REVIEW

Meta-Analysis: 13-Year Follow-up of Psychotherapy Effects on Youth Depression

Dikla Eckshtain, PhD, Sofie Kuppens, PhD, Ana Ugueto, PhD, Mei Yi Ng, PhD, Rachel Vaughn-Coaxum, PhD, Katherine Corteselli, EdM, John R. Weisz, PhD

Objective: Youth depression is a debilitating condition that constitutes a major public health concern. A 2006 meta-analysis found modest benefits for psychotherapy versus control. Has 13 more years of research improved that picture? We sought to find out.

Method: We searched PubMed, PsychINFO, and Dissertation Abstracts International for 1960 to 2017, identifying 655 randomized, English-language psychotherapy trials for individuals aged 4 to 18 years. Of these, 55 assessed psychotherapy versus control for youth depression with outcome measures administered to both treatment and control conditions at post ($\kappa = 53$) and/or follow-up ($\kappa = 32$). Twelve study and outcome characteristics were extracted, and effect sizes were calculated for all psychotherapy versus control comparisons. Using a three-level random-effects model, we obtained an overall estimate of the psychotherapy versus control difference while accounting for the dependency among effect sizes. We then fitted a three-level mixed-effects model to identify moderators that might explain variation in effect size within and between studies.

Results: The overall effect size (*g*) was 0.36 at posttreatment and 0.21 at follow-up (averaging 42 weeks after posttreatment). Three moderator effects were identified: effects were significantly larger for interpersonal therapy than for cognitive behavioral therapy, for youth self-reported outcomes than parent-reports, and for comparisons with inactive control conditions (eg, waitlist) than active controls (eg, usual care). Effects showed specificity, with significantly smaller effects for anxiety and externalizing behavior outcomes than for depression measures.

Conclusion: Youth depression psychotherapy effects are modest, with no significant change over the past 13 years. The findings highlight the need for treatment development and research to improve both immediate and longer-term benefits.

Key words: depression, children, adolescents, psychotherapy, meta-analysis

J Am Acad Child Adolesc Psychiatry 2019;■(■):■–■.

In 2011, the Grand Challenges in Global Mental Health initiative ranked unipolar depression as the number one challenge, more than twice as severe as the number two challenge (alcohol use disorders) in disability-adjusted life years.[1] Rates of depression accelerate during the school-age years, with one recent estimate of prevalence at 3% to 5% during ages 8 to 14 years and increasing to 20% during ages 14 to 17 years;[2] another report estimates that nearly 14% of adolescents will meet criteria for a depressive disorder before age 18 years.[3] Depression in young people is persistent and severe,[4] has high rates of relapse,[4,5] increases suicide risk,[6,7] and is associated with comorbid disorders[8] and functional impairment.[4,5,9,10] The psychiatric and psychosocial sequelae of depression in the school years persist well into adulthood.[11,12] Clearly, effective treatments are needed for depression in children and adolescents (herein "youths").

The National Institute of Mental Health (NIMH) emphasized in its 2007 and 2015 Strategic Plans the need for better treatments for mental disorders (unpublished data, March 2015 and April 2008).[13,14] For youth depression, psychotherapy has been recommended as the first-line treatment of choice.[15] However, meta-analytic findings have highlighted the challenge for psychotherapy with young people who are depressed. A meta-analysis of randomized controlled trials (RCTs) of youth psychotherapies, by Weisz *et al.* in 2006,[16] encompassing published studies and dissertations through December 2004, found that effects of psychotherapies for youths who are depressed were modest in their strength, breadth, and durability. The mean effect size (Hedges *g*) was 0.34 at posttreatment, dropping to 0.28 at follow-up assessments. An effect of 0.34 corresponds to a probability of 0.59 (versus chance at 0.50) that the average youth in the treatment group would be better off after treatment than the average youth in the control group.[17] These results suggest a need for more effective treatments for youth depression.

Two other relevant articles have been published since that 2006 meta-analysis.[16] One of these was a network meta-analysis (NMA) of youth depression treatment and prevention studies, conducted by Zhou *et al.* that included studies up to July 2014.[18] A primary finding was that only cognitive-behavioral therapy (CBT) and interpersonal therapy (IPT) were significantly more beneficial than most control conditions. Several procedural differences between Zhou *et al.* and Weisz *et al.* make it difficult to compare the two sets of findings or to derive an overall picture of the mean effect size for youth depression psychotherapies. Unlike Weisz et al., Zhou *et al.* did the following: included only interventions that used manualized or structured psychotherapy; excluded studies that used combinations of psychotherapies; used the procedure of selecting only one outcome measure per study for analysis rather than including all relevant outcome measures; consistent with the NMA approach, focused on specific comparisons among nodes ($n = 13$) and did not produce an overall effect size; and, using standard Bayesian NMA procedures, combined both direct comparisons of treatment conditions within the same RCT and indirect comparisons across RCTs. A limitation of such indirect comparisons is that they are not protected by randomization, so their validity rests on the assumption that all influential effect modifiers (eg, patient demographic, symptom severity, therapist characteristics, treatment setting, and duration) are matched across studies. As that assumption is not likely to be valid in psychotherapy research, NMA, although useful in a number of ways, is not a substitute for meta-analyses that restrict comparison to groups formed through random assignment—namely, comparisons within RCTs.

A second relevant meta-analysis was a broad synthesis of youth psychotherapy RCTs carried out by Weisz *et al.* of published studies between 1963 and 2013, encompassing treatments for depression, anxiety, attention deficit hyperactivity disorder (ADHD), and conduct-related disorders and problems.[19] This meta-analysis reported a mean depression treatment ES of $g = 0.29$ at posttreatment, 0.22 at follow-up. However, the breadth of focus restricted the identification of moderators of treatment effectiveness for depression-specific studies. Another limitation was that only published articles were included, creating a risk that estimated effects might have been inaccurate due to publication bias.

In the present meta-analysis, we provide an updated picture of youth psychotherapy effects, focusing specifically on depression treatment RCTs, and complementing the prior reports noted above by adding the following features: (1) We extended our literature search through the end of 2017, which is 3.5 years beyond Zhou *et al.* (2015),[18] 4 years beyond Weisz *et al.* (2017),[19] and 13 years beyond Weisz *et al.* (2006),[16] with a 64% increase in the number of studies (from 35 to 55). (2) Unlike Weisz *et al.* (2017),[19] we included not only published studies but also dissertations, which have been shown to provide an estimate of effects free of publication bias while also meeting research quality standards, possibly due to dissertation committee supervision.[20] (3) We did not exclude treatments that used combinations of psychotherapies. (4) We tested a fuller array of potential moderators of treatment effects at both the between-study and within-study levels, and with greater specificity than in the previous meta-analyses, using all of the moderators that were included in Weisz *et al.* (2017),[19] but focusing specifically on depression, as well as adding the moderators of study location and treatment type that were not included in Weisz *et al.* (2006).[16] (5) We included tests of all depression measures rather than selecting one for each study, and we adjusted for ES dependency by using a multilevel meta-analytic approach. (6) We included all nondepression mental health outcome measures, to test the extent to which depression treatment effects were specific to depression versus generalizing across other dimensions of mental health. (7) We restricted our analyses to direct within-study comparisons, given the concern that psychotherapy studies are likely to vary in distributions of effect modifiers (either reported or unreported) in ways that could affect the validity of indirect comparisons. Given the increased number of studies in recent years and continuing efforts by researchers to improve treatments and test new approaches, it is possible that effects have grown larger in more recent years. In addition, meta-analyses should be updated periodically to provide a current picture of youth psychotherapy effects,[21] especially considering the large increase in the number of studies and the improvement in the meta-analytic approach in comparison to the Weisz *et al.* (2006) meta-analysis.[16]

The main research questions of this meta-analysis are as follows: (1) What is the overall posttreatment effect of psychotherapy on youth depression, and are there differences in effect size between or within studies? (2) Can any differences in effect sizes be explained by study or outcome characteristics, considering 12 candidate moderators identified as important in prior research on youth depression? We tested the following four primary candidate moderators: (1) informant: youth, parent, and other (therapist, teacher, and others such as clinical interviewer), as intervention outcomes can differ depending on the informant.[16,19,22] (2) Treatment format: individual, group, and mixed individual and group. These were the most commonly used formats. (3) Treatment type: CBT, IPT for adolescents (IPT-A), and CBT and additional treatment (Table 1[23-87]). CBT and

TABLE 1 Characteristics of the 55 Youth Depression Psychotherapies Included in the Meta-Analysis

Study	Type of Sample	Mean Sample Size Used to Compute ESs at Post Treatment	Mean Age/Age Group	% Boys	Type of Treatment(s)	Treatment Protocol(s)	Type of Control Group(s)	Attrition (%)	Nondepression Measures Assessed	Overall ES	Depression ES Post Treatment	Depression ES Follow-up
Ackerson *et al.* (1998)[23]	Diagnosed community sample	22	15.9	36.4	CBT bibliotherapy		Waitlist	26.67		−0.02	−0.02	
Asarnow *et al.* (2005)[24] [Follow-up: Asarnow *et al.* (2009)[25]]	Symptomatic primary care sample	331	17.2	22	CBT	CWD-A	Enhanced usual care	20.81		0.18	0.21	0.15
Asarnow *et al.* (2002)[26]	Symptomatic school sample	23	10	35	CBT + family education intervention		Waitlist	Not reported		−0.30	−0.30	
Bolton *et al.* (2007)[27]	Symptomatic internally displaced youths in war-affected northern Uganda	175	14.97	43	IPT-A; client-centered intervention		Waitlist	16.46		0.79	0.79	
Brent *et al.* (1997)[28] [Follow-up: Brent *et al.* (1998)[29]]	Diagnosed recruited plus referred outpatient sample	56.85	15.6	24.3	CBT; non-behavioral family intervention	Systemic behavioral family therapy	Nondirective supportive therapy	19.91	Anxiety problems, Conduct problems	0.15	0.21	0.21
Charkhandeh *et al.* (2016)[30]	Diagnosed outpatient sample	124	15.26	46.28	CBT; Reiki therapy		Waitlist	0		1.32	1.34	
Clarke *et al.* (1995)[31]	Symptomatic high school sample	116	15.3	30	CBT	CWD-A	No treatment	22.67	Anxiety problems, conduct problems	0.08	0.31	0.10
Clarke *et al.* (2001)[32]	Symptomatic at-risk sample	57.5	14.6	35.6	CBT + nondescript usual care	CWD-A	HMO usual care	38.83	Conduct problems	0.09	0.14	0.15
Clarke *et al.* (2002)[33]	Diagnosed offspring of parents who are depressed from HMO	88	15.3	31.1	CBT	CWD-A	HMO usual care	0	Conduct problems	−0.008	−0.02	0.03
Clarke *et al.* (1999)[34]	Diagnosed recruited sample	57.75	16.2	29.2	CBT; CBT + parent group	CWD-A	Waitlist	27.37	Conduct problems	0.17	0.25	

(*continued*)

TABLE 1 Continued

Study	Type of Sample	Mean Sample Size Used to Compute ESs at Post Treatment	Mean Age/Age Group	% Boys	Type of Treatment(s)	Treatment Protocol(s)	Type of Control Group(s)	Attrition (%)	Nondepression Measures Assessed	Overall ES	Depression ES Post Treatment	Depression ES Follow-up
Curtis (1992)[35]	Diagnosed high school sample	19.5	15.8	26.1	CBT	CWD-A	Waitlist	15.22	Conduct problems	0.77	1.50	
Dana (1998)[36]	Symptomatic special education youths	18	10.49	87	CBT + social skills	CWD-A + Skillstreaming the Elementary School Child	No treatment	5.26	Conduct problems	0.17	0.15	
De Cuyper *et al.* (2004)[37]	Symptomatic school sample	18	10	25	CBT	Taking Action	Waitlist	18.18	Anxiety problems	0.21	0.27	0.93
Diamond *et al.* (2002)[38]	Diagnosed youths referred by parents or school	32	14.9	22	Non-behavioral family intervention	ABFT	Waitlist	Not reported	Anxiety problems, Conduct problems	0.65	0.57	
Doerr (1984)[39]	Symptomatic middle school sample	27	12.93	36.11	CBT		Usual care	25		1.14		1.14
Ettelson (2002)[40]	Symptomatic high school sample	20	15.2	44	CBT		Supportive contact waitlist	20	Anxiety problems	0.81	0.89	
Fischer (1995)[41]	Symptomatic youths in detention	16	12–17	87.5	CBT		Attention placebo	0		0.37	0.37	
Fleming *et al.* (2012)[42]	Symptomatic adolescents excluded from mainstream education	30	14.9	56	Computerized CBT	SPARX	Waitlist	6.25	Anxiety problems	0.47	0.75	
Garber *et al.* (2009)[43] [Follow-up: Brent *et al.* (2015)[44]]	Symptomatic offspring of parents who are depressed	293.5	14.8	41.5	CBT	CWD-A	Usual care	7.12		−0.27	−0.29	−0.26
Gillham *et al.* (2006)[45]	Symptomatic early adolescents from primary care	192.43	11.5	46.86	CBT	Penn Resiliency Program	Usual care	28.99		0.17	0.07	0.20
Gillham *et al.* (2012)[46]	Symptomatic school sample	235	12.5	52	CBT; CBT + parent involvement	Penn Resiliency Program	Usual care	12.48	Anxiety problems	0.03	0.04	

(*continued*)

TABLE 1 Continued

Study	Type of Sample	Mean Sample Size Used to Compute ESs at Post Treatment	Mean Age/Age Group	% Boys	Type of Treatment(s)	Treatment Protocol(s)	Type of Control Group(s)	Attrition (%)	Nondepression Measures Assessed	Overall ES	Depression ES Post Treatment	Depression ES Follow-up
Gillham *et al.* (2007)[47]	Symptomatic middle school sample	289.79	12.13	53.95	CBT	Penn Resiliency Program	Psychotherapy placebo; No treatment	37.61		0.06	0.02	0.07
Hickman (1994)[48]	Diagnosed youths attending a day treatment program	9	9.6	88.89	Social skills intervention		Usual care	0		0.71	0.61	0.81
Israel and Diamond (2012)[49]	Diagnosed youths referred by clinic	20	15.6	45	Non-behavioral family intervention	ABFT	Usual care	0		0.89	0.89	
Jeong *et al.* (2005)[50]	Symptomatic school sample	40	16	0	Dance movement therapy		Waitlist	0	Anxiety problems, conduct problems	0.09	-0.04	
R.C. Kahn (1989)[51]	Symptomatic youths abroad	29	15.97	47	Psychodynamic		No treatment	Not reported		−0.19	−0.37	0.001
J.S. Kahn *et al.* (1990)[52]	Symptomatic school sample	28.67	12.1	48.5	CBT + non-behavioral parent training; relaxation training; modeling intervention	CWD-A	Waitlist	15.69		0.71	0.67	0.79
Lewinsohn *et al.* (1990)[53]	Diagnosed recruited sample	39	16.2	39	CBT; CBT + parent group	CWD-A	Waitlist	24.26	Anxiety problems, conduct problems	0.76	0.83	
Liddle and Spence (1990)[54]	Symptomatic school sample	21	9.2	67.7	Social skills intervention		Attention placebo; waitlist	0		0.40	0.51	0.29
Listug-Lunde *et al.* (2013)[55]	Symptomatic middle school sample	15.5	12.44	62.5	CBT	CWD-A	Usual care	18.42	Anxiety problems	−0.12	−0.11	−0.37
Luby *et al.* (2012)1[56]	Diagnosed community sample	43	4.5	62.79	Behavioral parent training	Parent-Child Interaction Training	Psychotherapy placebo	20.37	Attention-deficit/ hyperactivity problems, Conduct problems	0.07	0.15	
McCarty *et al.* (2013)[57]	Symptomatic middle school sample	103.25	12.7	43.5	CBT	Positive Thoughts and Actions	Psychotherapy placebo	13.96	Attention-deficit/ hyperactivity problems, Conduct problems	0.28	0.27	

(*continued*)

TABLE 1 Continued

Study	Type of Sample	Mean Sample Size Used to Compute ESs at Post Treatment	Mean Age/Age Group	% Boys	Type of Treatment(s)	Treatment Protocol(s)	Type of Control Group(s)	Attrition (%)	Nondepression Measures Assessed	Overall ES	Depression ES Post Treatment	Depression ES Follow-up
McLaughlin (2010)[58]	Symptomatic school sample	22	11.82	59	CBT	CWD-A	Usual care	4.35		0.25	0.25	
Merry *et al.* (2012)[59]	Symptomatic adolescent sample	143	15.56	34.22	Computerized CBT	SPARX	Usual care	23.53	Anxiety problems	0.17	0.29	0.16
Moldenhauer (2004)[60]	Symptomatic primary care sample	19	14.6	27	CBT	CWD-A	Healthy lifestyle class	26.92	Anxiety problems, conduct problems	0.14	-0.16	
Mufson *et al.* (1999)[61]	Diagnosed outpatient sample	32	15.8	27.1	IPT-A		Clinical monitoring	33.33		0.48	0.48	
Reynolds and Coats (1986)[62]	Symptomatic high school sample	16.5	15.7	36.7	CBT; relaxation training		Waitlist	17.54	Anxiety problems	1.27	1.58	1.29
Rohde *et al.* (2004)[63]	Diagnosed sample in juvenile justice system	88	15.1	65	CBT	CWD-A	Tutoring and life skills training	5.38	Anxiety problems, attention-deficit/ hyperactivity problems, conduct problems	0.000025	0.26	-0.08
Rohde *et al.* (2014)[64] [Follow-up: Rohde *et al.* (2015)[65]]	Symptomatic high school sample	251	15.5	32	CBT; CBT bibliotherapy	CWD-A	Educational brochure	0		0.06	0.14	0.03
Rosselló and Bernal (1999)[66]	Diagnosed, school-referred sample	34	14.7	46	CBT; IPT-A		Waitlist	27.67	Conduct problems	0.33	0.35	
Sanford *et al.* (2006)[67]	Diagnosed community sample	29.5	15.84	35.48	Family psychoeducation + usual care		Usual care	4.84		0.48	0.54	0.44
Santomauro *et al.* (2016)[68]	Symptomatic community sample with comorbid autism spectrum disorders	20	15.75	60	CBT	Exploring Depression	Waitlist	13.04		−0.27	−0.27	
Stark *et al.* (1987)[69]	Symptomatic school sample	18.5	11.2	57.1	CBT (self-control training); CBT (problem solving)		Waitlist	2.63	Anxiety problems	0.43	0.67	

(*continued*)

TABLE 1 Continued

Study	Type of Sample	Mean Sample Size Used to Compute ESs at Post Treatment	Mean Age/Age Group	% Boys	Type of Treatment(s)	Treatment Protocol(s)	Type of Control Group(s)	Attrition (%)	Nondepression Measures Assessed	Overall ES	Depression ES Post Treatment	Depression ES Follow-up
Stasiak *et al.* (2014)[70]	Symptomatic high school sample	34	15.2	58.82	Computerized CBT	The Journey	Attention placebo	0		0.10	0.26	-0.06
Stice *et al.* (2007)[71]	Symptomatic high school and college sample	75.75	18.4	30	CBT; CBT bibliotherapy; supportive intervention	Feeling Good	Psychotherapy placebo; waitlist	0		0.14	0.13	0.14
Stice *et al.* (2008)[72] [Follow-up: Stice *et al.* (2010)[73]]	Symptomatic high school sample	169.2	15.6	44	CBT; CBT bibliotherapy; supportive expressive intervention	CWD-A; Feeling Good	Assessment only	0	Anxiety problems	0.24		0.54
Szigethy *et al.* (2007)[74] [Follow-up: Thompson *et al.* (2012)[75]]	Diagnosed youths with comorbid pediatric inflammatory bowel disease	34.33	14.99	49	CBT	PASCET for Physical Illness	Treatment as usual plus an information sheet about depression	16.26	Anxiety problems, attention-deficit/ hyperactivity problems	0.49	0.69	0.37
Treatment of Adolescents with Depression Study (TADS) Team (2004)[76] [Follow-ups: Kennard *et al.* (2006)[77]; Vitiello *et al.* (2009)[78]]	Diagnosed outpatient sample	223	14.6	45.6	CBT		Medication placebo	0	Anxiety problems	0.09	0.07	−0.06
Vostanis *et al.* (1996a)[79] [Follow-ups: Vostanis *et al.* (1996b)[80]; Vostanis *et al.* (1998)[81]]	Diagnosed outpatient sample	55.67	12.7	44	CBT		Psychotherapy placebo	Not reported	Anxiety problems, conduct problems	−0.08	0.35	−0.23
Weisz *et al.* (2009)[82]	Diagnosed community sample	42.29	11.77	44	CBT	PASCET	Usual care	25.81	Conduct problems	0.11	0.13	

(*continued*)

TABLE 1 Continued

Study	Type of Sample	Mean Sample Size Used to Compute ESs at Post Treatment	Mean Age/Age Group	% Boys	Type of Treatment(s)	Treatment Protocol(s)	Type of Control Group(s)	Attrition (%)	Nondepression Measures Assessed	Overall ES	Depression ES Post Treatment	Depression ES Follow-up
Weisz *et al.* (1997)[83]	Symptomatic school children	48	9.6	54.2	CBT	PASCET	No treatment	0		0.33	0.34	0.32
Yang *et al.* (2016)[84]	Diagnosed school sample	28.67	14.96	44.44	Active attention bias modification		Placebo attention bias modification	36.3	Anxiety problems	0.27	−0.04	0.79
Young *et al.* (2006)[85]	Symptomatic school sample	40	13.4	14.6	IPT-A		School counseling	2.44		1.23	1.53	1.10
Young *et al.* (2010)[86]	Symptomatic high school sample	50.67	14.51	40.3	IPT-A		School counseling	11.11		0.28	0.68	-0.63
Yu and Seligman (2002)[87]	Symptomatic school sample	211	11.8	55.45	CBT	Penn Optimism Program	No treatment	4.09		0.3265	0.25	0.37

Note: ABFT = attachment-based family therapy; CBT = cognitive-behavioral therapy; CWD-A = Adolescent Coping With Depression Course; IPT-A = Interpersonal Psychotherapy for Depressed Adolescents; ES = Hedges g; PASCET = Primary and Secondary Control Enhancement Training.

IPT-A were selected because these two approaches currently meet criteria for a well-established intervention for youth depression.[88] Studies combining CBT with additional treatment were included, as they were the second most common category. (4) Control condition: no treatment/waitlist, psychotherapy and medication placebo, and usual care (UC) treatment (treatment used in usual practice). Consistent with Weisz *et al.*,[16] we compared treatments to both passive and active control conditions, including UC that is considered to be an important externally valid control condition to evaluate.[89] As a secondary focus, we tested eight additional moderators highlighted in prior research:[16,19,22,90-95] year of study, study location (North America and outside of North America), participant engagement (recruited [recruited from those not seeking or receiving mental health services independently of the study] and referred [recruited from clients seeking or receiving outpatient or inpatient mental health services or school-based mental health services]), ethnicity (white sample [≥50% white] and nonwhite sample [<50% white]), sex (majority male participants [>50% male participants] and majority female participants [>50% female participants]), developmental period (consistent with the Weisz *et al.* [2006][16] definition of this variable, childhood [mean age <13 years] and adolescence [mean age ≥13 years]), diagnosis requirement (all participants required to meet formal diagnostic criteria and not required [some or none of the participants met diagnostic criteria]), treatment setting (clinical [inpatient hospital, day treatment program/partial day hospital, nonuniversity outpatient hospital/clinic, or combinations of more than one of these settings] and nonclinical [university or research/laboratory outpatient hospital/clinic, school, community setting such as a summer camp, home, or combinations of more than one of these settings]). (3) How lasting are psychotherapy effects? (4) Are psychotherapy effects specific to depression-related outcomes, or do they generalize to other outcomes, including anxiety and externalizing problems?

METHOD

Data Sources and Study Selection

Our search focused on RCTs testing youth psychotherapies for depression, including peer-reviewed publications and dissertations. We searched PsycINFO and PubMed for January 1960 to December 2017. The PsycINFO search used 21 search terms linked to psychological therapy (eg, psychother-, counseling) that had been used in previous youth therapy meta-analyses, crossed with outcome-assessment topic and age-group constraints. PubMed's indexing system (Medical Subject Headings [MeSH]) uses different keywords for the same concepts, and we used

Mental Disorders with these limits: clinical trial, child, published in English, and human subjects. Additional search methods included examining reviews and meta-analyses of youth psychotherapy research, reference trails in the identified reports, and additional studies by psychotherapy researchers whom we contacted. To address publication bias, we followed meta-analyses guidelines that recommend including unpublished studies of acceptable methodological quality by including dissertations.[20] Dissertations are appropriate because they are (1) free of publication bias; (2) reliably identifiable through a systematic search; and (3) strong in methodological quality, perhaps due to faculty committee supervision.[20] Dissertations were identified using Dissertations Abstracts International using the same search terms as for the published literature search.

Study and measure inclusion criteria were as follows: (a) participants selected and treated for depression; (b) random assignment of youths to treatment versus control conditions, with at least one of the treatment conditions being psychotherapy (pharmacotherapy alone or in combination with psychotherapy were excluded); (c) mean youth age of 4 to 18 years; (d) outcome measures administered to both treatment and control conditions at post and/or follow-up measurements; and (e) English language. Our operational definition of depression included either a depressive disorder diagnosis (*DSM* or *International Statistical Classification of Diseases and Related Health Problems* [*ICD*]) or elevated symptoms (eg, clinical range scores on standardized measures). Both diagnosis and elevated symptoms were included for the following reasons: (a) both are common and often used in the youth treatment outcome literature[92]; (b) youths with elevated symptoms have been shown to experience serious impairment[96,97]; (c) elevated symptoms often prompt more referrals than formal diagnosis[98,99]; and (d) formal diagnostic categories and criteria (using *DSM* and *ICD*) have varied markedly across the years. Using these selection criteria, the meta-analysis includes 55 randomized controlled trials, both peer-reviewed articles and dissertations. Figure 1 shows the study search and identification flowchart.

Data Extraction, Coding, and Processing

We coded studies for multiple study and sample characteristics and assessed intercoder agreement. Seven coders each coded 20 to 30 randomly selected studies independently. The most experienced coder, an RCT researcher with a doctorate in clinical psychology, was the master coder against which other coders were compared. Other coders were clinical psychology postdoctoral fellows and graduate students. We included continuous codes attaining intraclass correlation coefficients (ICCs) within or above the "excellent" range ($\geq$0.75) according to Cicchetti and Sparrow,[100] and categorical codes attaining kappas (κ) within Cohen's "substantial" (0.61–0.80) and "almost perfect" (>0.80) ranges.[101]

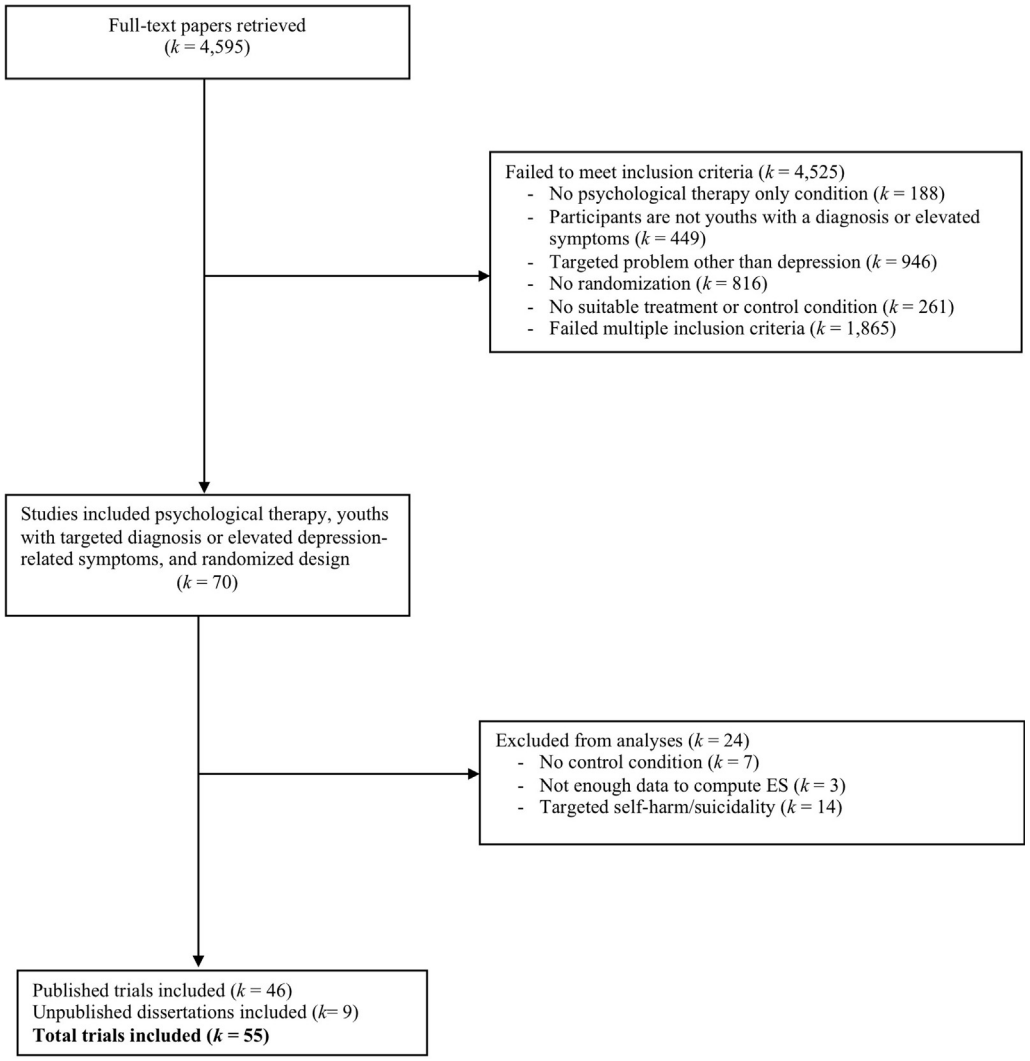
Intercoder agreement was assessed for control conditions (waitlist/no-treatment, psychotherapy, and pill placebo, and UC treatment in which therapists used whatever treatments they employed in their usual practice; $\kappa = 0.85$); whether a diagnosis was required for study inclusion (diagnosis required versus not required; $\kappa = 0.84$); participant engagement (referred versus recruited for the study; $\kappa = 0.62$); treatment setting (clinical versus nonclinical; $\kappa = 0.64$); percentage of participants who were male (ICC = 0.96) and who were white (ICC = 0.87), and mean age (ICC = 0.99), with all 3 then dichotomized for analysis into majority ($\geq$50%) male participants versus female participants, majority ($\geq$50%) white versus minority, and majority children (mean <13 years) versus adolescents; informant (youth, parent, and other; $\kappa = 0.87$); and treatment type ($\kappa = 0.85$) then collapsed into 5 categories (CBT, Interpersonal Psychotherapy for Adolescents [IPT-A], and CBT combined with other treatments, other youth-focused behavioral treatments [treatments whose components are based on behavioral or learning principles, and for which learning is a key mechanism through which change is hypothesized to occur], and other youth-focused nonbehavioral treatments [treatments whose components are based on nonbehavioral psychological principles and in which insight is a key mechanism through which change occurs); must be designed by the authors as a genuine therapy intended to produce change; and do not include interventions described as a "control" group, or a group designed "to control for" attention/nonspecific factors; eg, psychodynamic, nonbehavioral family intervention).

We also coded methodological rigor, as indicated by participant blinding (subjects not being aware of assessment or being able to influence the assessment; $\kappa = 0.62$), attrition (sample size ICC = 0.99, then coded as the percentage in the initial participants available to compute ES for that target problem at the particular measurement point), objective (behavioral counts, such as homework completion) versus subjective (self- or other-report, eg, by family member or school, treatment, or research staff) measure ($\kappa = 0.87$), power (sample size ICC = 0.99, then coded into adequate power [sample n = 128, providing power of 0.80 to detect an ES of 0.50 with $\alpha = 0.05$] versus inadequate); and, for the treatment conditions, presence of pre-therapy training (mean κ = 0.74), adherence/fidelity checks (mean $\kappa = 0.77$), and treatment manual or structured guide (mean $\kappa = 0.60$).

Effect Size Calculation

Effect sizes were calculated as Cohen's *d*,[101] assessing the standardized mean difference between treatment and

FIGURE 1 Flowchart Showing Study Retrieval, Review, Exclusion, and Inclusion



control conditions, divided by the pooled SD on measures of the problem targeted in treatment. Effect size calculations used either data reported or data provided by study authors whom we contacted to obtain data not provided in the written reports. We used the following procedure for studies that did not provide means and standard deviations: (1) transforming data to *d* using Lipsey and Wilson procedures when studies reported other metrics (eg, frequencies)[102]; (2) assigning the minimum *d* that would produce that significance level given the sample size (0.37% of our cases) when studies reported only *p* values or significant effects (assumed to reflect $p < 0.05$ if not otherwise stated); (3) assigning $d = 0$ when studies reported only a nonsignificant effect[103] (3.73% of our cases). All ES values were adjusted using the Hedges small sample correction,[104] which yields an unbiased estimate of the population standardized mean difference (*g*). We used data from observed cases/completers to compute effect sizes. If data were provided both in terms of intent-to-treat and observed cases/completers, data of observed cases/completers were used.

Data Synthesis

Meta-analytic Approach. The assumption of effect size independence was violated because 94.5% of the studies yielded multiple ESs stemming from multiple outcome measures, multiple informant reports, and/or multiple time points. We used a multi-level approach that allowed us to include all ESs in nonaggregated form per study to model this dependency. A 3-level random-effects model encompassed the sampling variation for each ES (level 1), within-study variation (level 2), and between-study variation (level 3), with variance partitioning coefficients

reflecting the percentage of variation that lies at a particular level. This extension of the commonly used random-effects meta-analytic model was used to examine main treatment effects. We further fitted a 3-level mixed effects model to identify moderators that might explain variation in treatment effects between and within studies by adding study (level 3) and outcome (level 2) characteristics. Continuous or dichotomous moderators were tested using a Wald test. For categorical moderators with more than two categories, the omnibus F test was used in conjunction with pairwise comparisons to examine which subgroup mean ESs differed significantly. Because parameter estimates are poor with a small number of studies, moderator tests were conducted only with categories that contained at least five studies. The percentage of explained variance reflected the decrease in variance with the addition of the particular moderator to the model. Parameters were estimated using the restricted maximum likelihood procedure implemented in SAS PROC HPMIXED with observed ESs weighted by the inverse of the sampling variance.

Publication Bias. Because bias against submitting and publishing null or negative findings could inflate mean treatment effects, we addressed risk of publication bias in three ways. First, we included unpublished dissertations, as discussed above. Second, we compared the mean ES for published studies versus dissertations; the difference was not significant [$t(555) = -1.01$; $p = .314$]. Third, we created a funnel plot,[105] with standard error plotted on the vertical axis as a function of ES on the horizontal axis. With absent publication bias, the plot should resemble an inverted funnel with studies distributed symmetrically around the mean ES. The Egger weighted regression test[106] showed that our plot was symmetrical at posttreatment [$t(51) = 0.80$, $p = .428$], but was asymmetrical at follow-up [$t(30) = 2.34$, $p = .026$]. Applying the trim-and-fill procedure[107] revealed that five studies were missing at the left side of the plot, resulting in a slight reduction of the adjusted ES, suggesting that some impact of publication bias for follow-up studies.

Risk of Bias. Because less rigorous studies have been found to overestimate mean treatment effects, we assessed methodological rigor using the following risk of bias criteria suggested by others[108,109]: (1) subject blindness to assessment, (2) participant attrition, (3) measurement objectivity, (4) adequately powered study and, for the treatment conditions, (5) presence of pretherapy training (therapists received training in the use of the particular psychotherapy as part of the study regardless of general or prior knowledge or use of the psychotherapy), (6) adherence/fidelity checks, and (7) treatment manual or structured guide. The impact of these 7 criteria on treatment effects was assessed using 3-level mixed models with Bonferroni adjustment ($p < .007$) to address the risk of chance findings. None of the risk of bias criteria had a significant impact on ES, except for the blinding criterion [$t(553) = -5.38$; $p < .001$] which showed a significantly lower mean ES when subjects were blind to assessment [$g = 0.09$, $t(553) = 1.35$, $p = .179$] compared to when this criterion was not met [$g = 0.34$, $t(553) = 7.09$, $p < .001$].

RESULTS

Study Pool

Our search yielded 55 RCTs (46 published trials and 9 dissertations) from 1984 to 2017 that met the inclusion criteria (Figure 1), generating 389 ESs. The studies included 4,560 participants (mean number to compute ES = 77.76; SD = 83.83). Mean age was 13.87 years (SD = 2.52), mean percentage of male participants was 44.00 (SD =17.11), and 49.1% of the study samples were majority white. On average, the treatment protocols specified 14.13 number of sessions (SD = 5.35); mean number of sessions planned in treatment was not significantly related to posttreatment ES [$t(172) = 0.49$, $p = .624$]. Table 1 provides detailed information about each study.

Overall Posttreatment Effect

The 53 studies reporting depression outcomes at posttreatment produced 222 dependent ESs. Mean posttreatment ES was 0.36 [95% CI = 0.25–0.47, $t(221) = 6.21$, $p < .001$]. Between-study variance ($\sigma_v^2 = 0.121$, $\chi^2(1) = 44.0$, $p < .001$) and within-study variance ($\sigma_u^2 = 0.050$, $\chi^2(1) = 53.2$, $p < .001$) were significant, with mean observed sampling (residual) variance of 0.136. Of the total variance, 39.4% was attributable to between-study differences and 16.2% to within-study differences.

For comparison to the findings of the 2006 meta-analysis,[16] we compared the overall effect size for studies through 2004 (the endpoint for studies included in the 2006 meta-analysis) and studies after 2004. The 28 studies through 2004, reporting depression outcomes at posttreatment, produced 147 dependent ESs, and the mean posttreatment ES was 0.39 [95% CI = 0.23–0.54, $t(220) = 4.86$, $p < .001$]. The 25 studies after 2004, reporting depression outcomes at posttreatment, yielded 75 dependent ESs with a mean posttreatment ES of 0.32 [95% CI = 0.16–0.49, $t(220) = 3.84$, $p < .001$]. There was no significant difference between the 2 study pools [$t(220) = -0.53$, $p = .599$].

Moderators of Posttreatment Effectiveness

As indicated in Table 2, we tested eight study-level moderators and four ES-level moderators. Three of the 12

TABLE 2 Results of Moderator Analyses Based on Three-level Mixed Effects Models of 222 Dependent Effect Sizes From 55 Studies at Posttreatment

			Subgroup Analysis			Moderator Test	
Moderator	**κ of Studies**	**No. of ESs**	**ES (*g*)**	**95% CI**		**Test Statistic**	***p***
Study-Level Moderators (Third Level)							
Year of Study	**53**	**222**				$t(220) = -1.80$	.074
Study location	**53**	**222**				$t(220) = 0.29$	.772
North America	39	188	0.35***	0.22	0.48		
Outside North America	14	34	0.39**	0.15	0.62		
Participant engagement	**50**	**207**				$t(205) = 0.58$	.564
Recruited	40	177	0.34***	0.21	0.48		
Referred	10	30	0.44**	0.16	0.72		
Ethnicity[a]	**40**	**172**				$t(170) = 1.16$	.247
White sample (≥50% white)	27	129	0.30***	0.15	0.45		
Nonwhite sample (<50% white)	13	43	0.46***	0.23	0.69		
Sex[a]	**53**	**222**				$t(220) = 0.85$	.397
Majority male participants (≥50% male participants)	16	49	0.28*	0.06	0.50		
Majority female participants (>50% Female participants)	37	173	0.39***	0.26	0.52		
Developmental period	**52**	**216**				$t(214) = 1.23$	.219
Childhood (mean age <13 y)	18	64	0.25*	0.05	0.46		
Adolescence (mean age ≥13 y)	34	152	0.41***	0.27	0.55		
Diagnosis requirement	**34**	**147**				$t(145) = -0.39$	.700
Required of all participants	18	97	0.41***	0.23	0.59		
Not required	16	50	0.35**	0.15	0.56		
Treatment Setting	**38**	**134**				$t(132) = -0.62$	.539
Clinical	8	31	0.33*	0.03	0.63		
Nonclinical	30	103	0.44***	0.26	0.61		
Effect Size−Level Moderators (Second Level)							
Informant	**53**	**221**				$F_{2,218} = 13.67$	<.001
Youth[b]	51	158	0.39***	0.28	0.51		
Parent[b,c]	12	25	-0.06	-0.26	0.14		
Other[c]	16	38	0.46***	0.29	0.63		
Treatment format	**52**	**219**				$F_{2,200} = 1.61$	.200
Individual	16	68	0.41***	0.21	0.60		
Group	29	107	0.34***	0.20	0.49		
Mixed individual and group	7	28	0.56***	0.31	0.80		
Treatment type[d]	**45**	**193**				$F_{2,174} = 3.59$	.030
CBT[b]	34	137	0.31***	0.18	0.44		
IPT-A[b]	5	14	0.78***	0.43	1.13		
CBT and additional treatment	7	26	0.45***	0.22	0.69		
Control condition	**52**	**215**				$F_{2,212} = 3.91$	.022
No treatment/waitlist[b]	28	114	0.49***	0.34	0.64		

(continued)

moderators explained a significant amount of variation in treatment benefit at posttreatment (Figure 2).

Informant. There was a significant difference in mean posttreatment effect for the informant moderator. Pairwise comparisons revealed significantly smaller mean effects for parent reports versus youth self-reports [$t(218) = 4.81$, $p < .001$] and "other reports" including siblings, peers, teacher, and/or therapist as the informant [$t(218) = 4.87$, $p < .001$]. The difference between youth self- and other reports was not significant [$t(218) = 0.85$, $p = .394$]. The informant moderator explained 22.5% of the within-study variance in ES.

Treatment Type. Mean treatment effect size relative to control conditions differed significantly according to the type of treatment used. Pairwise comparisons revealed a

TABLE 2 Continued

Moderator	κ of Studies	No. of ESs	Subgroup Analysis ES (*g*)	95% CI		Moderator Test Test Statistic	*p*
Psychotherapy and medication placebo[b]	12	47	0.16	-0.05	0.37		
Usual care treatment	15	54	0.29**	0.10	0.49		

Note: *Boldface type indicates moderators. Some moderators were missing for certain studies. Each study can contribute multiple effect sizes; thus study sample size across subgroups can exceed the total study sample size for the ES-level moderators. CBT = cognitive-behavioral therapy; ES = effect size; g = Hedges g; IPT = Interpersonal Psychotherapy for Adolescents.*

[a]*We also tested ethnicity (percent nonwhite) and sex (percent female participants) as continuous variables. Both tests were nonsignificant [ethnicity* $t(170) = -0.47$, $p = .638$; *sex* $t(220) = -0.44$, $p = .664$*].*

[b,c]*Within each moderator having more than two subgroups, identical superscript letters indicate significant (*$p < .05$*) pairwise comparisons between subgroups.*

[d]*The other behavioral and nonbehavioral treatment categories were excluded from the moderator analysis because of the limited number of studies.*

*$p < .05$; **$p < .01$; ***$p < .001$.

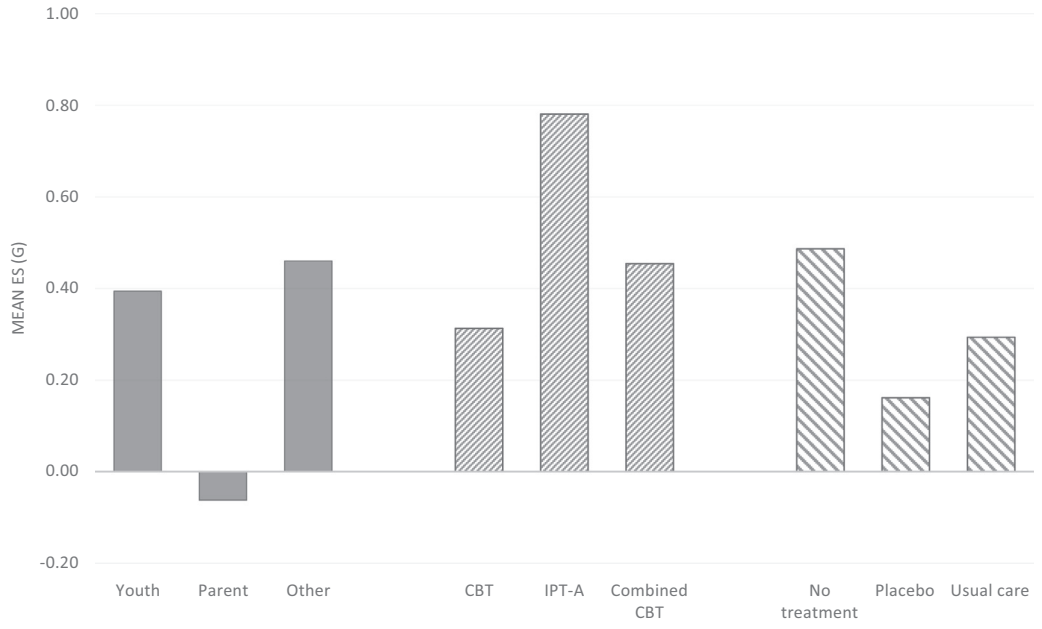
significantly larger mean effect relative to control conditions for IPT-A than for CBT [means 0.78 versus 0.31, $t(174) = 2.48$, $p = .014$]. The mean effect size relative to control conditions for CBT combined with another treatment versus CBT only [$t(174) = 1.19$, $p = .236$] or IPT-A [$t(174) = 1.52$, $p = .131$] was not significantly different. The treatment type moderator explained 12.8% of the between-study variance and 1.8% of the within-study variance in ES.

Because IPT-A is used only with adolescents, we carried out a secondary analysis that included only the CBT studies with majority adolescent samples (κ = 26). That comparison yielded a similar finding with a significantly larger mean effect compared to control conditions for IPT-A ($g = 0.79$) than for CBT [$g = 0.34$; $t(139) = 2.18$, $p = .031$].

Control Condition

Analyses revealed a significant difference in mean posttreatment effect for the control condition moderator. Pairwise comparisons revealed a significant larger mean effect for the waitlist/no-treatment condition compared to placebo control

FIGURE 2 Effect Sizes Associated With the Levels of 3 Moderators



Note: *(a) Informant: youth, parent, and other; (b) Treatment type: cognitive-behavioral therapy (CBT), Interpersonal Psychotherapy for Adolescents (IPT-A), and cognitive-behavioral therapy (CBT) combined with additional treatment (combined CBT); and (c) Control condition: no treatment/waitlist, psychotherapy or medication placebo, and usual care treatment.*

groups [ie, conditions designed to control for nonspecific factors such as receiving attention and expecting benefit[110]; $t(212) = 2.69$, $p = .008$]. The usual care condition did not differ significantly from placebo [$t(218) = 1.52$, $p = .131$] or waitlist/no-treatment [$t(218) = 0.90$, $p = .369$]. The control condition moderator explained 13.4% of the between-study variance and 2.1% of the within-study variance in ES. It should be noted that a sensitivity analysis revealed that findings regarding the placebo condition were virtually identical with pill placebo included[111] versus excluded.

Controlling for Potential Confounding. We also examined whether controlling for other moderators altered the effect of each primary moderator, to address potential confounding among moderators. A model including informant, treatment format, control condition, participant engagement, sex, and developmental period yielded similar findings for informant ($F_{2,175} = 15.36$, $p < .001$), treatment format ($F_{2,175} = 1.96$, $p = .14$), and control condition ($F_{2,175} = 3.08$, $p = .05$), although the number of studies with data for all moderators was somewhat smaller ($\kappa = 45$). In contrast, controlling for treatment format, informant, control condition, participant engagement, gender, and developmental period rendered the treatment type effect nonsignificant ($F_{2,131} = 1.05$, $p = .35$). It should be noted that the pairwise differences were in the same direction, yet less pronounced, and based on a much smaller ($\kappa = 36$) number of studies with data on each moderator.

Durability of Psychotherapy Effects

The 32 studies reporting follow-up assessments generated 167 dependent ESs. Mean follow-up ES was 0.23 [95% CI = 0.09–0.36, $t(166 = 3.25$, $p = .001$]. Between-study variance [$\sigma_v^2 = 0.119$, $\chi^2(1) = 47.9$, $p < .001$] and within-study variance [$\sigma_v^2 = 0.017$, $\chi^2(1) = 12.7$, $p < .001$] were significant, with residual (sampling) variance of 0.089. Some 52.9% of the total variance was attributable to between-study differences, and 7.6% to within-study differences. This mean follow-up ES was significantly lower than the mean posttreatment ES [$t(387) = 2.01$, $p = .045$].

A total of 30 studies reported both posttreatment and follow-up assessments generating 274 dependent ESs. The time lag between post-treatment and follow-up averaged 41.98 weeks (SD = 36.11), although it should be noted that this information was available for only 18 studies. Mean posttreatment ES was 0.28 [95% CI = 0.15–0.41, $t(272 = 4.21$, $p < .001$], whereas mean follow-up ES was 0.21 [95% CI = 0.08–0.34, $t(272 = 3.20$, $p = .002$]. This mean follow-up ES was not significantly different from the mean posttreatment ES of 0.28 reported for the same study pool [$t(272) = 1.87$, $p = .062$].

Generalizability of Posttreatment Effectiveness

Treatment of depression produced benevolent effects on measures of anxiety [$g = 0.22$, 95% CI = 0.07–0.36, $t(309) = 2.84$, $p < .005$] and externalizing behavior [ie, conduct and ADHD symptoms; $g = 0.14$, 95% CI = −0.02 to 0.30, $t(309) = 1.69$, $p < .092$], suggesting some generalization to other mental health conditions. When we fitted a 3-level mixed-effects model across all outcomes at posttreatment, we found a significant difference in psychotherapy effects according to the outcome type ($F_{3,2309} = 5.97$, $p = .003$). Pairwise comparisons revealed significantly larger effects for depression outcomes versus anxiety [$t(554) = 3.18$, $p = .001$] and externalizing outcomes [$t(554) = 2.96$, $p = .003$]. The difference in mean treatment effect between anxiety and externalizing outcomes was not significant [$t(554) = 0.16$, $p = .872$].

DISCUSSION

This meta-analysis includes the most comprehensive collection to date of peer-reviewed published and dissertation youth depression RCTs. Despite the 63% increase in the number of studies over a similarly structured 2006 youth depression meta-analysis,[16] from 35 to 55 studies, the magnitude of the psychotherapy effect size remained essentially unchanged, at 0.36, versus 0.34 in 2006. Also, there was no significant difference in ES between studies through 2004 and studies after 2004, suggesting no significant change in effectiveness since the previous synthesis. The mean ES of 0.36 falls midway between Cohen benchmarks for small and medium effects,[101] translating to a probability of 60% that a randomly selected youth receiving psychotherapy would be better off after treatment than a randomly selected youth in a control condition.[17] These findings suggest that (1) the posttreatment effectiveness of psychotherapy for youth depression has remained relatively unchanged over more than a dozen years, and (2) marked improvement is needed.

In comparison to the posttreatment mean of 0.36, we found a significantly smaller mean ES of 0.21 in follow-up assessments averaging about 42 weeks after posttreatment; the post versus follow-up difference in ES was highly significant. When we confined our analysis to the 30 studies with assessments at both immediate posttreatment and later follow-up (as in the 2006 depression meta-analysis[16]), the mean follow-up effect ($g = 0.21$) was not significantly different from the mean effect after treatment completion ($g = 0.28$), but was slightly smaller than the follow-up ES of

0.28 reported in the 2006 meta-analysis.[16] This suggests a need to focus treatment research not only on symptom reduction during treatment but also on prevention of relapse and resurgence of symptoms.[112] Current approaches to targeting relapse and recurrence of symptoms include, for example, continued practice of learned skills after treatment ends, scheduling of booster sessions, and planning for future check-ups (the "dental model").[113] Our findings suggest that more may be needed to preserve gains and to prevent slippage.

We found evidence that depression treatment can produce benevolent effects on nondepression symptoms, but we also found that effects for depression outcomes were significantly larger than effects for anxiety and externalizing behavior outcomes. So, our findings point to specificity, combined with some degree of generalization to conditions other than depression. This could be viewed as good news, particularly in light of evidence showing co-occurrence of depression with other youth disorders and problems, especially anxiety.[114,115] Our findings suggest that depression treatment alone may help address such co-occurrence, but larger effects are needed to have a genuine impact on comorbidities. Indeed, the recent emergence of transdiagnostic treatments designed to encompass depression and other disorders and problems that often co-occur reflects efforts to expand treatment benefit beyond what is produced by single disorder therapies.[116,117] For example, the Modular Approach to Therapy for Children with Anxiety, Depression, Trauma, or Conduct Problems (MATCH)[116] uses treatment modules designed to address symptoms of four diagnostic clusters, and includes a decision flowchart that guides module selection and sequence. In studies with clinically referred youths, MATCH has been found to outperform standard manualized treatments[118] and treatment by community clinicians who had years of training in standard manualized treatments.[119]

Our look at study characteristics revealed differences in treatment effects depending on the type of psychotherapy and control conditions used. Both CBT and IPT-A produced significant effects, significantly smaller for CBT than for IPT-A. CBT, with the largest number of studies in our meta-analysis ($\kappa = 34$), is the most widely disseminated evidence-based psychotherapy for youth depression. It focuses in part on modifying depressive cognitions and increasing engagement in rewarding activities.[120] IPT-A, a newer evidence-based psychotherapy, with only 5 studies in our meta-analysis ($\kappa = 5$), focuses in part on teaching communication and interpersonal skills that are needed to increase affiliation, to develop close attachment relationships, and to manage interpersonal stressors related to depression.[121] Our findings provide the strongest support for IPT-A; however, additional studies are needed. Controlling for potential confounders, we found that the CBT−IPT-A difference shrank and was no longer significant, perhaps due in part to reduced power.

We also found that treatments produced larger effects when compared with inert control conditions (ie, no treatment and waitlist), which was perhaps not surprising, considering that this type of comparison controls only for the passage of time. Passive control conditions were in fact the most common form of control group across the studies ($\kappa = 28$). When compared with psychotherapy and medication placebo, the least-often used type of control group ($\kappa = 12$), the psychotherapy effects were small and not significant. When compared with usual care ($\kappa = 15$), psychotherapy effects were significant but smaller ($g = 0.29$), which was somewhat encouraging, given that usual care is typically an active intervention intended to provide genuine benefit. That said, the nature and dose of usual care can vary so widely across youths, therapists, and studies that better documentation is needed in future research for proper interpretation of such comparisons (see also Spielmans *et al.* [2010] and Weisz *et al.* [2013]).[89,122]

An additional moderator of treatment effectiveness was the assessment informant. The effect size for youth self-reports was positive and significantly larger than that for parent-reports, which was nonsignificant: parent reports showed no posttreatment outcome difference between treatment groups and control groups. One possible explanation for this distressing finding may be that parental depression, often associated with youth depression[123] and thus potentially present in the studies that we included, has been linked to poorer intervention outcomes.[124-126] Whatever the reason, it certainly cannot be regarded as good news that, according to parents' reports, the psychotherapy that their children received for depression produced no measurable benefit. On a more positive note, larger effects were produced by reports from "other" informants, including a number of well-established measures such as the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS), Hamilton Depression Rating Scale, Bellevue Index of Depression, and Children's Depression Rating Scale—measures regarded by many as valid indicators of depression treatment response. Overall, these findings illustrate the need for a multi-informant assessment approach[22] to convey how the level of treatment benefit varies depending on the eye of the beholder.

Consistent with the Weisz *et al.* (2006) meta-analysis,[16] there was no significant effect for treatment format,

suggesting that the benefits of psychotherapy, although modest, are similar across individual, group, and mixed formats. Interestingly, there was no significant effect of any of the study-level moderators, including ethnicity, sex, developmental period, study location, recruitment, requirement of diagnosis, treatment setting, and year of study. These tests were reasonably powered (ethnicity $\kappa = 40$, diagnosis requirement $\kappa = 34$, and treatment setting $\kappa = 38$), suggesting that the psychotherapies tested to date may work relatively similarly across the population variations, but even more amply powered tests would strengthen confidence in that conclusion.

Our overall finding that psychotherapy for youth depression shows modest effects that do not differ markedly across different population groups may be explained in part by the heterogeneity of mechanisms underlying youth depression. Different mechanisms may require different psychotherapy approaches[127,128] instead of a single psychotherapy protocol for all youths who are depressed. The different mechanisms, and clinical presentations, of youth depression might require personalized treatment in which the psychotherapy is matched to the individual youth. Consider, for example, the two psychotherapies that produced significant effects in this meta-analysis, CBT and IPT-A. For some youths, targeting behavioral engagement and cognitions using CBT may be particularly effective; for others, targeting interpersonal skills using IPT-A may work better. Personalized treatment is the cornerstone of the NIH Precision Medicine Initiative, central to the NIMH Strategic Plan,[13] and a key component of the Institute of Medicine (IOM) report on building successful treatments.[129] Our capacity to do such treatment tailoring can be strengthened by an expanding database of RCTs synthesized through meta-analysis, with particular attention to interactions among patient characteristics and treatment approaches.

Certain limitations of our meta-analysis suggest possible directions for future research. First, as previously noted, only about half of the studies included follow-up assessments. Considering the need to prevent relapse and return of depressive symptoms,[112] future studies should include follow-up assessments wherever possible. Second, we found evidence of publication bias in follow-up studies; this may suggest a need for researchers to publish results even when they do not support the psychotherapy being tested, as well as attention by editors and editorial boards to processes of journal review that could bias the follow-up manuscript review process, perhaps tilting toward acceptance of encouraging findings and rejection of disappointing findings. We also found that effect size was lower when subjects were blind to assessment ($g = 0.09$ versus 0.34), suggesting the need for attention to methodological rigor in primary studies. Third, inclusion of only English-language studies may have limited generalizability of the findings across national boundaries. Fourth, only eight studies in our collection included measures of suicidality, of which two used imputed effect size because the studies reported only a nonsignificant effect; therefore, any synthesis focusing on assessing treatment effects on suicidality would be unreliable. Considering that suicidality is the second leading cause of adolescent deaths in the United States[130] and that depression increases risk for suicide,[131] future RCTs of youth depression treatments should include measures of suicidality. Fifth, studies often did not include systematic assessment or reporting of comorbid disorders, preventing us from fairly testing comorbidity as a moderator. Sixth, we coded psychotherapies as active treatments based on the authors' definitions; however, for numerous interventions, limited reporting of treatment procedures and rationale in the articles ruled out reliable coding of whether the therapies should be regarded as bona fide according to the criteria laid out by Wampold *et al.* (1997).[132] Also, we used coding procedures that have been developed over many years of review of the psychotherapy literature, adhering to the codes and procedures used in the 2006 meta-analysis to ensure fair comparison, while assessing and reporting intercoder agreement. However, we recognize that coding psychotherapy into theoretical categories, like CBT,[133] can be tricky and may affect the results of a meta-analysis. Seventh, we were unable to reliably fit a mixed-effects model that included all moderators simultaneously because of missingness across moderators that substantially decreased the study pool. As holds for virtually any meta-analysis, we therefore may not have controlled for all relevant confounders. Eighth, the analyses were based mainly on observed cases/completers and not intent-to-treat; therefore, the results may generalize only to treatment completers. Finally, we noted that substantial variance in outcome remained unexplained by our analyses. This suggests that important moderators—including some that could not be coded reliably (eg, researcher allegiance)—may have gone undetected; identifying what those moderators may be remains a key challenge for future research.

In conclusion, this meta-analysis of youth depression psychotherapy RCTs, following up on a similarly structured meta-analysis published 13 years ago,[16] with a substantial increase in the number of studies included, showed a strikingly similar pattern of findings, with a similarly modest level of treatment benefit. The findings suggest a persistent and profound need to strengthen the immediate and longer-term benefits of psychotherapies for youths who are depressed. This, in turn, highlights the need for creativity by treatment developers who specialize in youth interventions.

A puzzling, and perhaps limiting, characteristic of the only two empirically supported treatments for youth depression, to date, is that both—CBT and IPT-A—are essentially junior versions of treatments originally developed for adults. Surely it is possible to develop treatment approaches that begin with a focus on children and adolescents, focusing on the distinctive ways in which they experience and recover from depression, and to build interventions accordingly. Such an approach, combined with an emphasis on strategies for personalizing, might make a genuine difference in the level of benefit afforded by treatment. Whatever the approach, a critical goal suggested by our findings is that the next meta-analysis of youth psychotherapy effects will bring the kind of good news we that would all like to share with troubled youths and their families.

Accepted April 12, 2019.

Dr. Eckshtain is with Massachusetts General Hospital, Harvard Medical School, Boston. Dr. Kuppens is with KU Leuven, Belgium, and Karel de Grote University College, Antwerp, Belgium. Dr. Ugueto is with the McGovern Medical School, The University of Texas Health Science Center at Houston. Dr. Ng is with Florida International University, Miami. Dr. Vaughn-Coaxum is with University of Pittsburgh, PA. Ms. Corteselli and Dr. Weisz are with Harvard University, Cambridge, MA.

Dr. Weisz was funded by the Norlien Foundation.

This article is part of a special series devoted to the subject of depression, the presidential initiative of AACAP President Karen Dineen Wagner, MD, PhD. The series covers current topics in depression, including programs that have initiated depression screening for youth, processes by which youth who screen positive for depression receive treatment, and the identification and treatment of depression in primary care settings. The series was edited by Guest Editor Laura Richardson, MD, MPH, and Deputy Editor Elizabeth McCauley, PhD, ABPP.

The authors wish to thank Elizabeth R. Wolock, BA, Research Assistant from the Laboratory for Youth Mental Health in the Department of Psychology at Harvard University, and Morgan Westine, BS, Volunteer from Simmons College, for assistance in manuscript preparation.

Disclosure: Dr. Eckshtain has received funding from the National Institute of Mental Health. Dr. Weisz has received funding from the Institute of Education Sciences, the Norlien Foundation, and the Connecticut Health and Development Institute. Drs. Kuppens, Ugueto, Ng, Vaughn-Coaxum, and Ms. Corteselli report no biomedical financial interests or potential conflicts of interest.

Correspondence to Dikla Eckshtain, PhD, Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, 55 Fruit Street, Boston, MA 02114; e-mail: deckshtain@mgh.harvard.edu

https://doi.org/10.1016/j.jaac.2019.04.002

REFERENCES

1. Collins PY, Patel V, Joestl SS, March D, Insel TR, Daar AS. Grand challenges in global mental health. Nature. 2011;475:7-30.
2. Hankin BL, Young JF, Abela JRZ, *et al*. Depression from childhood into late adolescence: influence of gender, development, genetic susceptibility, and peer stress. J Abnorm Psychol. 2015;124:803-816.
3. Merikangas KR, He J, Burstein M, *et al*. Lifetime prevalence of mental disorders in US adolescents: results from the National Comorbidity Survey Replication—Adolescent Supplement (NCS-A). J Am Acad Child Adolesc Psychiatry. 2010;49:980-989.
4. Birmaher B, Arbelaez C, Brent D. Course and outcome of child and adolescent major depressive disorder. Child Adolesc Psychiatr Clin N Am. 2002;11:619-638.
5. Kovacs M. Presentation and course of major depressive disorder during childhood and later years of the life span. J Am Acad Child Adolesc Psychiatry. 1996;35:705-715.
6. Shaffer D, Pfeffer CR, Bernet W, Arnold V, Beitchman J, Benson RS. Practice parameter for the assessment and treatment of children and adolescents with suicidal behavior. J Am Acad Child Adolesc Psychiatry. 2001;40:24S-51S.
7. Brent DA. Pediatric depression. In: Mann JJ, McGrath PJ, Roose SP, eds. Clinical Handbook for the Management of Mood Disorders. New York, NY: Cambridge University Press; 2013:146-164.
8. Avenevoli SKE, Kessler RC, Merikangas KR. Epidemiology of depression in children and adolescents. In: Abela JRZ, Hankin BL, eds. Handbook of Depression in Children and Adolescents. New York, NY: Guilford Press; 2008:6-32.
9. Glied S, Pine DS. Consequences and correlates of adolescent depression. Arch Pediatr Adolesc Med. 2002;156:1009-1014.
10. Lewinsohn PM, Roberts RE, Seeley JR, Rohde P, Gotlib IH, Hops H. Adolescent psychopathology II: psychosocial risk factors for depression. J Abnorm Psychol. 1994;103:302-315.
11. Lewinsohn PM, Rohde P, Klein DN, Seeley JR. Natural course of adolescent major depressive disorder: I. Continuity into young adulthood. J Am Acad Child Adolesc Psychiatry. 1999;38:56-63.
12. Weissman MM, Wolk S, Goldstein RB, *et al*. Depressed adolescents grown up. JAMA. 1999;281:1707-1713.
13. U.S. Department of Health and Human Services, National Institute of Mental Health. NIMH Strategic Plan for Research. 2015. Available at: http://www.nimh.nih.gov/about/strategic-planning-reports/index.shtml. Accessed October 11, 2018.
14. U.S. Department of Health and Human Services, National Institute of Mental Health. NIMH Strategic Plan for Research. 2008. Available at: http://www.nimh.nih.gov/about/strategic-planning-reports/index.shtml. Accessed October 11, 2018.
15. American Academy of Child and Adolescent Psychiatry (AACAP). Practice parameter for the assessment and treatment of children and adolescents with depressive disorders. J Am Acad Child Adolesc Psychiatry. 2007;46:1503-1526.
16. Weisz JR, McCarty CA, Valeri SM. Effects of psychotherapy for depression in children and adolescents: a meta-analysis. Psychol Bull. 2006;132:132-149.
17. McGraw KO, Wong SP. A common language effect size statistic. Psychol Bull. 1992;111:361-365.
18. Zhou X, Hetrick SE, Cuijpers P, *et al*. Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: a systematic review and network meta-analysis. World Psychiatry. 2015;14:207-222.
19. Weisz JR, Kuppens S, Ng MY, *et al*. What five decades of research tells us about the effects of youth psychological therapy: a multilevel meta-analysis and implications for science and practice. Am Psychol. 2017;72:79-117.
20. McLeod BD, Weisz JR. Using dissertations to examine potential bias in child and adolescent clinical trials. J Consult Clin Psychol. 2004;72:235-251.
21. Atkins D, Fink K, Slutsky J. Better information for better health care: the Evidence-based Practice Center program and the Agency for Healthcare Research and Quality. Ann Intern Med. 2005;142:1035-1041.
22. De Los Reyes A, Augenstein TM, Wang M, *et al*. The validity of the multi-informant approach to assessing child and adolescent mental health. Psychol Bull. 2015;141:858-900.
23. Ackerson J, Scogin F, McKendree-Smith N, Lyman RD. Cognitive bibliotherapy for mild and moderate adolescent depressive symptomatology. J Consult Clin Psychol. 1998;66:685-690.
24. Asarnow JR, Jaycox LH, Duan N, *et al*. Effectiveness of a quality improvement intervention for adolescent depression in primary care clinics: a randomized controlled trial. JAMA. 2005;293:311-319.
25. Asarnow JR, Jaycox LH, Tang L, *et al*. Long-term benefits of short-term quality improvement interventions for depressed youths in primary care. Am J Psychiatry. 2009;166:1002-1010.
26. Asarnow JR, Scott CV, Mintz J. A combined cognitive-behavioral family education intervention for depression in children: a treatment development study. Cognit Ther Res. 2002;26:221-229.

27. Bolton P, Bass J, Betancourt T, *et al*. Interventions for depression symptoms among adolescent survivors of war and displacement in Northern Uganda: a randomized controlled trial. JAMA. 2007;298:519-527.
28. Brent DA, Holder D, Kolko D, *et al*. A clinical psychotherapy trial for adolescent depression comparing cognitive, family, and supportive therapy. Arch Gen Psychiatry. 1997;54:877-885.
29. Brent DA, Kolko DJ, Birmaher B, *et al*. Predictors of treatment efficacy in a clinical trial of three psychosocial treatments for adolescent depression. J Am Acad Child Adolesc Psychiatry. 1998;37:906-914.
30. Charkhandeh M, Talib MA, Hunt CJ. The clinical effectiveness of cognitive behavior therapy and an alternative medicine approach in reducing symptoms of depression in adolescents. Psychiatry Res. 2016;239:325-330.
31. Clarke GN, Hawkins W, Murphy M, Sheeber LB, Lewinsohn PM, Seeley JR. Targeted prevention of unipolar depressive disorder in an at-risk sample of high school adolescents: a randomized trial of group cognitive intervention. J Am Acad Child Adolesc Psychiatry. 1995;34:312-321.
32. Clarke GN, Hornbrook M, Lynch F, *et al*. A randomized trial of a group cognitive intervention for preventing depression in adolescent offspring of depressed parents. Arch Gen Psychiatry. 2001;58:1127-1134.
33. Clarke GN, Hornbrook M, Lynch F, *et al*. Group cognitive-behavioral treatment for depressed adolescent offspring of depressed parents in a health maintenance organization. J Am Acad Child Adolesc Psychiatry. 2002;41:305-313.
34. Clarke GN, Rohde P, Lewinsohn PM, Hops H, Seeley JR. Cognitive-behavioral treatment of adolescent depression: efficacy of acute group treatment and booster sessions. J Am Acad Child Adolesc Psychiatry. 1999;38:272-279.
35. Curtis SE. Cognitive-behavioral treatment of adolescent depression: effects on multiple parameters [dissertation]. Logan, Utah: Utah State University; 1992.
36. Dana EC. A cognitive-behavioral intervention for conduct-disordered and concurrently conduct-disordered and depressed children [dissertation]. Garden City, NY: Adelphi University; 1998.
37. De Cuyper S, Timbremont B, Braet C, De Backer V, Wullaert T. Treating depressive symptoms in schoolchildren: a pilot study. Eur Child Adolesc Psychiatry. 2004;13:105-114.
38. Diamond GS, Reis BF, Diamond GM, Siqueland L, Isaacs L. Attachment-based family therapy for depressed adolescents: a treatment development study. J Am Acad Child Adolesc Psychiatry. 2002;41:1190-1196.
39. Doerr ER. Anticipation training as a counseling model with depressed adolescents [dissertation]. St. Louis, MO: Saint Louis University; 1984.
40. Ettelson RG. The treatment of adolescent depression [dissertation]. Normal, IL: Illinois State University; 2002.
41. Fischer SA. Development and evaluation of group cognitive-behavioral therapy for depressed and suicidal adolescents in juvenile detention [dissertation]. Eugene, OR: University of Oregon; 1995.
42. Fleming T, Dixon R, Frampton C, Merry S. A pragmatic randomized controlled trial of computerized CBT (SPARX) for symptoms of depression among adolescents excluded from mainstream education. Behav Cogn Psychother. 2012;40:529-541.
43. Garber J, Clarke GN, Weersing VR, *et al*. Prevention of depression in at-risk adolescents: a randomized controlled trial. JAMA. 2009;301:2215-2224.
44. Brent DA, Brunwasser SM, Hollon SD, *et al*. Effect of a cognitive-behavioral prevention program on depression 6 years after implementation among at-risk adolescents: a randomized clinical trial. JAMA Psychiatry. 2015;72:1110-1118.
45. Gillham JE, Hamilton J, Freres DR, Patton K, Gallop R. Preventing depression among early adolescents in the primary care setting: a randomized controlled study of the Penn Resiliency Program. J Abnorm Child Psychol. 2006;34:203-219.
46. Gillham JE, Reivich KJ, Brunwasser SM, *et al*. Evaluation of a group cognitive-behavioral depression prevention program for young adolescents: a randomized effectiveness trial. J Clin Child Adolesc Psychol. 2012;41:621-639.
47. Gillham JE, Reivich KJ, Freres DR, *et al*. School-based prevention of depressive symptoms: a randomized controlled study of the effectiveness and specificity of the Penn Resiliency Program. J Consult Clin Psychol. 2007;75:9-19.
48. Hickman KA. Effects of social skills training on depressed children attending a behavioral day treatment program [dissertation]. Hempstead, NY: Hofstra University; 1994.
49. Israel P, Diamond GS. Feasibility of attachment based family therapy for depressed clinic-referred Norwegian adolescents. Clin Child Psychol Psychiatry. 2013;18:334-350.
50. Jeong Y, Hong S, Lee MS, Park M, Kim Y, Suh C. Dance movement therapy improves emotional responses and modulates neurohormones in adolescents with mild depression. Int J Neurosci. 2005;115:1711-1720.
51. Kahn RC. The effect of a group support intervention program on depression, social adjustment, and self-esteem of adolescents in an overseas American international school [dissertation]. Washington, DC: Catholic University of America; 1989.
52. Kahn JS, Kehle TJ, Jenson WR, Clark E. Comparison of cognitive-behavioral, relaxation, and self-modeling interventions for depression among middle-school students. Sch Psychol Rev. 1990;19:196-211.
53. Lewinsohn PM, Clarke GN, Hops H, Andrews JA. Cognitive-behavioral treatment for depressed adolescents. Behav Ther. 1990;21:385-401.
54. Liddle B, Spence SH. Cognitive-behaviour therapy with depressed primary school children: a cautionary note. Behav Psychother. 1990;18:85-102.
55. Listug-Lunde L, Vogeltanz-Holm N, Collins J. A cognitive-behavioral treatment for depression in rural American Indian middle school students. Am Indian Alsk Native Ment Health Res. 2013;20:16-34.
56. Luby J, Lenze S, Tillman R. A novel early intervention for preschool depression: findings from a pilot randomized controlled trial. J Child Psychol Psychiatry. 2012;53:313-322.
57. McCarty CA, Violette HD, Duong MT, Cruz RA, McCauley E. A randomized trial of the positive thoughts and action program for depression among early adolescents. J Clin Child Adolesc Psychol. 2013;42:554-563.
58. McLaughlin CL. Evaluating the effect of an empirically-supported group intervention for students at-risk for depression in a rural school district [dissertation]. Kent, OH: Kent State University; 2010.
59. Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MFG. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. BMJ. 2012;344:1-16.
60. Moldenhauer Z. Adolescent depression: a primary care pilot intervention study [dissertation]. Rochester, NY: University of Rochester; 2004.
61. Mufson L, Weissman MM, Moreau D, Garfinkel R. Efficacy of interpersonal psychotherapy for depressed adolescents. Arch Gen Psychiatry. 1999;56:573-579.
62. Reynolds WM, Coats KI. A comparison of cognitive-behavioral therapy and relaxation training for the treatment of depression in adolescents. J Consult Clin Psychol. 1986;54:653-660.
63. Rohde P, Clarke GN, Mace DE, Jorgensen JS, Seeley JR. An efficacy/effectiveness study of cognitive-behavioral treatment for adolescents with comorbid major depression and conduct disorder. J Am Acad Child Adolesc Psychiatry. 2004;43:660-668.
64. Rohde P, Stice E, Shaw H, Brière FN. Indicated cognitive behavioral group depression prevention compared to bibliotherapy and brochure control: acute effects of an effectiveness trial with adolescents. J Consult Clin Psychol. 2014;82:65-74.
65. Rohde P, Stice E, Shaw H, Gau JM. Effectiveness trial of an indicated cognitive–behavioral group adolescent depression prevention program versus bibliotherapy and brochure control at 1- and 2-year follow-up. J Consult Clin Psychol. 2015;83:736-747.
66. Rosselló J, Bernal G. The efficacy of cognitive-behavioral and interpersonal treatments for depression in Puerto Rican adolescents. J Consult Clin Psychol. 1999;67:734-745.
67. Sanford M, Boyle M, McCleary L, *et al*. A pilot study of adjunctive family psychoeducation in adolescent major depression: feasibility and treatment effect. J Am Acad Child Adolesc Psychiatry. 2006;45:386-395.
68. Santomauro D, Sheffield J, Sofronoff K. Depression in adolescents with ASD: a pilot RCT of a group intervention. J Autism Dev Disord. 2016;46:572-588.
69. Stark KD, Reynolds WM, Kaslow NJ. A comparison of the relative efficacy of self-control therapy and a behavioral problem-solving therapy for depression in children. J Abnorm Child Psychol. 1987;15:91-113.
70. Stasiak K, Hatcher S, Frampton C, Merry SN. A pilot double blind randomized placebo controlled trial of a prototype computer-based cognitive behavioural therapy program for adolescents with symptoms of depression. Behav Cogn Psychother. 2014;42:385-401.
71. Stice E, Burton E, Bearman SK, Rohde P. Randomized trial of a brief depression prevention program: an elusive search for a psychosocial placebo control condition. Behav Res Ther. 2007;45:863-876.
72. Stice E, Rohde P, Seeley JR, Gau JM. Brief cognitive-behavioral depression prevention program for high-risk adolescents outperforms two alternative interventions: a randomized efficacy trial. J Consult Clin Psychol. 2008;76:595-606.
73. Stice E, Rohde P, Gau JM, Wade E. Efficacy trial of a brief cognitive–behavioral depression prevention program for high-risk adolescents: effects at 1- and 2-year follow-up. J Consult Clin Psychol. 2010;78:856-867.
74. Szigethy E, Kenney E, Carpenter J, *et al*. Cognitive-behavioral therapy for adolescents with inflammatory bowel disease and subsyndromal depression. J Am Acad Child Adolesc Psychiatry. 2007;46:1290-1298.
75. Thompson RD, Craig A, Crawford EA, *et al*. Longitudinal results of cognitive behavioral treatment for youths with inflammatory bowel disease and depressive symptoms. J Clin Psychol Med Settings. 2012;19:329-337.
76. TADS Team. Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for Adolescents with Depression Study (TADS) randomized controlled trial. JAMA. 2004;292:807-820.
77. Kennard B, Silva S, Vitiello B, *et al*. Remission and residual symptoms after short-term treatment in the Treatment of Adolescents with Depression Study (TADS). J Am Acad Child Adolesc Psychiatry. 2006;45:1404-1411.

78. Vitiello B, Silva SG, Rohde P, *et al*. Suicidal events in the Treatment for Adolescents with Depression Study (TADS). J Clin Psychiatry. 2009;70:741-747.
79. Vostanis P, Feehan C, Grattan E, Bickerton W. Treatment for children and adolescents with depression: lessons from a controlled trial. Clin Child Psychol Psychiatry. 1996a;1: 199-212.
80. Vostanis P, Feehan C, Grattan E, Bickerton W. A randomised controlled out-patient trial of cognitive-behavioural treatment for children and adolescents with depression: 9-month follow-up. J Affect Disord. 1996b;40:105-116.
81. Vostanis P, Feehan C, Grattan E. Two-year outcome of children treated for depression. Eur Child Adolesc Psychiatry. 1998;7:12-18.
82. Weisz JR, Southam-Gerow M, Gordis EB, *et al*. Cognitive–behavioral therapy versus usual clinical care for youth depression: an initial test of transportability to community clinics and clinicians. J Consult Clin Psychol. 2009;77:383-396.
83. Weisz JR, Thurber CA, Sweeney L, Proffitt VD, LeGagnoux GL. Brief treatment of mild-to-moderate child depression using primary and secondary control enhancement training. J Consult Clin Psychol. 1997;65:703-707.
84. Yang W, Zhang JX, Ding Z, Xiao L. Attention bias modification treatment for adolescents with major depression: a randomized controlled trial. J Am Acad Child Adolesc Psychiatry. 2016;55:208-218.
85. Young JF, Mufson L, Davies M. Efficacy of Interpersonal Psychotherapy—Adolescent Skills Training: an indicated preventive intervention for depression. J Child Psychol Psychiatry. 2006;47:1254-1262.
86. Young JF, Mufson L, Gallop R. Preventing depression: a randomized trial of Interpersonal Psychotherapy—Adolescent Skills Training. Depress Anxiety. 2010;27: 426-433.
87. Yu DL, Seligman MEP. Preventing depressive symptoms in Chinese children. Prev Treat. 2002;5. AAI9953624.
88. Weersing VR, Jeffreys M, Do MT, Schwartz KTG, Bolano C. Evidence base update of psychosocial treatments for child and adolescent depression. J Clin Child Adolesc Psychol. 2017;46:11-43.
89. Weisz JR, Kuppens S, Eckshtain D, Ugueto AM, Hawley KM, Jensen-Doss A. Performance of evidence-based youth psychotherapies compared with usual clinical care: a multilevel meta-analysis. JAMA Psychiatry. 2013;70:750-761.
90. Casey RJ, Berman JS. The outcome of psychotherapy with children. Psychol Bull. 1985;98:388-400.
91. Kazdin AE, Bass D, Ayers WA, Rodgers A. Empirical and clinical focus of child and adolescent psychotherapy research. J Consult Clin Psychol. 1990;58:729-740.
92. Weisz JR, Kazdin AE. Evidence-Based Psychotherapies for Children and Adolescents (2nd ed.). New York, NY: Guilford Press; 2010.
93. Weisz JR, Weiss B, Alicke MD, Klotz ML. Effectiveness of psychotherapy with children and adolescents: a meta-analysis for clinicians. J Consult Clin Psychol. 1987;55: 542-549.
94. Weisz JR, Weiss B, Han SS, Granger DA, Morton T. Effects of psychotherapy with children and adolescents revisited: a meta-analysis of treatment outcome studies. Psychol Bull. 1995;117:450-468.
95. Weisz JR, Sandler IN, Durlak JA, Anton BS. Promoting and protecting youth mental health through evidence-based prevention and treatment. Am Psychol. 2005;60: 628-648.
96. Weisz JR. Psychotherapy for Children and Adolescents: Evidence-Based Treatments and Case Examples. New York, NY: Cambridge University Press; 2004.
97. Costello EJ, Angold A, Keeler GP. Adolescent outcomes of childhood disorders: the consequences of severity and impairment. J Am Acad Child Adolesc Psychiatry. 1999; 38:121-128.
98. Jensen AL, Weisz JR. Assessing match and mismatch between practitioner-generated and standardized interview-generated diagnoses for clinic-referred children and adolescents. J Consult Clin Psychol. 2002;70:158-168.
99. Weisz JR, Ugueto AM, Cheron DM, Herren J. Evidence-based youth psychotherapy in the mental health ecosystem. J Clin Child Adolesc Psychol. 2013;42:274-286.
100. Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. Am J Ment Defic. 1981; 86:127-137.
101. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York, NY: Erlbaum; 1988.
102. Lipsey M, Wilson D. Practical Meta-Analysis (Applied Social Research Methods). Thousand Oaks, CA: Sage; 2000.
103. Smith ML, Glass GV, Miller TI. The Benefits of Psychotherapy. Baltimore, MD: Johns Hopkins University Press; 1980.
104. Hedges L, Olkin I. Statistical Methods for Meta-Analysis. New York, NY: Academic Press; 1985.
105. Torgerson CJ. Publication bias: The Achilles' heel of systematic reviews? Br J Educ Stud. 2006;54:89-102.
106. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315:629-634.
107. Duval S, Tweedie R. Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. Biometrics. 2000;56:455-463.
108. Higgins JP, Green S, Scholten R. Cochrane Handbook for Systematic Reviews of Interventions. Chichester, UK: Wiley-Blackwell. 2008:31-46.
109. Kocsis JH, Gerber AJ, Milrod B, *et al*. A new scale for assessing the quality of randomized clinical trials of psychotherapy. Compr Psychiatry. 2010;51: 319-324.
110. Herbert JD, Gaudiano BA. Moving from empirically supported treatment lists to practice guidelines in psychotherapy: the role of the placebo concept. J Clin Psychol. 2005;61:893-908.
111. TADS Team. Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for Adolescents with Depression Study (TADS) randomized control trial. JAMA. 2004;292:807-820.
112. Kennard B, Emslie GJ, Mayes TL, Hughes JL. Relapse and recurrence in pediatric depression. Child Adolesc Psychiatr Clin N Am. 2006;15:1057-1079.
113. Simons AD, Rohde P, Kennard B, Robins M. Relapse and recurrence prevention in the Treatment for Adolescents with Depression Study. Cogn Behav Pract. 2005;12: 240-251.
114. Angold A, Costello EJ, Erkanli A. Comorbidity. J Child Psychol Psychiatry. 1999; 40:57-87.
115. Garber JWV. Comorbidity of anxiety and depression in youth: Iimplications for treatment and prevention. Clin Psychol: Sci Pract. 2010;17:293-306.
116. Chorpita BF, Weisz JR. Modular Approach to Therapy for Children with Anxiety, Depression, Trauma, or Conduct Problems (MATCH-ADTC). Satellite Beach, FL: PracticeWise; 2009.
117. Ehrenreich-May J, Kennedy SM, Sherman JA, *et al*. Unified Protocols for Transdiagnostic Treatment of Emotional Disorders in Children and Adolescents: Therapist Guide. New York, NY: Oxford University Press; 2018.
118. Weisz JR, Chorpita BF, Palinkas LA, *et al*. Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: a randomized effectiveness trial. Arch Gen Psychiatry. 2012;69:274-282.
119. Chorpita BF, Daleiden EL, Park AL, *et al*. Child STEPs in California: a cluster randomized effectiveness trial comparing modular treatment with community implemented treatment for youth with anxiety, depression, conduct problems, or traumatic stress. J Consult Clin Psychol. 2017;85:13-25.
120. Weersing VR, Brent DA. Cognitive behavioral therapy for depression in youth. Child Adolesc Psychiatr Clin N Am. 2006;15:939-957.
121. Mufson LDK, Moreau D, Weissman MM. Interpersonal Psychotherapy for Depressed Adolescents. 2 ed. New York, NY: Guilford Press; 2004.
122. Spielmans GI, Gatlin ET, McFall JP. The efficacy of evidence-based psychotherapies versus usual care for youths: controlling confounds in a meta-reanalysis. Psychother Res. 2010;20:234-246.
123. Lovejoy MC, Graczyk PA, O'Hare E, Neuman G. Maternal depression and parenting behavior: a meta-analytic review. Clin Psychol Rev. 2000;20:561-592.
124. Garber J, Clarke GN, Weersing VR, *et al*. Prevention of depression in at-risk adolescents: a randomized controlled trial. JAMA. 2009;301:2215-2224.
125. Beardslee WR, Brent DA, Weersing VR, *et al*. Prevention of depression in at-risk adolescents: longer-term effects. JAMA Psychiatry. 2013;70:1161-1170.
126. Eckshtain D, Marchette LK, Schleider J, Weisz JR. Parental depressive symptoms as a predictor of outcome in the treatment of child depression. J Abnorm Child Psychol. 2018;46:825-837.
127. Ng MY, Weisz JR. Annual research review: Building a science of personalized intervention for youth mental health. J Child Psychol Psychiatry. 2016;57:216-236.
128. Simon GE, Perlis RH. Personalized medicine for depression: can we match patients with treatments? Am J Psychiatry. 2010;167:1445-1455.
129. Institute of Medicine (IOM). Psychosocial Interventions for Mental and Substance Use Disorders: A Framework for Establishing Evidence-Based Standards. Washington, DC: National Academies Press; 2015.
130. Curtin SC. Warner M, Hedegaard H. Increase in suicide in the United States, 1999–2014. NCHS Data Brief. 2016;241. Centers for Disease Control and Prevention. Available at: http://www.cdc.gov/nchs/products/databriefs/db241.htm. Accessed October 11, 2018.
131. Brent DABM, Bridge J, Chen T, Chiappetta L. Age- and sex-related risk factors for adolescent suicide. J Am Acad Child Adolesc Psychiatry. 1999;38: 1497-1505.
132. Wampold BE, Mondin GW, Moody M, Stich F, Benson K, Ahn H. A meta-analysis of outcome studies comparing bona fide psychotherapies: empiricially, 'all must have prizes. Psychol Bull. 1997;122:203-215.
133. Wampold BE, Flückiger C, Del Re AC, *et al*. In pursuit of truth: a critical examination of meta-analyses of cognitive behavior therapy. Psychother Res. 2017; 27:14-32.