

Between- and Within-Person Relations Between Treatment Modules and Symptom Improvements Across Six Randomized Controlled Trials

Katherine E. Ventura-Conerly¹, Patrick Mair¹, Christian Webb^{2, 3}, Emma Hill³,
Bruce F. Chorpita⁴, Sally N. Merry⁵, and John R. Weisz¹

¹ Department of Psychology, Harvard University

² Center for Depression, Anxiety and Stress Research, McLean Hospital, Belmont, Massachusetts, United States

³ Department of Psychiatry, Harvard Medical School, Harvard University

⁴ Department of Psychology, University of California, Los Angeles

⁵ Department of Psychological Medicine, Auckland School of Medicine, University of Auckland

Objective: Most youth psychotherapies contain multiple treatment elements; little is known about their relative effectiveness. We assessed symptom improvements associated with treatment elements, represented by modules within modular psychotherapy. **Method:** Data from six clinical trials of the modular approach to therapy for children with anxiety, depression, trauma, or conduct problems modular psychotherapy were combined ($N = 490$; 5,403 sessions; 6–15 years) to test effects of modules grouped into seven common treatment principles: feeling calm (e.g., relaxation strategies), increasing motivation (e.g., contingency management), repairing thoughts (e.g., cognitive restructuring), solving problems (e.g., problem solving), trying the opposite (e.g., exposure), engagement/psychoeducation (e.g., building rapport), and future planning (e.g., planning skill use). Multilevel models with autoregressive covariance controlled for previous symptoms, session number, and baseline symptoms; accounted for temporality with each session occurring prior to associated outcomes; importantly, we modeled associations between outcomes and between- and within-person use of treatment principles. Measures included weekly youth- and caregiver-reported internalizing, externalizing, and total symptoms, plus idiographic top problems. **Results:** A between-person effect linked future planning ($Bs = -.369$ to $-.368$; $ps < .05$) to better outcomes. Better within-person effects were observed for increasing motivation ($Bs = -.087$ to $-.057$; $ps < .05$), trying the opposite ($Bs = -.087$ to $-.056$; $ps < .05$), and future planning ($B = -.146$; $p < .001$). Six outcomes showed worse within-person effects for engagement/psychoeducation ($Bs = .036$ – $.099$; $ps < .05$); efforts to build engagement drove this finding; those efforts were associated with less use of skills-focused principles. **Conclusions:** Increasing motivation, trying the opposite, and future planning were associated with especially good outcomes; therapists' engagement attempts were associated with less focus on skill building and poorer immediate outcomes. Results may inform youth psychotherapy decision-making research and practice.

What is the public health significance of this article?

Modules associated with increasing motivation, trying the opposite, and future planning were associated with especially strong immediate outcomes, while modules associated with engagement/psychoeducation were associated with relatively poor outcomes. The association between engagement/psychoeducation and poorer immediate outcomes seemed to be mostly driven by “getting to know you” activities, not by psychoeducation, and use of these modules was negatively correlated with use of skill-building modules. When combined with other information (e.g., assessment) and clinical judgment, these findings may guide treatment element selection and prioritization in youth psychotherapy.

Keywords: child and adolescent psychotherapy, anxiety and posttraumatic stress disorders, depressive disorders, conduct disorders, decision making

Supplemental materials: <https://doi.org/10.1037/ccp0000972.supp>

Jami F. Young served as action editor.

Katherine E. Ventura-Conerly  <https://orcid.org/0000-0003-3559-2491>

The authors have not published, posted, or submitted any papers using this data set and related individual primary outcome studies (from which the data used is drawn) are cited. Data are not publicly available; it comes from six different studies, and the authors of those studies should be contacted regarding access to their data. Data requests will be considered on an individual basis by the authors. Analytic code for main analyses is available on the Open Science Framework at <https://osf.io/w4qxm>.

This research was supported by the Charles J. Gelso, Psychotherapy Research Grant from American Psychological Association Division 29: Society for the Advancement of Psychotherapy and the U.S. National Institute of Mental Health, National Institutes of Health (NIH) (F31MH134555; principal investigator: Katherine E. Ventura-Conerly). Emma Hill's work on this project is also supported by the Passionate Leaders Project at Simmons University. Katherine E. Ventura-Conerly reports funding from the Templeton World Charity Foundation, the Harvard University Dean's Competitive Fund, the U.S. National Institute of Mental Health (F31MH134555), and the Fund for Innovation in Development. Christian Webb is partially supported by NIMH

continued

In the past six decades, over 450 randomized controlled trials (RCTs) of youth (i.e., child and adolescent) psychotherapies have been conducted, revealing overall moderate effects on clinical symptoms (Weisz et al., 2017, 2019). However, these psychotherapies vary considerably in their effectiveness (Fonagy et al., 2015; Weisz et al., 2017, 2019) perhaps in part because of variation in the specific “treatment elements” (e.g., exposure, cognitive restructuring) they include (Fitzpatrick et al., 2023; Leijten et al., 2021; Mulder et al., 2017). Because most psychotherapies include multiple treatment elements, the independent effects of the different elements are typically unclear (Fitzpatrick et al., 2023; Venturo-Conerly et al., 2023).

Because little is known about which treatment elements are most strongly associated with outcomes (Leijten et al., 2021; Mulder et al., 2017), clinical judgment is often used to select treatment elements, sometimes informed by other inputs such as supervision, assessment, reviews of past literature, flowcharts, measurement-based care (Scott & Lewis, 2015), and client input (Venturo-Conerly et al., 2023). A theory-driven and data-driven approach to understanding associations between treatment elements and outcomes could usefully complement current approaches to clinical decision making by adding predictive data to the clinical information available to clinicians. Such a contribution may be particularly relevant to flexible, modular psychotherapies (Chorpita et al., 2005; Ng & Weisz, 2016; Venturo-Conerly et al., 2023; Weisz & Chorpita, 2012; Weisz et al., 2015) such as modular approach to therapy for children with anxiety, depression, trauma, or conduct problems (MATCH) (Chorpita & Weisz, 2009), in which therapists select from a menu of therapeutic modules (e.g., relaxation strategies, behavioral activation) across a series of treatment sessions for each client (Venturo-Conerly et al., 2023; Weisz & Chorpita, 2012). Typically, modules are selected throughout treatment, session-by-session, depending on client baseline characteristics, newly emerging problems, and response to session content (Venturo-Conerly et al., 2023)—a process designed to fit the dynamic conditions of clinical practice (Chorpita et al., 2015).

Several RCTs have supported the effectiveness of modular psychotherapies (Bennett et al., 2024; Chorpita et al., 2017; Weisz et al., 2012), but others have not shown modular psychotherapies to outperform usual care (Merry et al., 2020; Weisz, Bearman, et al., 2020). One explanation suggested for the differing findings is that

they reflect variations across trials in support for and quality of clinicians’ decision making (Venturo-Conerly et al., 2023). Indeed, the literature on clinical decision making (Cohen et al., 2021; Leijten et al., 2021; Ng & Weisz, 2016) shows little empirical guidance for selection of specific treatment elements. A recent review of all flexible, modular youth psychotherapies found that all protocols proposed using clinical judgment when choosing modules (Venturo-Conerly et al., 2023), 40% proposed considering which problems are most urgent or severe and 20% proposed considering which problems are most tractable. Baseline assessment (95%), measurement-based care (65%), and relevant theories (55%) were also suggested to inform treatment target and element selection but often with few specifics proposed. Despite evidence that statistical models of archival data significantly outperform clinician judgment (Ægisdóttir et al., 2006; Meehl, 1954, 1986), no protocols used statistical analyses of archival data to guide module selection.

We took initial steps toward building such statistical models. We modeled relations between use of treatment modules and subsequent clinical symptom scores, using data from trials of MATCH, the most widely used and studied youth modular treatment. RCTs testing each MATCH module individually would be infeasible and would not represent actual practice because modules are used in combination. So, we used an approach capable of rigorously documenting module-outcome associations outside of the context of an RCT (Webb et al., 2019). To support true causal relationships, one must identify statistically significant associations between treatment procedures and outcomes, account for temporality (i.e., treatment module used prior to symptom score changes), and establish nonspuriousness (i.e., account for third-variable confounds that might influence both treatment module use and symptoms). To approximate these features outside of an RCT, one may model the extent to which associations can be accounted for by within-person effects, controlling for between-person effects, to assess separately: (a) to what extent between-person variance in receipt of treatment modules (throughout a treatment episode) predicts symptoms and (b) to what extent within-person (week-to-week) variance in receipt of treatment modules predicts subsequent symptoms (Feeley et al., 1999). By separating between-person and within-person variance, and by accounting for symptoms in previous sessions, one may

R01MH116969, NCCIH R01AT011002, the Tommy Fuss Fund, and a National Alliance for Research on Schizophrenia and Depression Young Investigator Grant from the Brain and Behavior Research Foundation. The authors also extend our gratitude to Malia Clark, Sanjana Ramrajvel, Michael Spitzer, Hannah Broderick, Roger Saumure, Alyssa Cameron, and Katie Cabrera, who assisted with data collation and preparation and with logistical aspects of article preparation. The randomized effectiveness trials that were the sources of data for this study were supported by funding to John R. Weisz from the John D. and Catherine T. MacArthur Foundation (Grant 07-90231-000-HCD), Casey Family Programs (Grant 02200), the Annie E. Casey Foundation (Grant 211.0004), the Child Health and Human Development Institute, and the Norlien Foundation, and to Sally N. Merry from the Health Research Council, New Zealand. Finally, Katherine E. Venturo-Conerly reports working for the Shamiri Institute, a 501(c)(3) nonprofit organization. John R. Weisz has received royalties from Cambridge University Press, the Guilford Press, and PracticeWise. Bruce F. Chorpita has received royalties from PracticeWise and is a board member of PracticeWise.

Katherine E. Venturo-Conerly played a lead role in conceptualization, formal analysis, funding acquisition, investigation, methodology, software, writing—

original draft, and writing—review and editing, a supporting role in data curation and project administration, and an equal role in validation and visualization. Patrick Mair played a supporting role in conceptualization, data curation, formal analysis, investigation, methodology, software, supervision, visualization, and writing—review and editing and an equal role in validation. Christian Webb played a supporting role in conceptualization, formal analysis, funding acquisition, investigation, methodology, software, supervision, validation, and writing—review and editing. Emma Hill played a supporting role in data curation, formal analysis, software, writing—original draft, and writing—review and editing and an equal role in validation and visualization. Bruce F. Chorpita played a supporting role in data curation and writing—review and editing. Sally N. Merry played a supporting role in data curation and writing—review and editing. John R. Weisz played a lead role in data curation, resources, and supervision, a supporting role in conceptualization, funding acquisition, investigation, methodology, and writing—review and editing, and an equal role in project administration.

Correspondence concerning this article should be addressed to Katherine E. Venturo-Conerly, Department of Psychology, Harvard University, 33 Kirkland Avenue, Cambridge, MA 02138, United States. Email: kventuroconerly@g.harvard.edu

eliminate potential between-person confounds that might influence module-use-to-symptom-score associations (e.g., confounds such as therapist tendencies to select particular modules overall or for particular clients). Inclusion of a between-person term representing overall amount of a module received across treatment accounts for third variables that influence treatment modules selection across treatment (e.g., clinician preferences, client preferences, and client presenting problems), thus allowing for more accurate estimates of the pure associations between within-person module use and outcomes (for more details, see Webb et al., 2019). Without a randomized design, one cannot be certain that third-variable confounds are not present, but by accounting for likely causes of spuriousness, one may have considerably stronger confidence in modeled associations. Modeling how specific modules relate to change in symptom severity may therefore, in combination with other clinically relevant information, help inform clinical decision making. This rationale informed the methods used in the present study.

Method

Data were combined from six RCTs (Harmon et al., 2021; Merry et al., 2020; Weisz, Bearman, et al., 2020; Weisz et al., 2012, 2018; Weisz, Thomassin, et al., 2020) of MATCH (Chorpita & Weisz, 2009), for youths aged 6–15. The MATCH treatment manual is published and available for purchase (Chorpita & Weisz, 2009); it is copyrighted and cannot legally be posted freely online. All study procedures for these trials were approved by the relevant IRBs. This study was preregistered on the Open Science Forum (OSF) at <https://osf.io/mr87g>. See supplement for study updates made after preregistration.

Measures

The 33 MATCH modules (Tables 1 and 2) are components of psychotherapies (e.g., exposure) for common youth problems (anxiety, depression, trauma, conduct). Modules can be mixed and matched flexibly within each session at the therapist's discretion to personalize therapy (Chorpita & Weisz, 2009; Weisz et al., 2012). We used standard module use reports by therapists after each treatment session to investigate associations between therapists' use of modules and subsequent youth symptom scores. The validity of therapist reports of module use is supported by prior research showing concordance with independent coding of MATCH session recordings (Ward et al., 2013).

To analyze relationships between treatment modules and subsequent symptom severity, the 33 MATCH modules were grouped into overarching treatment principles (Goldfried, 2009) derived from theory and a recent review and meta-analysis (Fitzpatrick et al., 2023) of youth RCTs (Comer et al., 2019; Southam-Gerow & Prinstein, 2014; Weersing et al., 2017) that identified five empirically supported principles of change (Weisz & Bearman, 2020; Weisz et al., 2017) ESPCs: (a) feeling calm—reducing tension and arousal (e.g., progressive muscle relaxation), (b) increasing motivation—using contingencies to increase desirable and decrease undesirable behavior (e.g., rewards), (c) repairing thoughts—changing unhelpful cognitions (e.g., reappraisal of depressive thoughts), (d) solving problems—learning effective problem solving procedures, and (e) trying the opposite—practicing the positive opposite of unhelpful behaviors (e.g., exposure). All MATCH modules fit one of these five principles except those in two categories that were added for this study:

engagement/psychoeducation (e.g., getting acquainted and getting-to-know-you activities, engaging parents, general conversation and social support, and joint activities to build rapport, and psychoeducation for children and caregivers) and future planning (e.g., planning posttherapy skill use).

Eight dependent variables, consisting of youth- and caregiver-report versions of four measures, internalizing symptoms, externalizing symptoms, and total symptoms, and idiographic top problems, were analyzed as separate outcomes. The dependent variables included approximately weekly measures of internalizing symptoms, externalizing symptoms, and total symptoms assessed via separate youth- and caregiver-report versions of the brief problem monitor (Piper et al., 2014), behavior and feelings survey (Weisz, Vaughn-Coaxum, et al., 2020), and brief problem checklist (Chorpita et al., 2010). These measures are all highly correlated with scores on the child behavior checklist and youth self report (Achenbach & Rescorla, 2001), measure the same constructs, and can be used interchangeably when standardized, as in this study. They also have strong reliability and validity (see Supplemental Material). An additional measure assessed weekly changes in severity of idiographic top problems. This was the top problems assessment (Weisz et al., 2011), which provided weekly severity ratings by each youth and caregiver (on a standardized scale) for the problems identified as most important by the youth and caregiver (separately) at the start of therapy. The top problems assessment shows good reliability and validity, and to add important clinical information as part of measurement-based care in clinical practice (see Supplemental Material).

Sample

Four hundred ninety youths received MATCH in one of six RCTs (Harmon et al., 2021; Merry et al., 2020; Weisz, Bearman, et al., 2020; Weisz et al., 2012, 2018; Weisz, Thomassin, et al., 2020). Mean (*SD*) age was 10.14 (2.36) years; 43.5% were born female, and 59.40% were White, 10.40% Black/African American, 11.2% Latino/Hispanic, 14.70% multiracial, 1.40% Asian, and 2.20% another unspecified race/ethnicity. Total treatment sessions averaged 11.03 (*SD* = 8.21) and total modules averaged 14.82 (*SD* = 14.78). The mean 11.03 sessions in this trial is lower than the reported mean of 16.54 planned sessions from a large youth psychotherapy meta-analysis (Weisz et al., 2017); this may, however, reflect true and generalizable differences between standardized and flexible, modular treatments, or between planned research protocols and implementation in clinical practice. Session attendance was encouraged weekly, but actual session frequency varied in practice (see supplement for more on session and outcome measurement timing). See Table 3.

Analyses

Main analyses focused on weekly use (yes/no) of each of the seven treatment principles. Each use of a treatment principle in a session was matched with the temporally closest subsequent report of youth, and separately caregiver, outcomes. When no outcome data were available that had been collected between the day after a given session and the day of the following session, the outcome data associated with that session was imputed instead using the missForest function in R, using five iterations of 100 trees (Kokla et al., 2019; Stekhoven & Bühlmann, 2012). Across all weekly symptom measures, 13.16%–15.77% of

Table 1
Module Descriptions and Use Information

Module name	Module description	No. of sessions with module received ($N_{\text{sessions}} = 5,403$)	How far through treatment module was received (see note) M (SD)	No. of clients who received each module ($N = 490$)	M (SD) no. of times module was received per client	M (SD) no. of times module was received among clients who received the module one + times
Getting acquainted	Client and clinician complete “get to know you” activities and overview the structure and content of treatment.	480 (8.88%)	0.26 (0.25)	297 (60.61%)	0.98 (1.16)	1.62 (1.09)
(Anxiety) fear ladder	Child and parent identify and rate a “ladder” of specific fears.	333 (6.16%)	0.43 (0.28)	128 (26.12%)	0.69 (1.52)	2.63 (1.95)
(Anxiety) learning about anxiety child	Child receives psychoeducation about anxiety and its treatment.	295 (5.46%)	0.41 (0.24)	121 (24.69%)	0.6 (1.36)	2.41 (1.76)
(Anxiety) learning about anxiety parent	Parent receives psychoeducation about anxiety and its treatment.	126 (2.33%)	0.45 (0.26)	82 (60.61%)	0.26 (0.73)	1.54 (1.10)
(Anxiety) practicing	Child practices exposure to feared situations or things using a thermometer to rate fear/anxiety.	475 (8.79%)	0.63 (0.22)	89 (18.16%)	0.97 (2.78)	5.34 (4.41)
(Anxiety) maintenance	Child reviews skills learned in treatment and considers how they might be used in the future. Parent learns about identifying relapses.	62 (1.15%)	0.82 (0.22)	36 (7.35%)	0.13 (0.56)	1.72 (1.21)
(Anxiety) cognitive STOP	Child learns to identify anxious self-talk in anxious situations, and switch to coping self-talk.	138 (2.55%)	0.71 (0.21)	48 (9.80%)	0.28 (1.28)	2.85 (3.08)
Wrap up	Child reviews progress and creates a “commercial” for their program. Parents and child plan to continue treatment progress posttherapy.	100 (1.85%)	0.94 (0.14)	75 (15.31%)	0.2 (0.56)	1.33 (0.72)
(Trauma) safety planning	Child learns skills for maintaining personal safety, including where and when to ask for help.	29 (0.54%)	0.53 (0.31)	13 (2.65%)	0.06 (0.4)	2.23 (1.09)
(Trauma) narrative	Child creates a written narrative of the traumatic event.	43 (0.80%)	0.68 (0.24)	9 (1.84%)	0.09 (0.75)	4.78 (2.21)
(Depression) learning about depression child	Child receives psychoeducation about depression and its treatment.	352 (6.51%)	0.39 (0.27)	174 (35.51%)	0.71 (1.3)	2.01 (1.48)
(Depression) learning about depression parent	Parent receives psychoeducation about depression and its treatment.	143 (2.65%)	0.43 (0.26)	110 (22.45%)	0.29 (0.63)	1.31 (0.66)
(Depression) problem solving	Child learns five STEPS for effective problem solving (e.g., “Say what the problem is; Think of solutions”).	473 (8.75%)	0.59 (0.25)	170 (34.69%)	0.97 (1.86)	2.78 (2.21)
(Depression) activity selection	Child learns about how actions and feelings are associated and identifies 10 mood-boosting activities.	261 (4.83%)	0.57 (0.25)	122 (24.90%)	0.53 (1.28)	2.12 (1.78)
(Depression) learning to relax	Child learns about associations between feelings and physical tension, and progressive muscle relaxation.	187 (3.46%)	0.58 (0.24)	120 (24.49%)	0.38 (0.81)	1.55 (0.94)
(Depression) quick calming	Child learns to identify areas on chronic body tension, and learns the ReST concept: relax muscles, slow breathing, think of a peaceful place.	223 (4.13%)	0.67 (0.22)	123 (25.10%)	0.46 (1.06)	1.81 (1.43)
(Depression) presenting a positive self	Child learns how acting positive can improve mood and relationships, role-plays the positive self, and identifies situations in which they can practice it.	155 (2.87%)	0.67 (0.2)	74 (15.10%)	0.32 (1.1)	2.09 (2.08)
(Depression) cognitive BLUE	Child learns about negative thinking traps and learns that thoughts are just guesses.	207 (3.83%)	0.68 (0.2)	74 (15.10%)	0.42 (1.4)	2.76 (2.58)
(Depression) cognitive TLC	Child discusses some skills for addressing negative thoughts and feelings such as talking to a friend and refocusing on something more positive.	97 (1.80%)	0.75 (0.17)	56 (11.43%)	0.2 (0.7)	1.75 (1.25)

(table continues)

Table 1 (continued)

Module name	Module description	No. of sessions with module received ($N_{\text{sessions}} = 5,403$)	How far through treatment module was received (see note) M (SD)	No. of clients who received each module ($N = 490$)	M (SD) no. of times module was received per client	M (SD) no. of times module was received among clients who received the module one + times
(Depression) plans for coping	Child reviews skills learned in treatment and considers how they might use them in the future to maintain treatment gains. Parent learns about identifying relapses.	150 (2.78%)	0.86 (0.14)	64 (13.06%)	0.31 (1.09)	2.34 (2.09)
(Conduct) engaging parents	Clinician builds rapport with the parent. Parents identify goals for therapy. Clinician outlines their role and the parent's role in addressing child conduct problems.	346 (6.40%)	0.36 (0.29)	190 (38.78%)	0.71 (1.45)	1.82 (1.84)
(Conduct) learning about behavior	Parent receives psychoeducation about child disruptive behavior and its treatment.	383 (7.09%)	0.43 (0.27)	178 (36.33%)	0.78 (1.63)	2.16 (2.09)
(Conduct) one-on-one time	Parents learn to improve the quality and frequency of their one-on-one interactions with their child.	515 (9.53%)	0.55 (0.26)	167 (34.08%)	1.04 (2.18)	3.07 (2.80)
(Conduct) praise	Parents learn about rewarding appropriate behaviors with labeled praise and attention.	408 (7.55%)	0.6 (0.24)	143 (29.18%)	0.82 (1.92)	2.83 (2.64)
(Conduct) active ignoring	Parents learn to intentionally ignore mildly disruptive behaviors (e.g., whining) to decrease their frequency.	303 (5.61%)	0.62 (0.22)	115 (23.47%)	0.61 (1.48)	2.60 (2.04)
(Conduct) giving effective instructions	Parents learn when and how to give clear, effective, and simple instructions to increase compliance and learn common traps that decrease instruction effectiveness.	233 (4.31%)	0.66 (0.21)	98 (20.00%)	0.47 (1.24)	2.35 (1.81)
(Conduct) rewards	Parent learns about how positive reinforcement can increase the frequency of desired behaviors, and identifies behaviors to reward and possible rewards.	389 (7.20%)	0.66 (0.21)	103 (21.02%)	0.83 (2.19)	3.93 (3.25)
(Conduct) time out	Parent learns about benefits of and procedures for an effective time out.	109 (2.02%)	0.78 (0.17)	49 (10.00%)	0.22 (0.85)	2.20 (1.68)
(Conduct) making a plan	Parent identifies situations in which their child is likely to misbehave and plans how to address misbehaviors in these situations.	100 (1.85%)	0.85 (0.16)	54 (11.02%)	0.2 (0.75)	1.81 (1.49)
(Conduct) daily report card	Parents list concerns about behaviors at school, ask school personnel to provide daily reports about these concerns, and identify rewards for appropriate behaviors.	54 (1.00%)	0.73 (0.23)	20 (4.08%)	0.11 (0.71)	2.70 (2.34)
(Conduct) looking ahead	Parent reviews skills learned in treatment and considers how they might be used to maintain treatment gains.	78 (1.44%)	0.91 (0.16)	50 (10.20%)	0.16 (0.54)	1.54 (0.86)
(Conduct) booster	Parent learns about identifying relapses. Parents review current strategies for child conduct problems and discuss which to continue or replace.	18 (0.33%)	0.96 (0.08)	18 (3.67%)	0.05 (0.26)	1.28 (0.57)

Note. Multiple modules could be received in one session. The target problems in parentheses in the first column indicate the primary target problem associated with each module. Those modules without a target problem listed in parentheses are not assigned one consistent primary target problem. Column four, showing how far through treatment each module is received, is calculated by, each time a module is used with a client, dividing the session number in which it is used by the total number of sessions received by that client; then, overall means and standard deviations of these proportions for each module are calculated (and reported in the table) to represent when during a course of treatment modules are used on average and how much variance there is in when these modules are used. STOP = scared, thoughts, other thoughts, praise; BLUE = blaming myself, looking for bad news, unhappy guessing, exaggerating; TLC = talk to a friend, look for a silver lining, change the channel.

Table 2
Treatment Principle Descriptions and Use Information

Principle name	Modules included within treatment principle	No. of sessions with principle received (<i>N</i> _{sessions} = 5,403)	How far through treatment principle was received (see note) <i>M</i> (<i>SD</i>)	No. of clients who received each principle (<i>N</i> = 490)	<i>M</i> (<i>SD</i>) no. of times principle was received per client	<i>M</i> (<i>SD</i>) no. of times principle was received among clients who received the principle one + times
FIRST—Feeling calm	Learning to relax, quick calming	372 (6.89%)	0.63 (0.23)	159 (32.45%)	0.76 (1.4)	2.33 (1.54)
FIRST—Increasing motivation	One-on-one time, praise, active ignoring, giving effective instructions, rewards, time out, daily report card	1,234 (22.84%)	0.6 (0.25)	204 (41.63%)	2.62 (4.95)	6.30 (5.97)
FIRST—Repairing thoughts	Cognitive STOP, cognitive BLUE, cognitive TLC	417 (7.72%)	0.7 (0.2)	126 (25.71%)	0.84 (2.11)	3.29 (3.05)
FIRST—Solving problems	Problem solving, safety planning, making a plan	599 (11.09%)	0.63 (0.26)	215 (43.88%)	1.22 (2.01)	2.78 (2.21)
FIRST—Trying the opposite	Practicing, Activity Selection, Narrative, Presenting A Positive Self	918 (16.99%)	0.62 (0.23)	217 (44.29%)	1.88 (3.23)	4.24 (3.69)
Engagement/psychoeducation	Getting Acquainted, Fear Ladder, Learning About Anxiety Child, Learning About Depression Parent, Learning About Depression Parent, Learning About Behavior, Engaging Parents	1993 (36.89%)	0.38 (0.28)	470 (95.92%)	4.17 (3.21)	4.34 (3.15)
Future planning	Wrap Up, Plans For Coping, Looking Ahead, Maintenance, Booster	369 (6.83%)	0.88 (0.16)	164 (33.47%)	0.79 (1.54)	2.35 (1.86)

Note. Multiple principles could be received in one session. Column four, showing how far through treatment each principle is received, is calculated by, each time a principle is used with a client, dividing the session number in which it is used by the total number of sessions received by that client; then, overall means and standard deviations of these proportions for each principle are calculated (and reported in the table) to represent when during a course of treatment principles are used on average, and how much variance there is in when these principles are used. Getting acquainted is divided into two versions, one for anxiety and one for depression, in the MATCH manual; because these modules are nearly identical, they are combined into one getting acquainted module for this study. FIRST = feeling calm, increasing motivation, repairing thoughts, solving problems, trying the opposite; STOP = scared, thoughts, other thoughts, praise; BLUE = blaming myself, looking for bad news, unhappy guessing, exaggerating; TLC = talk to a friend, look for a silver lining, change the channel; MATCH = modular approach to therapy for children with anxiety, depression, trauma, or conduct problems.

Table 3
Sample Baseline Demographic Information

Characteristic	<i>M (SD)</i>
Baseline symptom raw scores	
Youth-rated top problems	5.04 (2.49)
Caregiver-rated top problems	6.43 (2.22)
Youth-rated total problems	6.40 (4.56)
Caregiver-rated total problems	9.76 (5.91)
Youth-rated internalizing	3.04 (2.71)
Caregiver-rated internalizing	4.25 (3.30)
Youth-rated externalizing	3.44 (2.69)
Caregiver-rated externalizing	5.54 (3.30)
Total sessions	11.03 (8.21)
Modules	14.82 (14.78)
Child age	10.14 (2.36)
Characteristic	%
Child sex	
Male	54.1
Female	43.5
NA	2.4
Child race/ethnicity	
White/Caucasian	59.4
Black/African American	10.4
Latino/Hispanic	11.2
Asian	1.4
Multicultural/mixed	14.7
Other	2.2
NA	0.6
Protocol	
Anxiety	13.7
Depression	23.9
Trauma	3.7
Conduct	34.1
NA	24.7

Note. The total population is 490 families. Each client is assigned a primary protocol by their clinical team based on their primary target problem. See the supplement for more on the baseline symptom measures, including their psychometric properties. Baseline symptom scores are the first consistently available symptom scores in the data set or the session one outcome measures. The range of total sessions is 1–47. NA = not available.

entries were missing from youth-report, and 8.49%–9.15% from caregiver-report. See supplement for more on the session-outcome matching process and for descriptives related to session and outcome timing.

For main analyses, linear mixed models were used to model the relationships between principle use and subsequent youth and caregiver symptoms. Associations between within-patient and between-patient principle use and subsequent symptoms were disaggregated (Curran & Bauer, 2011; Wang & Maxwell, 2015; Webb et al., 2019). These models produced estimates of within-person relationships between principle use and outcomes at each time point, while accounting for between-person effects. Specifically, in these models, each principle was separated into two predictor terms: one for between-person variance in principle use (i.e., mean use of a given principle for each person across treatment) and the other for the within-person variance in principle use (i.e., weekly presence or absence of principle use). This provided estimates within the same model of the associations between receipt of specific principles and symptoms *within* each participant from week-to-week, and the associations

between mean amount of specific principles received by each person *across all of treatment* and their symptom scores.

Additionally, session-by-session ratings of use of each principle (time T) served as predictors of symptoms at the closest time point after, (time $T + 1$, which fell between sessions). Models also accounted for baseline symptoms and the linear effect of time on symptoms (i.e., session number was included for detrending). Of note, in main models, baseline symptoms were defined as the first consistently available measurement point, which was after session one. In the supplement, models are also presented for those youth with baseline measures available before session one. They also included a random intercept and random slope for within-person variance in element use for each participant. Finally, to account for scores after the previous session (time $T - 1$), which could have influenced module selection (at time T) and scores (at time $T + 1$), an autoregressive covariance term was added to the models, as described in previous literature (Funatogawa & Funatogawa, 2012, 2018). Including a lagged dependent variable (i.e., the outcome value at time $T - 1$ as a predictor) was considered as an alternative, however, in mixed effects models, this can bias estimates (Allison, 2015). Therefore, we used mixed-effects longitudinal models specifically designed to account for outcomes at previous time points without bias using an autoregressive covariance term (Funatogawa & Funatogawa, 2012, 2018).

From these models, estimates were generated of the within- and between-person associations between use of each treatment principle and immediate symptom scores across all participants throughout the course of psychotherapy. For each of the seven treatment principles, eight models were run, one for each outcome, and a false discovery rate p value correction was also applied within the eight models relevant to a given treatment principle; results with the FDR correction are reported in the article body. For any treatment principles significantly associated with subsequent symptoms, as preregistered, additional exploratory models (without the false discovery rate correction) were run testing relationships between individual modules and subsequent symptoms. Exploratory models were also run including only those youths who had elevated baseline externalizing problems, and separately those with elevated baseline internalizing.

Results

Across all 490 youths with session ($N_{\text{sessions}} = 5,403$) and module use data, the most frequently used principle was engagement/psychoeducation (1,993 sessions); future planning was used least often (369 sessions). For more, see Tables 1 and 2. Within-person associations, between principles and immediate outcomes were significant for four of the seven principles (see Tables 4 and 5 for all findings and Figure 1 for plots of significant results).

Across six outcomes (caregiver-rated internalizing symptoms, total symptoms, and idiographic top problems, and youth-rated internalizing symptoms, total symptoms, and idiographic top problems), symptom scores were significantly higher (i.e., *worse*) on weeks when engagement/psychoeducation was used than when it was not ($B = .037-.099$; $ps < .05$). When applying a false discovery rate p value correction across all fixed effects for engagement/psychoeducation, the within-person effects remained significant for four of the measures, and marginally surpassed the threshold for the other two—that is, internalizing symptoms ($p = .057$) and caregiver-rated internalizing symptoms ($p = .074$).

Table 4*Between- and Within-Person Relations Between Modules Grouped by Treatment Principles and Caregiver-Report Symptoms*

Variable	Caregiver internalizing		Caregiver externalizing		Caregiver total problem		Caregiver top problems' assessment	
	Slope <i>B</i> (<i>SE</i>)	<i>p</i>	Slope <i>B</i> (<i>SE</i>)	<i>p</i>	Slope <i>B</i> (<i>SE</i>)	<i>p</i>	Slope <i>B</i> (<i>SE</i>)	<i>p</i>
Feeling calm within-person	−0.003 (.032)	.921	−0.019 (.036)	.600	−0.016 (.034)	.633	0.002 (.036)	.962
Feeling calm between-person	0.181 (.223)	.418	0.126 (.227)	.579	0.187 (.230)	.419	−0.204 (.246)	.408
Increasing motivation (within)	−0.057 (.026)	.030*	−0.075 (.032)	.018*	−0.087 (.029)	.003*	−0.024 (.031)	.429
Increasing motivation (between)	−0.016 (.086)	.853	0.004 (.090)	.967	−0.017 (.087)	.846	0.128 (.094)	.173
Repairing thoughts (within)	−0.044 (.036)	.218	−0.026 (.036)	.477	−0.046 (.036)	.196	−0.040 (.037)	.289
Repairing thoughts (between)	−0.240 (.206)	.245	−0.197 (.207)	.343	−0.232 (.210)	.270	−0.016 (.224)	.944
Solving problems (within)	0.016 (.030)	.598	0.034 (.030)	.260	0.030 (.032)	.345	−0.018 (.030)	.542
Solving problems (between)	−0.032 (.179)	.860	−0.030 (.183)	.871	−0.026 (.186)	.890	−0.060 (.200)	.764
Trying the opposite (within)	−0.028 (.025)	.261	−0.035 (.026)	.180	−0.037 (.026)	.157	−0.030 (.027)	.281
Trying the opposite (between)	0.235 (.128)	.068	−0.228 (.135)	.091	0.032 (.133)	.812	−0.017 (.141)	.906
Engagement/psychoeducation (within)	0.043 (.021)	.041*	0.037 (.024)	.122	0.053 (.022)	.017*	0.068 (.024)	.005*
Engagement/psychoeducation (between)	−0.032 (.084)	.776	0.105 (.085)	.217	0.051 (.085)	.550	−0.040 (.091)	.657
Future planning (within)	−0.062 (.036)	.082	−0.046 (.038)	.218	−0.065 (.037)	.077	−0.145 (.040)	<.001*
Future planning (between)	−0.453 (.256)	.077	−0.636 (.261)	.015*	−0.636 (.264)	.017*	−0.152 (.283)	.590

Note. The values shown in this table are the beta values for the association between weekly use (yes or no) of each treatment principle and the closest subsequent outcome (standardized). All models control for the baseline (defined as session one) symptom score on the outcome measure and for time in the form of session number (see Curran & Bauer, 2011 for more on detrending), and include a random intercept and slope term. These *p* values are uncorrected, corrected values are in text. More negative *B* values indicate better treatment outcomes immediately after use of the principle. *SE* = standard error.

* *p* < .05.

Within-person associations between use of future planning and subsequent symptom scores were significant on the caregiver-rated idiographic top problems measure; scores were lower (i.e., better) on weeks when future planning was used relative to when it was not (*B* = −0.145; *p* < .001). Additionally, for two measures (caregiver-rated externalizing symptoms, *B* = −.636, and total symptoms, *B* = −.636), greater use of future planning between-persons also predicted significantly better outcomes (*p* < .05). With a false discovery rate *p* value correction, these within-person and between-person effects remained significant.

Across caregiver-rated internalizing symptoms, externalizing symptoms, and total symptoms, scores were significantly (*B*s = −0.087 to −0.057; *ps* < .05) lower (i.e., better) on weeks when increasing motivation was used than when it was not. With a false discovery rate *p* value correction, two of these within-person effects remained significant, while the effect for caregiver internalizing symptoms became marginally significant (*p* = .063).

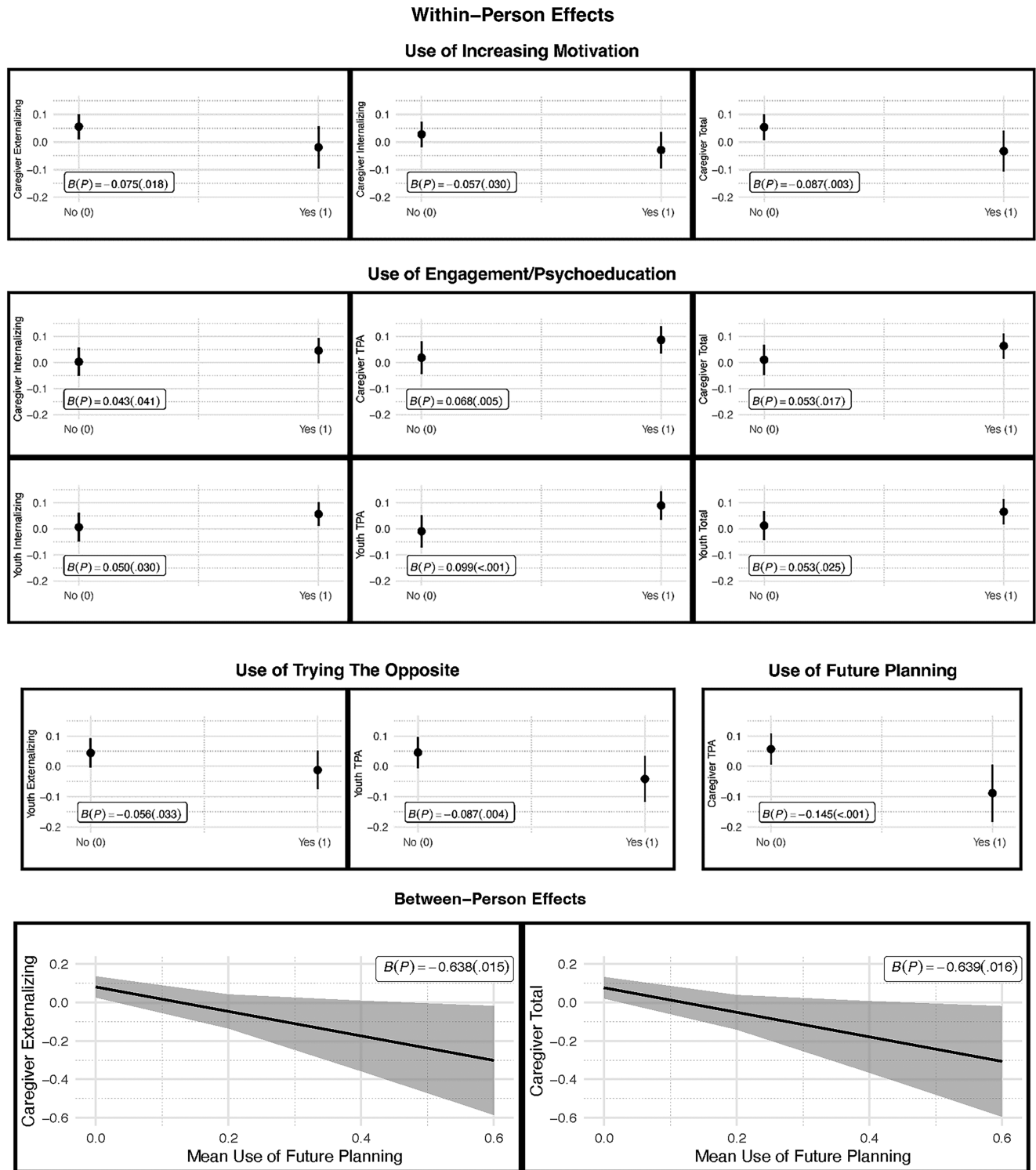
Within-person associations, between trying the opposite and subsequent symptom scores were significant (*p* < .05) across two outcomes. On measures of youth-rated externalizing symptoms and

Table 5*Between- and Within-Person Relations Between Modules Grouped by Treatment Principles and Youth-Report Symptoms*

Variable	Youth internalizing		Youth externalizing		Youth total problem		Youth top problems' assessment	
	Slope <i>B</i> (<i>SE</i>)	<i>p</i>	Slope <i>B</i> (<i>SE</i>)	<i>p</i>	Slope <i>B</i> (<i>SE</i>)	<i>p</i>	Slope <i>B</i> (<i>SE</i>)	<i>p</i>
Feeling calm within-person	0.011 (.035)	.764	−0.016 (.040)	.692	0.000 (.041)	.991	−0.078 (.041)	.057
Feeling calm between-person	0.298 (.211)	.159	0.252 (.227)	.268	0.323 (.216)	.135	−0.013 (.246)	.958
Increasing motivation (within)	−0.018 (.026)	.501	−0.033 (.030)	.269	−0.027 (.027)	.309	−0.052 (.034)	.128
Increasing motivation (between)	−0.090 (.082)	.272	−0.071 (.088)	.418	−0.091 (.082)	.269	0.076 (.095)	.423
Repairing thoughts (within)	−0.042 (.035)	.229	−0.008 (.039)	.837	−0.024 (.037)	.525	−0.023 (.044)	.607
Repairing thoughts (between)	−0.259 (.191)	.176	−0.138 (.207)	.507	−0.203 (.197)	.303	−0.272 (.223)	.224
Solving problems (within)	0.013 (.030)	.659	0.014 (.029)	.614	0.016 (.028)	.568	−0.035 (.032)	.279
Solving problems (between)	0.158 (.170)	.355	0.042 (.184)	.819	0.093 (.176)	.598	−0.159 (.199)	.424
Trying the opposite (within)	−0.024 (.026)	.369	−0.056 (.026)	.033*	−0.049 (.027)	.066	−0.087 (.030)	.004*
Trying the opposite (between)	0.118 (.121)	.327	−0.019 (.132)	.889	0.085 (.124)	.497	0.025 (.141)	.858
Engagement/psychoeducation (within)	0.050 (.023)	.030*	0.044 (.024)	.074	0.053 (.024)	.025*	0.099 (.026)	<.001*
Engagement/psychoeducation (between)	0.061 (.079)	.439	0.044 (.088)	.613	0.050 (.082)	.546	0.003 (.094)	.971
Future planning (within)	−0.027 (.043)	.530	0.007 (.039)	.859	−0.015 (.041)	.713	−0.030 (.051)	.549
Future planning (between)	−0.346 (.240)	.151	−0.281 (.262)	.284	−0.324 (.248)	.193	−0.057 (.284)	.840

Note. The values shown in this table are the beta values for the association between weekly use (yes or no) of each treatment principle and the closest subsequent outcome (standardized). All models control for the baseline (defined as session one) symptom score on the outcome measure and for time in the form of session number (see Curran & Bauer, 2011 for more on detrending) and include a random intercept and slope term. These *p* values are uncorrected, corrected values are in text. More negative *B* values indicate better treatment outcomes immediately after use of the principle. *SE* = standard error.

* *p* < .05.

Figure 1*Effects Plots for Significant Relations Between Treatment Modules Organized by Principles and Outcomes*

Note. These plots show graphical depictions of the associations presented in text and in Tables 4 and 5 between weekly use (yes or no) of each treatment principle and the closest subsequent outcome (standardized). Weekly principle use was coded as a factor for the purposes of this illustration. As is the case above, more negative B values indicate better treatment outcomes immediately after use of the principle.

idiographic top problems, scores were lower (i.e., better) on weeks when trying the opposite was used relative to when it was not ($B_s = -.087$ to $-.056$; $p_s < .05$). When a false discovery rate p value correction was applied, the within-person effects remained significant for youth-rated idiographic top problems and were just short of the threshold ($p = .074$) for youth-rated externalizing symptoms.

For three of the treatment principles (i.e., feeling calm, repairing thoughts, and solving problems) no between-person or within-person effects were statistically significant. In additional analyses of only those with elevated internalizing symptoms ($n = 385$) or, separately, externalizing ($n = 352$) symptoms, results were generally similar in direction, but some effects shifted across the threshold for significance, likely in part due to the smaller sample size (see [Supplemental Material](#)). Similarly, results of models of youth ($n = 358$) and caregiver ($n = 403$) outcomes for youth with presession one symptom scores generally were similar to results of main models, with some shifts across the significance threshold especially for trying the opposite, likely also in part due to the smaller and different sample.

Findings from models of individual modules within principles are presented in the [Supplemental Material](#). In summary, within increasing motivation, use of active ignoring and praise predicted lower (i.e., better; $p < .05$) youth-rated externalizing symptoms. Within trying the opposite, practicing (i.e., exposure) predicted better outcomes ($p < .05$) on many measures while trauma narrative significantly ($p < .05$) predicted one worse youth-report outcome. Within engagement/psychoeducation, within-person use of getting acquainted and engaging parents predicted worse scores ($p < .05$) across many outcomes. Within future planning, wrap up significantly ($p < .05$) predicted better outcomes on several measures and maintenance did for caregiver-rated internalizing symptoms.

Discussion

Within- and between-person use of seven common principles in youth psychotherapies were tested for associations with treatment outcomes, using a data set representing 490 youths (5,403 treatment sessions) in six RCTs of the flexible, modular MATCH psychotherapy (Harmon et al., 2021; Merry et al., 2020; Weisz, Bearman, et al., 2020; Weisz et al., 2012, 2018; Weisz, Thomassin, et al., 2020). The models were designed to account for temporality, previous symptom scores, and potential third-variable confounds at the subject-level, to shed light on module-outcome associations in a context in which random assignment to modules would be infeasible and unrepresentative of actual psychotherapy. Within-person use of increasing motivation, trying the opposite, and future planning predicted significantly better symptom scores, whereas within-person use of engagement/psychoeducation predicted significantly worse symptom scores. Feeling calm, repairing thoughts, and solving problems showed no significant associations with outcomes. Many of the significant effects remained significant after correction for multiple testing, suggesting robustness of study findings.

Some previous research has supported limited differential treatment element effects (Dour et al., 2013; Venturo-Conerly et al., 2022; Webb et al., 2019), while other studies have not and have underscored the difficulty of identifying active psychotherapy ingredients (Fitzpatrick et al., 2023; Lorenzo-Luaces, 2023). Still, few models of differential treatment element effects have been tested, and large sample sizes, sufficient variance element use, and complex models may be necessary to detect effects; such models can be

difficult to build without carefully cleaning and combining individual participant, session-by-session data from multiple studies. This study provided support for the ability of models that employ session-by-session data from several trials to identify differential treatment principle and module effects.

A distinctive feature of this study was the identification of both within-person and between-person effects. This added rigor by permitting us to adjust for potentially troublesome confounding. As one example, both significant within- and between-person effects were identified for use of future planning. This suggests that there were systematic differences between those who received different amounts of future planning (between-person effects) and suggests that on weeks when future planning was received, outcomes differed from weeks when it was not received (within-person effects). By including both effects in one model, it is possible to adjust for confounding between-person differences that may influence relations between receipt of future planning and outcomes, isolating the effect of future planning on weeks when it is used. Future planning was associated with significantly better scores on two caregiver-report measures between-person, and on one caregiver-report measure within-person. The significant between-person effect suggests that receiving more future planning may be associated with better symptom scores across treatment, perhaps because those who terminate treatment in a planned fashion (as opposed to dropping out early) are more likely to receive future planning, which generally comes late in MATCH. Importantly, because there was also a significant within-person effect identified, using future planning may also lead to better outcomes not only for such artifactual reasons, but also because it is particularly effective, maybe because it is important for initiating and maintaining use of treatment skills in everyday life. Overall, therefore, reviewing and planning for use of skills may be important for symptom scores, and it seems that similarly important for scores throughout treatment is completing a full, planned treatment course.

No other significant between-person effects were identified in this study, indicating that the person-level mean amount of each treatment principle (operationalized by grouping modules) other than future planning were not significantly associated with treatment outcomes. This could be for many reasons. For instance, it is possible that effects between-person are less consistent across youth than those effects within-person; reasons for receiving more or less of a certain treatment element, and the effectiveness of that element overall across treatment, may vary considerably across people. It is also possible that insufficient variance in outcomes or use of treatment elements across treatment reduced ability to detect between-person effects, but this seems unlikely in this particular study.

In this study, use of engagement/psychoeducation predicted worse subsequent symptoms. In theory this might reflect confounding (i.e., clinicians might work harder to build a relationship with more symptomatic youths, or on weeks when symptoms are more severe). However, the inclusion of (a) baseline symptoms as a covariate, (b) an autoregressive covariance term accounting for previous symptoms, and (c) a between-person effect designed specifically to control for confounding caused by subject-level or therapist-subject-relationship variables (e.g., a poorer therapeutic relationship) makes this explanation less likely. Instead, it seems that treatment outcomes were worse on weeks after engagement/psychoeducation was used relative to when it was not; this effect appeared to be driven most by “getting to know you” activities/discussions rather than psychoeducation, because only

the getting acquainted and engaging parents modules predicted worse scores ($p < .05$).

Considering past research supporting the importance of the therapeutic relationship (Lambert & Barley, 2001), we should stress that this finding will require replication and that it certainly does not indicate that building strong engagement or a strong relationship is harmful. It is possible, though, that an emphasis on engagement/psychoeducation that reduces therapists' emphasis on helping youths develop therapeutic skills may not be ideal. Consistent with this notion, we found that greater use of engagement/psychoeducation was highly negatively correlated with use of skills-based principles ($r = -.62$; $p < .001$), as was use of the modules getting acquainted and engaging parents ($r = -.51$; $p < .001$): more time spent building engagement was associated with less time spent building therapeutic skills. Of course, engagement building and skill building can be done in concert; youths may value a relationship through which they are acquiring useful skills. But our data suggest the possibility that when the two compete, treatment benefit may be undermined.

Additionally, the trying the opposite and increasing motivation principles were associated with better within-person weekly outcomes, suggesting that replacing unhelpful behaviors with more helpful behaviors and using contingencies to encourage desired behaviors and discourage undesired behaviors may be especially effective for immediate outcomes. Trying the opposite was associated only with better youth-rated outcomes, whereas increasing motivation was associated only with better caregiver-rated outcomes; this could be because caregivers are the primary agents implementing increasing motivation strategies, while youths often use trying the opposite with less caregiver involvement. The primary agents implementing a strategy may be more likely to report changes related to that strategy because of a combination of increased sensitivity to changes, more immediate awareness of changes in the short term, and expectancy effects which may influence their desire to experience and report changes. Additionally, interpreting results of specific symptom measures must be done cautiously because of overlap among symptom types and differences in caregiver and youth experiences of similar symptoms. Still, use of increasing motivation was associated significantly with lower levels of multiple types of symptoms, likely because modules grouped within increasing motivation are often used in practice to motivate change among youth with both internalizing and externalizing problems. Trying the opposite was associated with changes in externalizing and idiographic top problems, but it is still somewhat surprising that it was not associated with changes in internalizing problems; perhaps it more often results in changes over several sessions, with repeated use leading to gradual symptom reduction.

Within the trying the opposite principle, practicing (i.e., exposure) appeared to be particularly promising, while, surprisingly, use of trauma narrative (i.e., written exposure for trauma, used infrequently) was associated with worse outcomes. These results are generally consistent with past research supporting the strong effects of exposure (Parker et al., 2018), but trauma narrative (a form of exposure) may not have been used optimally, or may be associated with worse outcomes only in the short term, with improvements over multiple sessions, which the study methodology was not designed to detect. Among modules within increasing motivation, praise and active ignoring were associated with the strongest beneficial effects, suggesting that praising desired behaviors and

ignoring minor undesired behaviors may be particularly effective ways of increasing motivation. Of note, both principles (trying the opposite and increasing motivation) are relatively behavioral compared with others (i.e., repairing thoughts, solving problems, and feeling calm), consistent with a previous study (Webb et al., 2019) suggesting that behavioral approaches may outperform cognitive for youth depression. Similarly, these results are consistent with idiographic analyses showing especially strong effects of behavioral activation within a small sample of depressed youth receiving MATCH (Frederick et al., 2024).

This study has certain limitations. First, despite intensive efforts to control for potential confounds such as inclusion of between- and within-person effects, without randomization it will always be difficult to know with certainty that some undetected third variables did not influence findings. Relatedly, timing of use differed across modules, and although timing was accounted for, one cannot be certain that timing did not influence findings. Second, data on nonsymptom outcomes (e.g., therapeutic relationship quality) were not available; some modules such as those grouped within the engagement/psychoeducation principle might have untested effects on such outcomes. Additionally, mean scores at baseline on the clinical symptom measures were relatively low (not in the clinical range), which is likely a feature of the symptom scores among youth naturalistically recruited through community clinics into effectiveness trials, but also may limit ability to detect symptom reductions. Relatedly, as discussed in the supplement, not all measures had been fully validated for the full age ranges represented within the six trials included. Additionally, while the sample in many ways shows racial/ethnic demographics similar to the U.S. general population (*US Population by year, race, age, ethnicity, & more*, 2025), it includes small percentages of several racial/ethnic groups, especially Asian, Black/African American, and Hispanic/Latino people; findings are unlikely to be fully generalizable across demographic groups. Finally, all treatment module-outcome associations modeled in this study are short term, as is necessary for disentangling within- and between-person effects. Certain modules likely have different effects in the long term, when used in combination, or across multiple sessions, and this suggests a useful direction for future research.

Future research should involve testing how these short-term findings relate to longer term outcomes. Another important future direction will be testing the robustness of observed effects in other samples (e.g., for adults, in non-Western countries, or in therapy delivered by lay-providers). Relatedly, formal testing of moderators is outside the scope of this study, but will be important in future research; certain modules may be more effective at certain times during the course of therapy or for individuals with certain demographic or clinical characteristics. Indeed, past research suggests that individuals may respond quite differently to different modules, underscoring the importance of using clinical judgment and considering individual characteristics in addition to group-level findings when making treatment decisions (Frederick et al., 2024). Finally, although many common treatment procedures were included in this study, similar studies may be conducted with additional procedures (e.g., mindfulness) or using varying influential treatment variables other than module selection (e.g., therapist interpersonal styles; Anderson et al., 2016).

Overall, this study provided evidence for differential effects of different treatment procedures within youth psychotherapy and supported the possibility of detecting these differences when using

large samples with sufficient variance in treatment procedure use and rigorous analytic techniques. This study revealed favorable short-term clinical outcomes after use of future planning, trying the opposite, and increasing motivation, and unfavorable short-term clinical outcomes after use of engagement/psychoeducation. If evidence continues to accumulate in favor of more behavioral treatment approaches (e.g., exposure, behavioral activation, contingencies) for youth, clinicians may choose to emphasize those approaches more in their treatment of youth, particularly if they are uncertain about which type of treatment element is most likely to be effective for certain youth clients. Similarly, if evidence continues to suggest that use of exclusively engagement strategies over many sessions is associated with lower effectiveness, clinicians may select skills-based approaches when uncertain whether continued engagement or skills-based approaches would be best for a young client early in therapy. Broadly, there is no universal “best” decision in youth psychotherapy (Chorpita & Daleiden, 2018) and selection of treatment approaches will necessarily be influenced by many contextual factors. These findings, when combined with other related findings, clinically relevant theories, client outcomes, client preferences (Langer et al., 2022), and clinical judgment (Venturo-Conerly et al., 2023) may help guide clinician treatment decisions, particularly when clinicians are uncertain about which treatment procedures to emphasize with clients.

References

- Achenbach, T. M., & Rescorla, L. (2001). *Manual for the ASEBA school-age forms & profiles*. University of Vermont, Research Center for Children, Youth, & Families.
- Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G., & Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist*, 34(3), 341–382. <https://doi.org/10.1177/0011000005285875>
- Allison, P. (2015, June 3). *Don't put lagged dependent variables in mixed models*. Statistical Horizons.
- Anderson, T., McClintock, A. S., Himawan, L., Song, X., & Patterson, C. L. (2016). A prospective study of therapist facilitative interpersonal skills as a predictor of treatment outcome. *Journal of Consulting and Clinical Psychology*, 84(1), 57–66. <https://doi.org/10.1037/ccp0000060>
- Bennett, S. D., Cross, J. H., Chowdhury, K., Ford, T., Heyman, I., Coughtrey, A. E., Dalrymple, E., Byford, S., Chorpita, B., Fonagy, P., Moss-Morris, R., Reilly, C., Smith, J. A., Stephenson, T., Varadkar, S., Blackstone, J., Quartly, H., Hughes, T., Lewins, A., ... Shafran, R. (2024). Clinical effectiveness of the psychological therapy Mental Health Intervention for Children with Epilepsy in addition to usual care compared with assessment-enhanced usual care alone: A multicentre, randomised controlled clinical trial in the UK. *Lancet*, 403(10433), 1254–1266. [https://doi.org/10.1016/S0140-6736\(23\)02791-5](https://doi.org/10.1016/S0140-6736(23)02791-5)
- Chorpita, B. F., & Daleiden, E. L. (2018). Coordinated strategic action: Aspiring to wisdom in mental health service systems. *Clinical Psychology: Science and Practice*, 25(4), Article e12264. <https://doi.org/10.1111/cpsp.12264>
- Chorpita, B. F., Daleiden, E. L., Park, A. L., Ward, A. M., Levy, M. C., Cromley, T., Chiu, A. W., Letamendi, A. M., Tsai, K. H., & Krull, J. L. (2017). 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. *Journal of Consulting and Clinical Psychology*, 85(1), 13–25. <https://doi.org/10.1037/ccp0000133>
- Chorpita, B. F., Daleiden, E. L., & Weisz, J. R. (2005). Modularity in the design and application of therapeutic interventions. *Applied and Preventive Psychology*, 11(3), 141–156. <https://doi.org/10.1016/j.appsy.2005.05.002>
- Chorpita, B. F., Park, A., Tsai, K., Korathu-Larson, P., Higa-McMillan, C. K., Nakamura, B. J., Weisz, J. R., Krull, J., & the Research Network on Youth Mental Health. (2015). Balancing effectiveness with responsiveness: Therapist satisfaction across different treatment designs in the Child STEPs randomized effectiveness trial. *Journal of Consulting and Clinical Psychology*, 83(4), 709–718. <https://doi.org/10.1037/a0039301>
- Chorpita, B. F., Reise, S., Weisz, J. R., Grubbs, K., Becker, K. D., Krull, J. L., & the Research Network on Youth Mental Health. (2010). Evaluation of the Brief Problem Checklist: Child and caregiver interviews to measure clinical progress. *Journal of Consulting and Clinical Psychology*, 78(4), 526–536. <https://doi.org/10.1037/a0019602>
- Chorpita, B. F., & Weisz, J. R. (2009). *Modular approach to therapy for children with anxiety, depression, trauma, or conduct problems (MATCH-ADTC)*. PracticeWise.
- Cohen, Z. D., Delgadillo, J., & DeRubeis, R. J. (2021). Personalized treatment approaches. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change: 50th anniversary edition* (7th ed., pp. 673–703). John Wiley & Sons, Inc.
- Comer, J. S., Hong, N., Poznanski, B., Silva, K., & Wilson, M. (2019). Evidence base update on the treatment of early childhood anxiety and related problems. *Journal of Clinical Child and Adolescent Psychology*, 48(1), 1–15. <https://doi.org/10.1080/15374416.2018.1534208>
- Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, 62(1), 583–619. <https://doi.org/10.1146/annurev.psych.093008.100356>
- Dour, H. J., Chorpita, B. F., Lee, S., Weisz, J. R., & the Research Network on Youth Mental Health. (2013). Sudden gains as a long-term predictor of treatment improvement among children in community mental health organizations. *Behaviour Research and Therapy*, 51(9), 564–572. <https://doi.org/10.1016/j.brat.2013.05.012>
- Feeley, M., DeRubeis, R. J., & Gelfand, L. A. (1999). The temporal relation of adherence and alliance to symptom change in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*, 67(4), 578–582. <https://doi.org/10.1037/0022-006X.67.4.578>
- Fitzpatrick, O., Cho, E., Venturo-Conerly, K., Ugueto, A., Ng, M. Y., & Weisz, J. (2023). Examining principles of therapeutic change in youth psychotherapy: A meta-analysis. *Clinical Psychological Science*, 11(2), 326–344. <https://doi.org/10.1177/21677026221120230>
- Fonagy, P., Cottrell, D., Phillips, J., Bevington, D., Glaser, D., & Allison, E. (2015). *What works for whom? A critical review of treatments for children and adolescents* (2nd ed.). Guilford Press.
- Frederick, J., Ng, M. Y., Valente, M. J., Venturo-Conerly, K., & Weisz, J. R. (2024). What CBT modules work best for whom? Identifying subgroups of depressed youths by their differential response to specific modules. *Behavior Therapy*, 55(4), 898–911. <https://doi.org/10.1016/j.beth.2024.01.004>
- Funatogawa, I., & Funatogawa, T. (2012). An autoregressive linear mixed effects model for the analysis of unequally spaced longitudinal data with dose-modification. *Statistics in Medicine*, 31(6), 589–599. <https://doi.org/10.1002/sim.4456>
- Funatogawa, I., & Funatogawa, T. (2018). Autoregressive linear mixed effects models. In I. Funatogawa & T. Funatogawa (Eds.), *Longitudinal data analysis* (pp. 27–58). Springer. https://doi.org/10.1007/978-981-10-0077-5_2
- Goldfried, M. R. (2009). Searching for therapy change principles: Are we there yet? *Applied & Preventive Psychology*, 13(1–4), 32–34. <https://doi.org/10.1016/j.appsy.2009.10.013>
- Harmon, S. L., Price, M. A., Corteselli, K. A., Lee, E. H., Metz, K., Bonadio, F. T., Hersh, J., Marchette, L. K., Rodríguez, G. M., Raftery-Helmer, J.,

- Thomassin, K., Bearman, S. K., Jensen-Doss, A., Evans, S. C., & Weisz, J. R. (2021). Evaluating a modular approach to therapy for children with anxiety, depression, trauma, or conduct problems (MATCH) in school-based mental health care: Study protocol for a randomized controlled trial. *Frontiers in Psychology*, 12, Article 639493. <https://doi.org/10.3389/fpsyg.2021.639493>
- Kokla, M., Virtanen, J., Kolehmainen, M., Paananen, J., & Hanhineva, K. (2019). Random forest-based imputation outperforms other methods for imputing LC-MS metabolomics data: A comparative study. *BMC Bioinformatics*, 20(1), Article 492. <https://doi.org/10.1186/s12859-019-3110-0>
- Lambert, M. J., & Barley, D. E. (2001). Research summary on the therapeutic relationship and psychotherapy outcome. *Psychotherapy: Theory, Research, Practice, Training*, 38(4), 357–361. <https://doi.org/10.1037/0033-3204.38.4.357>
- Langer, D. A., Holly, L. E., Wills, C. E., Tompson, M. C., & Chorpita, B. F. (2022). Shared decision-making for youth psychotherapy: A preliminary randomized clinical trial on facilitating personalized treatment. *Journal of Consulting and Clinical Psychology*, 90(1), 29–38. <https://doi.org/10.1037/ccp0000702>
- Leijten, P., Weisz, J. R., & Gardner, F. (2021). Research strategies to discern active psychological therapy components: A scoping review. *Clinical Psychological Science*, 9(3), 307–322. <https://doi.org/10.1177/2167702620978615>
- Lorenzo-Luaces, L. (2023). Identifying active ingredients in cognitive-behavioral therapies: What if we didn't? *Behaviour Research and Therapy*, 168, Article 104365. <https://doi.org/10.1016/j.brat.2023.104365>
- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. University of Minnesota Press.
- Meehl, P. E. (1986). Causes and effects of my disturbing little book. *Journal of Personality Assessment*, 50(3), 370–375. https://doi.org/10.1207/s15327752jpa5003_6
- Merry, S. N., Hopkins, S., Lucassen, M. F. G., Stasiak, K., Weisz, J. R., Frampton, C. M. A., Bearman, S. K., Ugueto, A. M., Herren, J., Cribb-Su'a, A., Kingi-Uluave, D., Loy, J., Hartdegen, M., & Crengle, S. (2020). Effect of clinician training in the modular approach to therapy for children vs usual care on clinical outcomes and use of empirically supported treatments: A randomized clinical trial. *JAMA Network Open*, 3(8), Article e2011799. <https://doi.org/10.1001/jamanetworkopen.2020.11799>
- Mulder, R., Murray, G., & Rucklidge, J. (2017). Common versus specific factors in psychotherapy: Opening the black box. *The Lancet: Psychiatry*, 4(12), 953–962. [https://doi.org/10.1016/S2215-0366\(17\)30100-1](https://doi.org/10.1016/S2215-0366(17)30100-1)
- Ng, M. Y., & Weisz, J. R. (2016). Annual Research Review: Building a science of personalized intervention for youth mental health. *Journal of Child Psychology and Psychiatry*, 57(3), 216–236. <https://doi.org/10.1111/jcpp.12470>
- Parker, Z. J., Waller, G., Duhne, P. G. S., & Dawson, J. (2018). The role of exposure in treatment of anxiety disorders: A meta-analysis. *International Journal of Psychology & Psychological Therapy*, 18(1), 111–141.
- Piper, B. J., Gray, H. M., Raber, J., & Birkett, M. A. (2014). Reliability and validity the brief problem monitor, an abbreviated form of the Child Behavior Checklist. *Psychiatry and Clinical Neurosciences*, 68(10), 759–767. <https://doi.org/10.1111/pcn.12188>
- Scott, K., & Lewis, C. C. (2015). Using measurement-based care to enhance any treatment. *Cognitive and Behavioral Practice*, 22(1), 49–59. <https://doi.org/10.1016/j.cbpra.2014.01.010>
- Southam-Gerow, M. A., & Prinstein, M. J. (2014). Evidence base updates: The evolution of the evaluation of psychological treatments for children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, 43(1), 1–6. <https://doi.org/10.1080/15374416.2013.855128>
- Stekhoven, D. J., & Bühlmann, P. (2012). MissForest—Non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 28(1), 112–118. <https://doi.org/10.1093/bioinformatics/btr597>
- US population by year, race, age, ethnicity, & more. (2025, March 13). USAFacts. <https://usafacts.org/data/topics/people-society/population-and-demographics/our-changing-population/>
- Venturo-Conerly, K. E., Osborn, T. L., Alemu, R., Roe, E., Rodriguez, M., Gan, J., Arango, S., Wasil, A., Wasanga, C., & Weisz, J. R. (2022). Single-session interventions for adolescent anxiety and depression symptoms in Kenya: A cluster-randomized controlled trial. *Behaviour Research and Therapy*, 151, Article 104040. <https://doi.org/10.1016/j.brat.2022.104040>
- Venturo-Conerly, K. E., Reynolds, R., Clark, M., Fitzpatrick, O. M., & Weisz, J. R. (2023). Personalizing youth psychotherapy: A scoping review of decision-making in modular treatments. *Clinical Psychology: Science and Practice*, 30(1), 45–62. <https://doi.org/10.1037/cps0000130>
- Wang, L. P., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods*, 20(1), 63–83. <https://doi.org/10.1037/met0000030>
- Ward, A. M., Regan, J., Chorpita, B. F., Starace, N., Rodriguez, A., Okamura, K., Daleiden, E. L., Bearman, S. K., Weisz, J. R., & the Research Network on Youth Mental Health. (2013). Tracking evidence based practice with youth: Validity of the MATCH and standard manual consultation records. *Journal of Clinical Child and Adolescent Psychology*, 42(1), 44–55. <https://doi.org/10.1080/15374416.2012.700505>
- Webb, C. A., Stanton, C. H., Bondy, E., Singleton, P., Pizzagalli, D. A., & Auerbach, R. P. (2019). Cognitive versus behavioral skills in CBT for depressed adolescents: Disaggregating within-patient versus between-patient effects on symptom change. *Journal of Consulting and Clinical Psychology*, 87(5), 484–490. <https://doi.org/10.1037/ccp0000393>
- Weersing, V. R., Jeffreys, M., Do, M. T., Schwartz, K. T. G., & Bolano, C. (2017). Evidence base update of psychosocial treatments for child and adolescent depression. *Journal of Clinical Child and Adolescent Psychology*, 46(1), 11–43. <https://doi.org/10.1080/15374416.2016.1220310>
- Weisz, J. R., & Bearman, S. K. (2020). *Principle-guided psychotherapy for children and adolescents: The FIRST program for behavioral and emotional problems*. Guilford Press.
- Weisz, J. R., Bearman, S. K., Ugueto, A. M., Herren, J. A., Evans, S. C., Cheron, D. M., Alleyne, A. R., Weissman, A. S., Tweed, J. L., Pollack, A. A., Langer, D. A., Southam-Gerow, M. A., Wells, K. C., & Jensen-Doss, A. (2020). Testing robustness of child STEPs effects with children and adolescents: A randomized controlled effectiveness trial. *Journal of Clinical Child and Adolescent Psychology*, 49(6), 883–896. <https://doi.org/10.1080/15374416.2019.1655757>
- Weisz, J. R., & Chorpita, B. F. (2012). “Mod squad” for youth psychotherapy: Restructuring evidence-based treatment for clinical practice. In P. C. Kendall (Ed.), *Child and adolescent therapy: Cognitive-behavioral procedures* (4th ed., pp. 379–397). Guilford Press.
- Weisz, J. R., Chorpita, B. F., Frye, A., Ng, M. Y., Lau, N., Bearman, S. K., Ugueto, A. M., Langer, D. A., Hoagwood, K. E., & the Research Network on Youth Mental Health. (2011). Youth Top Problems: Using idiographic, consumer-guided assessment to identify treatment needs and to track change during psychotherapy. *Journal of Consulting and Clinical Psychology*, 79(3), 369–380. <https://doi.org/10.1037/a0023307>
- Weisz, J. R., Chorpita, B. F., Palinkas, L. A., Schoenwald, S. K., Miranda, J., Bearman, S. K., Daleiden, E. L., Ugueto, A. M., Ho, A., Martin, J., Gray, J., Alleyne, A., Langer, D. A., Southam-Gerow, M. A., Gibbons, R. D., & the Research Network on Youth Mental Health. (2012). Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: A randomized effectiveness trial. *Archives of General Psychiatry*, 69(3), 274–282. <https://doi.org/10.1001/archgenpsychiatry.2011.147>
- Weisz, J. R., Krumholz, L. S., Santucci, L., Thomassin, K., & Ng, M. Y. (2015). Shrinking the gap between research and practice: Tailoring and testing youth psychotherapies in clinical care contexts. *Annual Review of Clinical Psychology*, 11(1), 139–163. <https://doi.org/10.1146/annurev-clinpsy-032814-112820>

- Weisz, J. R., Kuppens, S., Ng, M. Y., Eckshtain, D., Ugueto, A. M., Vaughn-Coaxum, R., Jensen-Doss, A., Hawley, K. M., Krumholz Marchette, L. S., Chu, B. C., Weersing, V. R., & Fordwood, S. R. (2017). What five decades of research tells us about the effects of youth psychological therapy: A multilevel meta-analysis and implications for science and practice. *American Psychologist*, 72(2), 79–117. <https://doi.org/10.1037/a0040360>
- Weisz, J. R., Kuppens, S., Ng, M. Y., Vaughn-Coaxum, R. A., Ugueto, A. M., Eckshtain, D., & Corteselli, K. A. (2019). Are psychotherapies for young people growing stronger? Tracking trends over time for youth anxiety, depression, attention-deficit/hyperactivity disorder, and conduct problems. *Perspectives on Psychological Science*, 14(2), 216–237. <https://doi.org/10.1177/1745691618805436>
- Weisz, J. R., Thomassin, K., Hersh, J., Santucci, L. C., MacPherson, H. A., Rodriguez, G. M., Bearman, S. K., Lang, J. M., Vanderploeg, J. J., Marshall, T. M., Lu, J. J., Jensen-Doss, A., & Evans, S. C. (2020). Clinician training, then what? Randomized clinical trial of child STEPs psychotherapy using lower-cost implementation supports with versus without expert consultation. *Journal of Consulting and Clinical Psychology*, 88(12), 1065–1078. <https://doi.org/10.1037/ccp0000536>
- Weisz, J. R., Ugueto, A. M., Herren, J., Marchette, L. K., Bearman, S. K., Lee, E. H., Thomassin, K., Alleyne, A., Cheron, D. M., Tweed, J. L., Hersh, J., Raftery-Helmer, J. N., Weissman, A. S., & Jensen-Doss, A. (2018). When the torch is passed, does the flame still burn? Testing a “train the supervisor” model for the Child STEPs treatment program. *Journal of Consulting and Clinical Psychology*, 86(9), 726–737. <https://doi.org/10.1037/ccp0000331>
- Weisz, J. R., Vaughn-Coaxum, R. A., Evans, S. C., Thomassin, K., Hersh, J., Ng, M. Y., Lau, N., Lee, E. H., Raftery-Helmer, J. N., & Mair, P. (2020). Efficient monitoring of treatment response during youth psychotherapy: The behavior and feelings survey. *Journal of Clinical Child and Adolescent Psychology*, 49(6), 737–751. <https://doi.org/10.1080/15374416.2018.1547973>

Received September 13, 2024

Revision received June 25, 2025

Accepted June 27, 2025 ■