Does It Work in the Real World? The Effectiveness of Treatments for Psychological Problems in Children and Adolescents

Catherine M. Lee, Catherine Horvath, and John Hunsley
University of Ottawa

Despite the availability of hundreds of treatment studies in the scientific literature, we know little about whether these treatments work in regular practice. We present an updated review of treatment effectiveness studies for psychological problems in children and adolescents. A literature search yielded 20 recent articles describing effectiveness studies for the treatment of anxiety disorders, depression, and disruptive behavior problems. We compared data from these effectiveness studies with two benchmarks reported in meta-analyses of efficacy trials: the numbers of clients who completed services and the improvements found in those who completed services. All studies of the treatment of internalizing disorders reported completion rates above 80%; the majority of parenting interventions for the treatment of disruptive behavior problems reported that more than 75% of parents who began services completed them. The improvement rates reported in effectiveness studies for internalizing problems were comparable to the benchmarks reported in efficacy studies. There was greater variability in the treatment of disruptive behavior problems, with several studies outperforming the benchmark, and a smaller number yielding poorer results. Practitioners should be encouraged to see promising results that suggest evidence-based treatments for child and adolescent disorders can be effective when used in typical clinical settings.

Keywords: effectiveness, real world, evidence-based treatment

Psychologists want to use treatments that work in regular settings, with the types of clients who need services. Enthusiasm about the growing list of efficacious psychological treatments for emotional and behavioral problems in children and youth (e.g., Chorpita et al., 2011) is tempered by concern that treatments developed and tested under the highly controlled conditions of efficacy trials may not be robust enough to yield similar results when delivered in real-world settings by regular service providers to a clientele that may differ from that in randomized controlled trials (Baker-Ericzén, Hurlburt, Brookman-Frazee, Jenkins, & Hough, 2010). It is in this context that effectiveness studies (i.e., evaluations of the outcome of these treatments when provided to clients who are routinely referred for services) are likely to provide the kind of evidence psychologists need.

Although effectiveness studies were initially seen as the antithesis of efficacy studies, it is more accurate to think of the two designs as having different emphases. In efficacy studies, the focus is on minimizing threats to a study’s internal validity; in effectiveness studies, emphasis is placed on maximizing external validity, while maintaining an adequate level of internal validity. To provide an indication of treatment effects relative to what is typically found in randomized controlled trials (RCTs), data from effectiveness studies can be compared with a benchmark, or gold standard, established in efficacy trials. Originally developed as an aid for improving industrial and business practices (e.g., Camp, 1989), benchmarking is now commonly used in a range of organizations and institutions as part of a continuous quality improvement strategy.

Hunsley and Lee (2007) reviewed treatment effectiveness studies and compared their outcomes with benchmarks derived from relevant RCTs. They examined 21 studies of treatments for adult disorders and 14 studies of treatments for disorders of childhood and adolescence that provided information on completion rates and clinically significant outcomes. Hunsley and Lee’s review provided encouraging preliminary evidence of the transportability to real-world clinical settings of treatments with established efficacy. This article provides an update of their review, focusing specifically on psychological treatments for emotional and behavioral problems in children and adolescents. An update of this literature is timely given the substantial number of recently published effectiveness studies that examine treatments for child and adolescent clients. The evaluation of these treatments, vis-à-vis empirically...
based benchmarks, is crucial in determining the extent to which treatments with demonstrated efficacy are actually effective when used in typical clinical settings.

**Literature Search**

**Inclusion Criteria**

To be included in the review, a study had to meet six inclusion criteria. (1) The study had to examine a psychological treatment designed to reduce behavioral or emotional problems in children and adolescents. We did not include treatments for health problems such as asthma, diabetes, or pain. (2) The study was explicitly described as an effectiveness study or the article included statements that indicated that the study was designed to test whether a psychological treatment that had been shown to be efficacious in RCTs and meta-analyses would yield similar results when delivered in a regular or real-world clinical setting. We included studies that were associated with a university, as long as services were delivered by regular clinicians, but we excluded studies in which services were provided by graduate students (e.g., Fox & Holtz, 2009). Although it was not unusual for effectiveness studies to use treatment manuals, session checklists, and to include regular supervision, we considered that weekly review of videotapes and several hours a week of supervision exceeded what was commonly available in a real-world clinical setting and therefore excluded studies that reported such intense supervision (e.g., Gardner, Burton, & Klimes, 2006). (3) The study had to cite at least one supportive published efficacy study conducted prior to the effectiveness study. We did not require that identical treatment procedures be used in both the effectiveness and efficacy studies. Instead, we required that the treatment procedures were based on similar conceptual models and used treatment strategies similar to those that have been identified as core components in treatment of a particular problem (cf. Chorpita & Daleiden, 2009; Kaminski, Valle, Filene & Boyle, 2008). (4) Data were provided on the number of those who enrolled in services and the number who completed treatment. (5) Information on improvement in child or adolescent psychosocial functioning was provided. (6) Children and youth had been referred for treatment, or if they were recruited for the study through advertising, the client either met diagnostic criteria for a disorder or the mean pretreatment score of clients was above a clinical cutoff on an established measure such as the Eyberg Child Behavior Inventory (ECBI; Eyberg & Pincus, 1999), the Strengths and Difficulties Questionnaire (Goodman, 1997), or the Child Behavior Checklist (Achenbach & Rescorla, 2001).

**Article Search**

We examined the research published from the end of March 2006 to the end of December 2011 (i.e., since the period covered by Hunsley & Lee, 2007). First, a PsycINFO search was conducted of journal articles published in English using the terms “treatment effectiveness evaluation,” “treatment effectiveness,” “effectiveness trials,” and “psychotherapy.” In addition, we used the general terms “parent training,” as well as naming three specific treatments that have been frequently evaluated in efficacy trials of treatments for child and adolescent disorders, “Parent-Child Interaction Therapy,” “Incredible Years,” and “Triple P.” Second, we searched the tables of contents of 31 journals that frequently publish treatment research. A list of the journals consulted can be found in the Appendix.

In addition, as our third search strategy, we reviewed the reference lists of articles that were found to meet our inclusion criteria to obtain additional effectiveness studies that we had not identified with our initial two search strategies. As our fourth and final strategy, we also searched the references of recent reviews of child and adolescent treatment (e.g., Chorpita et al., 2011; De Los Reyes & Kazdin, 2009) for effectiveness studies.

For each potentially relevant article identified via these various strategies, we reviewed the abstract and, if necessary, the entire article, to determine whether the article met inclusion criteria. An article was retained for further consideration only if the first two authors agreed that all six inclusion criteria had been met; discrepancies were resolved by consultation with the third author. The original search yielded 148 records. After applying all of our selection criteria, our search yielded 20 child and adolescent treatment effectiveness studies meeting our criteria, published in English between April 2006 and December 2011.

**Completion and Outcome Benchmarks**

In the course of our article search we also identified a number of recent meta-analyses relevant to the conditions covered in our review. Once we had established the set of effectiveness studies, we then examined these meta-analyses of efficacy trials for the pertinent disorders to identify two benchmarks that are of particular relevance to clinicians: treatment completion and treatment improvement. In many cases, the meta-analyses did not provide such information so, for some disorders, there was only one meta-analysis pertinent for our purposes. In the event that there was more than one pertinent meta-analysis for a given disorder, we selected the most recent and comprehensive.

In examining the literature published since 2006, we found that the increasing use of the CONSORT Statement Flow Diagram (www.consort-statement.org) greatly facilitated our coding of treatment completion because articles that included a CONSORT flow diagram clearly showed the numbers of participants assigned to each treatment condition who began and completed treatment and who were available for posttreatment evaluation. Nevertheless, because there is no consensus on the definition “receiving the allocated intervention,” cross-study comparisons remain imprecise (Fernandez & Eyberg, 2009). We conducted independent-samples t tests (Statistics Calculator, 2012; Version 4.0), comparing the completion rates reported in the effectiveness studies with those reported in the meta-analytic benchmark. The other benchmark we extracted was the improvement observed following treatment. There was great variability in the improvement benchmarks described in the meta-analyses and reported in the effectiveness studies. These comparisons are described below.

**Comparing the Results of Treatment Effectiveness Studies With Benchmarks**

**Internalizing Disorders**

We identified seven studies that examined the effectiveness of treatment of internalizing disorders in children and adolescents; these studies are summarized in Table 1. These studies described
results for 394 young people who received treatment for internalizing problems. For internalizing disorders, we were able to obtain specific estimates of the percentage of clients whose improvement moved them from the clinical range to the nonclinical range. Typically the treatment outcome benchmark provided information on either (a) recovery/remission rates or (b) improvement rates (i.e., including both patients who recovered fully and those who experienced substantial clinical improvement but still remained symptomatic).

**Depression.** Three treatment-effectiveness studies for adolescent depression were identified. These reported on services offered in a range of settings, including a school and community mental health centers in the United States, as well as in a camp in Uganda for persons displaced by war. Potential participants were excluded from most studies. Because different specific treatments have been found to be efficacious for the various types of anxiety disorders, we selected one meta-analysis that reported benchmarks for mixed anxiety disorders and another for obsessive–compulsive disorder.

A meta-analysis by In-Albon and Schneider (2007) summarized results of 24 studies on cognitive–behavioral treatment of anxiety disorders other than OCD and PTSD and presented completion and improvement rates separately for individual, group, and family modalities. A comparison of data from the two effectiveness studies to these benchmarks indicated comparable completion data for family and group treatment and superior completion rates for individual cognitive behavioral therapy (CBT) for anxiety disorders. For children who received individual or group CBT, rates of recovery comparable to the benchmark were reported (i.e., in terms of recovery comparable to the benchmark were reported (i.e., in terms of improvement greater than the benchmark.

---

### Table 1

**Effectiveness Studies for Treating Internalizing Disorders in Children and Adolescents**

<table>
<thead>
<tr>
<th>Study</th>
<th>Site</th>
<th>Child age</th>
<th>Exclusion criteria</th>
<th>Treatment (n)</th>
<th>Completers</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Depression</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benchmark: Watanabe, Hunot, Omori, Churchill, &amp; Furukawa (2007)</td>
<td></td>
<td></td>
<td></td>
<td>85.8%</td>
<td>49.6%</td>
<td></td>
</tr>
<tr>
<td>Bolton et al. (2007): Uganda</td>
<td>DPC</td>
<td>14–17</td>
<td>Cognitive or physical disability</td>
<td>IPT (99)</td>
<td>82.8%</td>
<td>37.1% REC&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>Shirk et al. (2009): USA</td>
<td>School</td>
<td>14–18</td>
<td>BP: Psychotic</td>
<td>CBT (50)</td>
<td>91.0%</td>
<td>64.0% REC&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>Weisz et al. (2009): USA</td>
<td>CMHC</td>
<td>8–15</td>
<td>DD: Psychotic</td>
<td>CBT (32)</td>
<td>100%&lt;sup&gt;+&lt;/sup&gt;</td>
<td>70.0% REC&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

**Mixed anxiety disorders**

| Benchmark: In-Albon & Schneider (2007) | | | | | | |
| Bodden et al. (2008): Netherlands | CMHC | 8–18 | OCD; PTSD; ASD; MED; MR; SA; Untreated ADHD | CBT (64) | 87.1%<sup>a</sup> | 63.0% |
| Lau et al. (2010): Hong Kong | CMHC | 6–11 | Phobias | CBT (24) | | |

**Obsessive compulsive disorder**

| Benchmark: Abramowitz, Whiteside, & Deacon (2005) | | | | | | |
| Farrell, Schulp, & Boschen (2010): Australia | CMHC | 7–17 | Schizophrenia; MR | CBT (33) | 94.2%<sup>+</sup> | 63.0% REC<sup>+</sup> |
| Valderhaug et al. (2007): Norway | CMHC | 8–17 | Tourette’s; PDD | CBT (28) | 85.7%<sup>+</sup> | 50.0% REC<sup>+</sup> |

---

Note. Site: CMHC = community mental health clinic; DPC = displaced persons’ camp; Exclusion criteria: ASD = autism spectrum disorder; BP = bipolar disorder; DD = developmental disorder; MED = on medication to treat disorder; MR = mental retardation; OCD = obsessive-compulsive disorder; PDD = pervasive developmental disorder; SA = sexual abuse. Treatment: CBT = cognitive–behavioral therapy; Outcome: REC = recovered, no longer meets diagnostic criteria for the disorder.

<sup>a</sup>This benchmark is based on child and adolescents receiving pharmacotherapy or psychotherapy. The exclusion criteria were as follows: − = study result was significantly less than the benchmark; ns = study result was not significantly different from the benchmark; + = study result was significantly greater than the benchmark.
of the percentage of participating children who no longer met diagnostic criteria for the anxiety disorder for which they received services). The recovery rate at posttreatment was lower than the benchmark in the effectiveness study of family-based CBT. Because In-Albon and Schneider did not report the sample size for the follow-up recovery rates, it was not possible to compare the percentages statistically. The percentages reported by Bodden et al. (2008) indicate that the children continued to make gains at 3-month follow-up, and Lau, Chan, Li, and Au (2010) reported that gains were sustained at 3- and 6-month follow-ups.

Two studies of the treatment of pediatric OCD were compared with benchmarks reported in a meta-analysis of 10 studies of cognitive–behavioral therapy involving exposure and response prevention (Abramowitz, Whiteside, & Deacon, 2005). Analyses revealed that both completion and outcome data were comparable to the benchmarks. One study (Valderhaug, Larsson, Götestam, & Piancentini, 2007) reported continued gains at 6-month follow-up. Hunsley and Lee (2007) identified only one effectiveness study of the treatment of PTSD in children or adolescents following a natural disaster (Giannopoulou, Dikaikou, & Yule, 2006). Since 2007, another effectiveness study examining outcomes in children and youth receiving psychological services for PTSD following a natural disaster has been published (Jaycox et al., 2010). These studies were both conducted in “real-world” settings, included young people with clinically significant disorders and reported on completion and improvement data. However, there are no reviews reporting benchmarks for the treatment of PTSD in children exposed to natural disasters. We were therefore unable to include the two studies of PTSD related to natural disaster studies in our review.

### Disruptive Behavior Problems

We identified 13 studies examining the effectiveness of parent training in the treatment of disruptive behavior problems; these studies are summarized in Table 2. These studies included data on 1,118 parents who participated in parenting programs. The studies included a range of evidence-based parenting programs; the majority of these interventions were offered in community agencies, with two offered in a hospital and one at a university clinic. Ten of the reports were from Europe, two from Australia, and one from Hong Kong. The majority of the studies reported on services offered to treat preschool and elementary-school-age children, with only one focusing on services for adolescents. A third of the

<table>
<thead>
<tr>
<th>Study</th>
<th>Site</th>
<th>Child age</th>
<th>Exclusion criteria</th>
<th>Treatment</th>
<th>Completers</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark: Maughan et al. (2005)</td>
<td>78.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benchmark: Thomas &amp; Zimmer-Gembeck (2007)</td>
<td>78.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Axberg et al. (2007): Sweden</td>
<td>Diverse settings</td>
<td>3–9</td>
<td>N/R</td>
<td>IY (113)</td>
<td>73.8%*</td>
<td>Mother: $ES_w .67^* $Father: $ES_w .54^* $ $ES_w .73^* $</td>
</tr>
<tr>
<td>Baruch, Vrouw, &amp; Wells (2011): England</td>
<td>CMHC</td>
<td>10–17</td>
<td>N/R</td>
<td>PLL (123)</td>
<td>81.8%*</td>
<td></td>
</tr>
<tr>
<td>Breet et al. (2009): Belgium</td>
<td>University</td>
<td>4–7</td>
<td>None</td>
<td>PMT (34)</td>
<td>86.6%*</td>
<td>$ES_w .20^* $</td>
</tr>
<tr>
<td>Costin &amp; Chambers (2007): Australia</td>
<td>MHC</td>
<td>5–13</td>
<td>None</td>
<td>PMT (94)</td>
<td>83.9%*</td>
<td></td>
</tr>
<tr>
<td>Griffin et al. (2010): Ireland</td>
<td>Hospital &amp; CMHC</td>
<td>3–6</td>
<td>None</td>
<td>PPEYP (46)</td>
<td>70.8%*</td>
<td>$ES_w .45^* $</td>
</tr>
<tr>
<td>Hautmann et al. (2009): Germany</td>
<td>MHC</td>
<td>3–10</td>
<td>None</td>
<td>PEP (248)</td>
<td>77.7%*</td>
<td>$ES_w .39^* $</td>
</tr>
<tr>
<td>Hutchings et al. (2007): Wales</td>
<td>Community agencies</td>
<td>3–6</td>
<td>None</td>
<td>IY (104)</td>
<td>83.0%*</td>
<td>$ES_w .66^* $</td>
</tr>
<tr>
<td>Kleve et al. (2010): England</td>
<td>CAMHS</td>
<td>2–11</td>
<td>None</td>
<td>IY (128)</td>
<td>69.0%*</td>
<td>$ES_w .74^* $</td>
</tr>
<tr>
<td>Kling, Forster, Sundell, &amp; Melin (2010): Sweden</td>
<td>Social welfare clinics</td>
<td>3–10</td>
<td>No other rx</td>
<td>PMT (56)</td>
<td>96.5%*</td>
<td>$ES_w .72^* $</td>
</tr>
<tr>
<td>Leung et al. (2009): Hong Kong</td>
<td>Hospital</td>
<td>2–8</td>
<td>N/R</td>
<td>PCITT (48)</td>
<td>77.0%*</td>
<td>$ES_w .17^* $</td>
</tr>
<tr>
<td>Ogden &amp; Hagen (2008): Norway</td>
<td>Child service agencies</td>
<td>4–12</td>
<td>Severe MR; autism; parental problems</td>
<td>PMT (59)</td>
<td>88.0%*</td>
<td>$ES_w .16^* $</td>
</tr>
<tr>
<td>Phillips et al. (2008): Australia</td>
<td>CMHC</td>
<td>1.5–4.3</td>
<td>N/R</td>
<td>PCTT (43)</td>
<td>88.5%*</td>
<td>$ES_w 1.23^* $</td>
</tr>
<tr>
<td>Quinn et al. (2007): Ireland</td>
<td>Community clinics</td>
<td>4–7</td>
<td>N/R</td>
<td>PP (22)</td>
<td>95.6%*</td>
<td>$ES_w .49^* $</td>
</tr>
</tbody>
</table>

Note. Site: CMHC = community mental health clinic; Treatment: IY = Incredible Years; PLL = Parenting with Love and Limits; PMT = Parent Management Training; PEP = Prevention Program of Externalizing Problems; PP = Parents Plus; Outcome: F = father report; M = mother report. The exclusion criteria were as follows: = study result was significantly lower than the benchmark; ns = study result was not significantly different from the benchmark; + = study result was significantly greater than the benchmark.
studies did not provide information on exclusion criteria, whereas the majority reported that there were no exclusion criteria, and some studies explicitly examined the usefulness of the program when offered to families of children with serious comorbidity (e.g., Kleve et al., 2011; Quinn, Carr, Carroll, & O’Sullivan, 2007).

We used as our completion benchmark the data reported by Thomas and Zimmer-Gembeck (2007), who analyzed 24 studies of two widely adopted parenting interventions for parents of children aged 3–12 years with clinically significant externalizing problems: Parent-Child Interaction Therapy (PCIT: Eyberg & Robinson, 1982) and Triple P-Positive Parenting Program (Sanders, Markie-Dadds, Tully, & Bor, 2000). Fourteen of these studies reported data on attrition. Because the rates of attrition were significantly different in the three studies reporting on parent training delivered via a TV series, we excluded them from our calculations, yielding a completion benchmark of 78.9% based on data from 615 parents.

Analyses comparing the effectiveness studies with the benchmark indicated that 11 of the 13 effectiveness studies on treatment of disruptive behavior problems reported completion data that were comparable to the benchmark reported by Thomas and Zimmer-Gembeck (2007), one was inferior, and one reported superior completion rates. Although poorer completion rates were reported in a sample with no exclusion criteria in which one third of the children had a clinically defined neuro-developmental disorder, it is noteworthy that in this high-needs sample over two thirds of parents who began the program completed it (Kleve et al., 2011).

Unlike the effectiveness studies of treatments for anxiety and depression that routinely reported on children’s diagnostic status, studies of interventions for children and youth with disruptive behavior problems rarely included a formal diagnostic procedure. Only 50% of the effectiveness studies on treatments for disruptive behavior problems reported on some type of clinically significant change metric. The great variability in the exact definition and metric used across studies rendered cross-study comparisons of clinically significant change virtually impossible. Furthermore, we were unable to identify a recent meta-analysis of treatments for disruptive behavior disorders that reported an index of clinically significant change that we could use as a benchmark.

We therefore examined effect sizes reported in recent meta-analyses that we could use as a benchmark. In doing so, we had to be cognizant that about half of the effectiveness studies were single sample, prepost designs that yielded within-group effect sizes. Because effect sizes in such designs attribute all changes in scores to treatment effects, they are usually considerably larger than the values found with between-groups comparisons (see Lipsey & Wilson, 1993). Accordingly, we needed benchmarks for both between-groups and within-group effects. Such benchmarks were provided by Maughan, Christiansen, Jenson, Olympia, and Clark (2005) in their meta-analysis of 79 outcome studies that examined the effects of behavioral parent training on child and adolescent disruptive behavior problems. The mean effect size and confidence intervals presented in Table 2 were based on unbiased, weighted estimates, and, in the case of the between-groups values, were based on data comparing those who completed treatment with those who received no treatment. Although some of the effectiveness studies reported effect sizes, many did not. Therefore, we calculated effect size values (applying a correction for small sample sizes) for all studies reporting outcomes on the Eyberg Child Behavior Inventory Intensity scale, the Child Behavior Checklist Externalizing scale, or the Strengths and Difficulties Questionnaire Total Problems scale. In almost all instances when an effect size had been reported in a study, our calculated values were comparable to those reported in the original study, so Table 2 presents the values reported in the original studies. In one case though, the reported value greatly underestimated our calculation, so we reported our calculated value in the table for the Phillips, Morgan, Cawthorne, and Barnett (2008) study.

For the seven effectiveness studies that had used between-groups designs, we examined whether the effect sizes fell within the 95% confidence interval of the between-groups analyses reported by Maughan et al. (2005). Five effect sizes were above the upper end of the benchmark confidence interval, and two were below the lower end of the benchmark confidence interval. Two of the seven between-groups studies (Griffin, Guerin, Sharry, & Drumm, 2010, and Ogden & Hagen, 2008) compared a parent training intervention with treatment as usual. It would not be surprising if these studies yielded effect sizes smaller than the benchmark, as the benchmark was derived from comparisons of treatments to no-treatment control groups. Interestingly, although the effect sizes reported by Ogden and Hagen (2008) study were below the lower end of the confidence interval, the results of the Griffin et al. (2010) study were actually beyond the upper end of the confidence interval.

For the six effectiveness studies reporting within-group design results, we compared the effect sizes with the 95% confidence interval for the within-group analyses reported by Maughan et al. (2005). Because Axberg, Hansson, and Broberg (2007) reported separate results for data from mothers and fathers, there were seven effect sizes in total from the six studies: three fell within the 95% confidence interval of the benchmark, two were larger than the upper limit of the interval, and two were smaller than the lower limit of the interval. Four of the effectiveness studies also reported on follow-up effects, with further improvements noted at 3-month (Hautmann et al., 2009) and 5-month follow-up (Costin & Chambers, 2007; Griffin et al., 2010), and with the gains being maintained at 1-year follow-up in one study (Braet et al., 2009).

Conclusions and Implications

In the 5 years since the publication of the Hunsley and Lee (2007) review, the number of effectiveness studies of treatments for problems in children and adolescents has almost doubled, with the most marked increase in studies of interventions for children and adolescents with disruptive behavior problems. Our review identified studies in which a total of 394 young people received treatment for internalizing problems, and 1,118 parents participated in parenting programs delivered in regular clinical settings. This burgeoning literature reflects clinicians’ efforts to adopt evidence-based interventions and to determine whether they are helpful when delivered in a real-world setting. It is particularly noteworthy that the majority of the studies we reviewed addressed the transportability of interventions developed in North America to other countries, and for many, it included translation of materials into another language. Only two of the studies were based in North America.

We found that studies of real-world treatment of internalizing disorders were conducted on samples that excluded those with
severe mental disorders, such as psychosis, that could affect the young person’s capacity to participate in services. Studies found that more than 80% of those who begin services complete them. This figure is a considerable improvement on the completion rates in routine clinical services reported for youth by Nock and Ferriter (2005). Consistent with the trend reported by Hunsley and Lee (2007), treatment for depression yielded variable results, with some discouraging findings in a severely compromised sample in a displaced persons camp in Uganda, and more encouraging results showing that two thirds of young people treated in school or at a community clinic recovered. Again, consistent with the pattern found by Hunsley and Lee (2007) regarding treatments for anxiety disorders, over half the youth who completed treatment recovered from the disorder by the end of treatment. Studies that extended assessment to a follow-up period reported that clients maintained or continued to make gains in the months after services ended.

The large number of studies of interventions for disruptive behavior problems demonstrated very encouraging completion data. There was variability in treatment outcomes, but clear evidence that positive results can be obtained in a range of treatment settings, with samples of children experiencing multiple challenges. Such encouraging results provide substantial evidence that parenting programs for disruptive behavior problems can work very well in routine clinical practice.

As noted by Hunsley and Lee (2007), researchers must do more to make their findings directly relevant and meaningful to clinicians. For example, although there has been progress in the consistent reporting of the numbers of clients who complete services, there remains great variability in the precision with which treatment completion is defined (Fernandez & Eyberg, 2009). We acknowledge that it would be impossible to establish hard and fast definitions of what constitutes “treatment completion” because for many problems we do not yet know the minimum “dosage” of treatment that is required to obtain improvement. Nevertheless, it would be very helpful if researchers were to explain in their effectiveness study articles the way they define treatment completion.

Greater clarity could also be achieved in presenting treatment outcome data. Reports on treatments for depression and anxiety disorders routinely used some metric for clinically significant change indicating the proportions of clients who could be considered to have recovered from the problem for which services were sought. Unfortunately, there was little consistency in the definitions used, making comparisons difficult. Studies of the effectiveness of treatments for disruptive behavior problems are even more problematic in this regard. It is the exception, rather than the rule, for an article to indicate the percentage of youth who show clinically significant improvements following their parents’ involvement in interventions to help them deal with disruptive behavior problems. We recommend that clinical researchers routinely report the percentage of participants who achieve clinically significant change after receiving interventions. Further, we suggest that editorial policies of psychology journals should require the reporting of such information. Furthermore, we encourage editors to not underestimate the value of effectiveness studies based on quasi-experimental designs. We are convinced that clinical research is most likely to influence service delivery if it contains data that are of greatest direct relevance to practitioners. Overall, our findings suggest that psychologists can be cautiously optimistic about the clinical usefulness of treatments that have yielded positive results in replicated randomized controlled trials. As in our earlier review, the present pattern of findings suggests that clinicians who use evidence-based treatments when working with youth in real-world settings can achieve results comparable to those obtained in pertinent efficacy studies. However, the number of effectiveness studies is small compared with the number of efficacy trials, and effects of publication bias cannot be ruled out in evaluating the overall impact of effectiveness studies. We strongly encourage practitioners to conduct effectiveness studies, to determine the usefulness of an evidence-based treatment for the clientele they serve and to pay attention to effectiveness studies as they select treatments.

In addition, the findings from published studies also show that there is room for improvement with respect to the impact of our treatments. Innovation must address strategies to engage and retain in services the minority of children and adolescents (and their parents) who end services prematurely. Given the strong evidence that lower socioeconomic status families with significant parental psychopathology are at greater risk for dropping out of treatment, it must be a priority to develop services that are accessible and sensitive to the needs of these families (Fernandez & Eyberg, 2009). Furthermore, although many treatments are successful in helping children and adolescents to recover from their problems, there remain a substantial number of clients who do not respond positively to the evidence-based treatments currently available. Much work remains for psychologists, both in their roles as clinicians and as researchers, in better understanding how to help young people with psychological problems.

References


Appendix

Table of Contents Search of 31 Journals

Journals Searched

- American Journal of Psychiatry
- Archives of General Psychiatry
- Behavior Therapy
- Behaviour Research and Therapy
- Behavioral and Cognitive Psychotherapy
- British Journal of Clinical Psychology
- British Journal of Psychiatry
- British Medical Journal
- Child and Family Behavior Therapy
- Child and Adolescent Mental Health
- Clinical Child Psychology and Psychiatry
- Clinical Psychology and Psychotherapy
- Clinical Psychology: Science and Practice
- Cognitive Therapy and Research
- International Journal of Eating Disorders
- Journal of Child Psychology and Psychiatry
- Journal of Family Psychology
- Journal of Emotional and Behavioral Disorders
- Journal of the American Medical Association

- Journal of Anxiety Disorders
- Journal of Clinical Child and Adolescent Psychology
- Journal of Clinical Psychology
- Journal of Cognitive Psychotherapy
- Journal of Consulting and Clinical Psychology
- Journal of Nervous and Mental Disease
- Journal of the American Academy of Child and Adolescent Psychiatry
- Journal of Pediatric Psychology
- Mental Health Services Research
- Psychological Bulletin
- Psychotherapy
- Psychotherapy Research

Received April 18, 2012
Revision received August 31, 2012
Accepted September 4, 2012