, Christiana Kim, and Grace Jung for their good work assisting in the preparation of this manuscript.

**LITERATURE CITED**


Contents

Thinking About Rumination: The Scholarly Contributions and Intellectual Legacy of Susan Nolen-Hoeksema
Sonja Lyubomirsky, Kristin Layous, Joseph Chancellor, and S. Katherine Nelson ........ 1

Clinical Dysfunction and Psychosocial Interventions: The Interplay of Research, Methods, and Conceptualization of Challenges
Alan E. Kazdin ................................................................. 25

Recent Advances in Autism Research as Reflected in DSM-5 Criteria for Autism Spectrum Disorder
Catherine Lord and Somer L. Bishop ........................................ 53

A Theory of States and Traits—Revised
Rolf Steyer, Axel Mayer, Christian Geiser, and David A. Cole ....................... 71

Biological and Psychosocial Predictors of Postpartum Depression:
Systematic Review and Call for Integration
Ilona S. Yim, Lynlee R. Tanner Stapleton, Christine M. Guardino,
Jennifer Habn-Holbrook, and Christine Dunkel Schetter .......................... 99

Shrinking the Gap Between Research and Practice: Tailoring and Testing Youth Psychotherapies in Clinical Care Contexts
John R. Weisz, Lauren S. Krumholz, Lauren Santucci, Kristel Thomassin,
and Mei Yi Ng .................................................................. 139

Obsessive-Compulsive and Related Disorders: A Critical Review of the New Diagnostic Class
Jonathan S. Abramowitz and Ryan J. Jacoby ................................ 165

Clinical Features, Cognitive Biases, and Treatment of Body Dysmorphic Disorder
Angela Fang and Sabine Wilhelm ................................................. 187

The Development and Course of Bipolar Spectrum Disorders: An Integrated Reward and Circadian Rhythm Dysregulation Model
Lauren B. Alloy, Robin Nusslock, and Elaine M. Boland .............................. 213

Etiologic, Phenomenologic, and Endophenotypic Overlap of Schizophrenia and Bipolar Disorder
Godfrey D. Pearlson ......................................................... 251
Developmental Dyslexia
Robin L. Peterson and Bruce F. Pennington ............................................. 283

Chronic Traumatic Encephalopathy: Historical Origins
and Current Perspective
Philip H. Montenigro, Daniel T. Corp, Thor D. Stein, Robert C. Cantu,
and Robert A. Stern ................................................................. 309

Depression and Cognition in the Elderly
Sophia Wang and Dan G. Blazer ....................................................... 331

fMRI Functional Connectivity Applied to Adolescent
Neurodevelopment
Monique Ernst, Salvatore Torrisi, Nicholas Balderston, Christian Grillon,
and Elizabeth A. Hale ................................................................. 361

Emotion Regulation and Psychopathology
Gal Sheppes, Gaurav Suri, and James J. Gross .................................... 379

Self-Reported Experiences of Discrimination and Health: Scientific
Advances, Ongoing Controversies, and Emerging Issues
Tené T. Lewis, Courtney D. Cogburn, and David R. Williams .................. 407

Indexes
Cumulative Index of Contributing Authors, Volumes 2–11 .................... 441
Cumulative Index of Article Titles, Volumes 2–11 ................................. 445

Errata
An online log of corrections to Annual Review of Clinical Psychology articles may be
found at http://www.annualreviews.org/errata/clinpsy