


Special Issue Article

Identifying intervention strategies for preventing the mental health consequences of childhood adversity: A modified Delphi study

Leslie R. Rith-Najarian¹ , Noah S. Triplett², John R. Weisz¹ and Katie A. McLaughlin¹

¹Department of Psychology, Harvard University, Cambridge, MA, USA and ²Department of Psychology, University of Washington, Seattle, WA, USA

Abstract

Exposure to childhood adversity is a powerful risk factor for psychopathology. Despite extensive efforts, we have not yet identified effective or scalable interventions that prevent the emergence of mental health problems in children who have experienced adversity. In this modified Delphi study, we identified intervention strategies for effectively targeting both the neurodevelopmental mechanisms linking childhood adversity and psychopathology – including heightened emotional reactivity, difficulties with emotion regulation, blunted reward processing, and social information processing biases, as well as a range of psychopathology symptoms. We iteratively synthesized information from experts in the field and relevant meta-analyses through three surveys, first with experts in intervention development, prevention, and childhood adversity ($n = 32$), and then within our study team ($n = 8$). The results produced increasing stability and good consensus on intervention strategy recommendations for specific neurodevelopmental mechanisms and symptom presentations and on strength of evidence ratings of intervention strategies targeting youth and parents. More broadly, our findings highlight how intervention decision making can be informed by meta-analyses, enhanced by aggregate group feedback, saturated before consensus, and persistently subjective or even contradictory. Ultimately, the results converged on several promising intervention strategies for prevention programming with adversity-exposed youth, which will be tested in an upcoming clinical trial.

Keywords: childhood adversity, Delphi study, intervention development, prevention, psychopathology

(Received 29 October 2020; accepted 31 October 2020)

Introduction

Exposure to adversity is common in children, both in the USA and internationally. Approximately half of all US children will experience at least one form of adversity – such as physical, sexual, or emotional abuse, neglect, exposure to interpersonal violence, or chronic poverty – by the time they reach adulthood (Benjet et al., 2009; Green et al., 2010; Kessler et al., 2010; McLaughlin et al., 2012). Evidence from population-based and longitudinal studies demonstrates that experiences of adversity are associated with the onset of internalizing and externalizing psychopathology transdiagnostically, including anxiety, mood, disruptive behavior, posttraumatic stress, and substance use disorders (Cohen, Brown, & Smiles, 2001; Green et al., 2010; Kessler et al., 2010; MacMillan et al., 2001; McLaughlin, Conron, Koenen, & Gilman, 2010; McLaughlin et al., 2012; Weich, Patterson, Shaw, & Stewart-Brown, 2009). Exposure to childhood adversity explains about a third of mental disorder onsets in the USA (Green et al., 2010; McLaughlin et al., 2012) and is associated with increased risk for a wide range of chronic diseases and early mortality (Brown et al., 2009; Chen, Turiano, Mroczek, &

Miller, 2016; Felitti et al., 1998; Wegman & Stetler, 2009). A substantial body of work has characterized the developmental mechanisms through which experiences of adversity influence risk for psychopathology and chronic diseases (Cicchetti & Toth, 1995; Danese & McEwen, 2012; McLaughlin, 2016; McLaughlin, Colich, Rodman, & Weissman, 2020; McLaughlin & Lambert, 2017). Although knowledge of these mechanisms has grown dramatically, little progress has been made in translating those mechanisms into interventions to prevent the onset of psychopathology on subsequent consequences (McLaughlin, DeCross, Jovanovic, & Tottenham, 2019). Through our current work, we hope to carry on Edward Zigler's legacy of “be[ing] totally committed to the optimal development of each child” by developing such an intervention for children who have screened positive for adversity. In contrast to the scarcity of research available prior to the development of Head Start, considered “the prototype of effective early childhood intervention,” we are fortunate to now be able to draw on the substantial evidence base supporting early childhood intervention (Zigler & Styfco, 2001).

In theory, screening and early intervention could help adversity-exposed children by preventing future incidences of adversity along with its physical and mental health consequences. We agree with the belief that children who begin life healthy “grow to become contributing members of the society, [and] the small investment made in their early years will have compounded to reap a handsome dividend” (Zigler, 1996, p. 47). A 10% reduction in childhood adversity prevalence in the USA has been estimated to save \$105 billion annually (Bellis et al.,

Author for Correspondence: Leslie Rith-Najarian, Department of Psychology, Harvard University, Room 1006, William James Hall, 33 Kirkland St, Cambridge, MA 02138, USA; E-mail: rithnajarian@fas.harvard.edu

Cite this article: Rith-Najarian LR, Triplett NS, Weisz JR, McLaughlin KA (2021). Identifying intervention strategies for preventing the mental health consequences of childhood adversity: A modified Delphi study. *Development and Psychopathology* 33, 748–765. <https://doi.org/10.1017/S0954579420002059>

2019). The broader the scale of screening and prevention efforts, the greater the potential impact becomes. Pediatric primary care is a promising setting for wide-reaching adversity screening and prevention efforts, as the majority of US children see a primary care provider annually (Chevarley, 2003). Moreover, addressing child mental health concerns in a medical setting has myriad benefits, including reduced stigma, cost, and family discomfort, as well as improved outcomes (Asarnow, Rozenman, Wiblin, & Zeltzer, 2015; Godoy et al., 2017; Kolko & Perrin, 2014). Primary care settings are already increasingly integrating youth mental health care (Aupont et al., 2013), with approximately half of all visits now involving some type of behavioral or mental health concern (Martini et al., 2012). Simultaneously, there has been a movement encouraging pediatricians to screen for adversity within routine primary care visits (Garner et al., 2012; Pardee, Kuzma, Dahlem, Boucher, & Darling-Fisher, 2017). One encouraging example in primary care settings is the Safe Environment for Every Kid (SEEK) program (Dubowitz, 2014; Dubowitz, Feigelman, Lane, & Kim, 2009), which includes a validated short adversity screening tool and a brief intervention (5–15 min) of motivational interviewing and community resource referrals for families who endorse adversity experiences (Dubowitz et al., 2007; Dubowitz, Prescott, Feigelman, Lane, & Kim, 2008; Feigelman et al., 2009; Lane et al., 2007). Randomized controlled trials of SEEK have shown reduced rates of future child maltreatment (Dubowitz et al., 2009; Dubowitz, Lane, Semiatin, & Magder, 2012), and feasibility studies have shown high acceptability and satisfaction among providers and caregivers in primary care settings (e.g., Eismann, Theuerling, Maguire, Hente, & Shapiro, 2019). However, despite all this progress, one concern about adversity screening is the dearth of interventions that can prevent the mental health outcomes associated with adversity exposure (Campbell, 2020). This scarcity is problematic, as it has been recommended that youth receive evidence-based prevention strategies that are matched to their screened risks for psychopathology (Hankin, 2020). There thus remains a major challenge for adversity screening efforts in primary care or elsewhere: we must respond to those youth identified as adversity-exposed and yet we lack effective strategies for preventing the associated psychopathology.

The absence of effective preventive interventions for children exposed to adversity does not reflect a lack of research attention to this issue. Dozens of psychosocial prevention and early intervention programs for maltreated children have been developed and tested, but meta-analyses examining them have not produced substantive evidence of intervention effectiveness (Churchill, 2016; van IJzendoorn, Bakermans-Kranenburg, Coughlan, & Reijman, 2020). There are some effective interventions for children experiencing psychopathology after exposure to trauma (e.g., abuse, exposure to violence), such as trauma-focused cognitive behavioral therapy (TF-CBT; Cohen, Deblinger, Mannarino, & Steer, 2004; Cohen, Mannarino, & Iyengar, 2011; Dorsey et al., 2017). However, these interventions are treatment level, and youth psychological treatments typically involve a large number of sessions (averaging 16.5 in a recent meta-analysis; Weisz, Kuppens, et al., 2017), which is less practical for typical behavioral health session limits. More feasible for integration into primary care would be, for example, a brief (four-session) skills-based intervention that has demonstrated effective prevention of post-traumatic stress disorder (PTSD) symptoms in children who have experienced trauma (Berkowitz, Stover, & Marans, 2011). Nevertheless, adversity experiences are associated not only with

PTSD but also with transdiagnostic psychopathology, including a wide range of internalizing and externalizing problems (Green et al., 2010; Keyes et al., 2012; McLaughlin et al., 2012). Furthermore, such existing evidence-based interventions were designed as treatments for children who have already developed symptoms of psychopathology following traumatic events and thus primarily focus on symptom reduction rather than preventing the emergence of psychopathology. This distinction is important as the mechanisms that underlie symptom maintenance may differ from those that contribute to the emergence of psychopathology following experiences of adversity. As such, prevention practices that target the developmental pathways linking adversity to the onset of psychopathology – particularly mechanisms that are associated with transdiagnostic risk – may ultimately hold the most promise.

Many have argued that greater attention to the neurodevelopmental pathways linking childhood adversity with psychopathology is needed in order to improve our ability to effectively intervene with children who have encountered adversity (Bentovim et al., 2020; Cicchetti & Toth, 2009; Manly, Kim, Rogosch, & Cicchetti, 2001; Masten & Cicchetti, 2010; McLaughlin et al., 2019; van IJzendoorn et al., 2020). Over the past decade, a substantial body of evidence has emerged on these mechanisms. As noted in recent reviews of this literature, developmental pathways that have garnered the most consistent evidence for being influenced by experiences of adversity and that, in turn, are associated with the later onset of psychopathology include threat-related social information processing biases (e.g., hostile attribution bias), heightened emotional reactivity, difficulties with emotion regulation, and blunted reward processing (McLaughlin, 2016; McLaughlin et al., 2019). Threat-related social information processing biases have been observed most consistently among children exposed to trauma and forms of adversity involving threat (e.g., violence exposure, abuse), whereas blunted reward processing has been observed most consistently in children who have experienced forms of adversity involving deprivation, including neglect, caregiver separation, and food insecurity. Heightened emotional reactivity and difficulties with emotion regulation have been observed in children exposed to many forms of adversity (for a review, see McLaughlin et al., 2019). Correspondingly, McLaughlin et al. (2019) posit that these mechanisms could serve as important behavioral targets for interventions aimed at preventing psychopathology transdiagnostically. Zigler (2000) also suggested that research on brain development had important implications for early childhood interventions – specifically, that such research has made us more certain than ever that adverse environments alter processes such as learning and emotion regulation. Regardless of such suggestions, we know of no existing prevention program designed to specifically target the developmental processes that are influenced by exposure to adversity and, in turn, predict the later emergence of psychopathology.

Current Research

In this article we describe a series of studies investigating the development and testing of an intervention program through pediatric primary care settings to prevent the mental health consequences of exposure to childhood adversity. Our planned intervention will incorporate the evidence-based SEEK screening and case management approach (Dubowitz, 2014; Dubowitz et al., 2009). We will then extend services by offering a brief

Table 1. Specific intervention strategies within the FIRST principles

Feeling calm	Increasing motivation	Repairing thoughts	Solving problems	Trying the opposite
<ul style="list-style-type: none"> ■ Guided relaxation ■ Self-calming ■ Progressive muscle relaxation ■ Deep breathing ■ Guided imagery 	<ul style="list-style-type: none"> ■ Positive attending and praise ■ Differential attention ■ Effective instructions and house rules ■ Reinforcement with tangible rewards ■ Time-out and response cost ■ Self-reinforcement 	<ul style="list-style-type: none"> ■ Identifying unhelpful, distorted thoughts ■ Cognitive restructuring and thought challenging ■ Cognitive distraction ■ Attribution retraining 	<ul style="list-style-type: none"> ■ Identifying problems, conflicts, dilemmas ■ Generating a list of possible solutions ■ Weighing pros and cons to select best option 	<ul style="list-style-type: none"> ■ Exposure ■ Behavioral activation ■ Anger control skills ■ Role playing

intervention (four to six sessions) to families with children who screen positive for adversity exposure. As Zigler and colleagues have argued, parents and caregivers play a central role in children's behavioral development (Zigler, Pfannenstiel, & Seitz, 2008) and, correspondingly, their response to behavioral health interventions (Sun, Rith-Najarian, Williamson, & Chorpita, 2019). Accordingly, our brief intervention will introduce coping strategies for the children and adolescents, and parenting strategies for caregivers.

Though we ought not simply repackage treatment strategies for preventive purposes, we unfortunately lack strongly supported prevention strategies for adversity-exposed youth. In contrast, the development of new intervention strategies for youth mental health treatments seems to have reached saturation (Okamura *et al.*, 2020). As such, the vast array of intervention components from youth treatment programs may still provide some candidates for our prevention purposes, though other prevention strategies likely remain to be discovered through empirical research. Given that experiences of adversity are associated with psychopathology transdiagnostically (e.g., Keyes *et al.*, 2012), it is essential that intervention strategies are flexible enough to target a range of internalizing and externalizing problems and the variable needs of children and families. Accordingly, we selected two evidence-based interventions – each of which includes multiple intervention elements – from which to pull candidate intervention strategies: FIRST (Weisz & Bearman, 2020) and PCIT-ED (Luby, Barch, Whalen, Tillman, & Freedland, 2018).

Selection of candidate intervention strategies

FIRST (Feeling calm, Increasing motivation, Repairing thoughts, Solving problems, Trying the opposite)

FIRST is a good fit to our goals for several reasons. First, it is transdiagnostic, encompassing intervention procedures for an array of problems and concerns, including posttraumatic stress, anxiety, depressed mood, and conduct problems. Second, it is designed for personalization – for tailoring to fit the specific combination of needs of each individual child and family. Third, it is streamlined, efficient and accessible to clinicians, organizing many specific evidence-based procedures within five broad principles of change. These five principles are Feeling calm, Increasing motivation, Repairing thoughts, Solving problems, and Trying the opposite (see Table 1 for details). Clinicians using FIRST are encouraged to identify which of these principles may be most relevant to the child they are working with, and then apply that principle in their work with the child, drawing from the specific procedures grouped beneath that principle, as shown in Table 1.

Benchmarking trials of FIRST have shown rapid reductions in child problems and symptoms – assessed via standardized measures of internalizing and externalizing problems and via idiographic measures of the “top problems” most important to children and their caregivers – and with trajectories of improvement similar to those seen with markedly more complicated treatments that require substantially more training (Cho, Bearman, Woo, Weisz, & Hawley, 2020; Weisz, Bearman, Santucci, & Jensen-Doss, 2017).

PCIT (parent-child interaction therapy)

PCIT is an intervention focused on improving parent-child interactions that has demonstrated large and sustained effects for treating youth with disruptive behavior problems (Brestan & Eyberg, 1998; McNeil & Hembree-Kigin, 2010; Thomas & Zimmer-Gembeck, 2007). Critical for our purposes, PCIT has also been shown to reduce the rate of future maltreatment when delivered to parents (or caregivers) who had abused their children (Chaffin *et al.*, 2004; Timmer, Urquiza, Zebell, & McGrath, 2005). PCIT includes several intervention elements that are taught through behavioral observations of parent-child interactions and live coaching of these interactions by a therapist. PCIT includes two primary intervention components: child-directed interaction and parent-directed interaction. In child-directed interaction, parents learn a variety of techniques designed to enhance the parent-child relationship. These include techniques praise, reflecting, imitating, and describing the child's behavior, and providing contingent attention (i.e., attending to positive behaviors and ignoring negative behaviors). Parent-directed interaction provides parents with training in effective discipline strategies designed to increase child compliance and decrease disruptive behaviors. PCIT has been expanded to include additional elements designed to target internalizing problems (Carpenter, Puliafico, Kurtz, Pincus, & Comer, 2014). For our purposes, we focus on PCIT plus emotional development (PCIT-ED), which includes an additional module focused on parenting techniques to help children identify and regulate their emotional experiences (Lenze, Pautsch, & Luby, 2011). PCIT-ED has been shown to significantly reduce depression symptoms in both children and their caregivers (Luby *et al.*, 2018; Luby, Lenze, & Tillman, 2012). More importantly for prevention, another trial showed that, relative to a waitlist, PCIT-ED produced greater changes in reward processing, which were in turn associated with greater improvements in depression symptoms (Barch *et al.*, 2019). Thus, this research suggests that some interventions designed as treatments for youth psychopathology may also influence key developmental mechanisms involved in the link between adversity and psychopathology.

Study aims

The success of prevention and early intervention programs depends in large part on their design, and incorporated elements should be preventive as opposed to purely remedial (Zigler, 1996). On the one hand, many children and adolescents who screen positive for adversity exposure may already be symptomatic. Thus, these youth will likely be well served by intervention strategies from FIRST and PCIT-ED. On the other hand, many other youth who screen positive may not yet be experiencing clinically significant mental health symptoms. Therefore, we still need to identify which candidate intervention strategies are more likely to target not only symptoms but also the neurodevelopmental mechanisms that put a youth at risk of eventual symptom onset (i.e., threat-related social information processing biases, heightened emotional reactivity, difficulties with emotion regulation, and blunted reward processing; McLaughlin et al., 2019). Such an approach would be consistent with the American Psychological Association (APA) prevention guidelines to develop programs that reduce risks and promote human strengths, as opposed to strategies that focus only on symptom reduction (American Psychological Association, 2014).

To this end, we conducted a modified Delphi study to gather feedback on the intervention strategies that we selected as candidates for our prevention program for adversity-exposed youth. A lack of expert consensus about which youth mental health interventions are considered “evidence based” has been a longstanding issue related to the science-to-practice gap (Weisz, Sandler, Durlak, & Anton, 2005), and that issue is even more salient when developing a novel intervention. Consensus methods can enhance decision making in health research by systematically assessing expert recommendations for topics that have a dearth of scientific agreement or evidence (Trevelyan & Robinson, 2015). It has been recommended to combine the existing scientific evidence base with the current widespread professional norms (i.e., expert opinions) to maximize effective clinical care (Campbell, Roland, & Buetow, 2000). It has also been noted that, when using Delphi methods in mental health research, the quality of input from experts depends on the quality of input provided to them (e.g., systematic reviews, personal experience) (Jorm, 2015). Systematic reviews are a gold standard for assessing the current state of effectiveness for youth treatments and prevention programs (Weisz et al., 2005) and have been specifically recommended as a strong form of evidence for identifying effective components of child maltreatment interventions (Powell et al., 2015). We thus planned to synthesize feedback across survey rounds as informed by two sources of knowledge – experts and meta-analyses.

Our primary aim was to evaluate consensus on recommendations and strength of evidence ratings for the candidate intervention strategies. We recognize that reaching consensus on recommendations is not necessarily proof that those interventions are the “best” or “correct” recommendations (Trevelyan & Robinson, 2015). Still, even if there is not always a single best solution, it is helpful to prioritize options from a list of candidates (Hall, 2009). We also believe that a more extensive approach to planning intervention design is consistent with the cautionary notes of Zigler regarding “hasty” prevention development and implementation (Zigler, 1996; Zigler & Styfco, 2001). The findings of the current study could provide some foundational evidence to more strongly hypothesize which intervention strategies might be mechanism-targeting prevention approaches. The results of skills recommendations can be directly applied to

our intervention development and, later, their actual effects on mechanisms and symptoms empirically tested in future clinical trials of the prevention program. An exploratory additional aim was to gain insight into how evidence-informed decision making in intervention development unfolds.

Method

All procedures were reviewed and approved by the Institutional Review Board of Harvard University.

Overview of the Delphi approach

Traditional Delphi studies aim to gather expert input iteratively over multiple rounds until consensus is achieved, with the first round generating qualitative data that then informs the subsequent items (Hsu & Sandford, 2007). However, given that our aim was not necessarily to reach consensus but to assess for consensus, we conducted a modified Delphi study. First, we invited experts to recommend candidate intervention strategies based on the five broad principles from FIRST (Weisz & Bearman, 2020) and three parenting practices from PCIT-ED (Luby et al., 2018). Second, we invited a subset of experts – members of our broader study team – to identify relevant systematic reviews and then consider the strength of evidence for intervention strategies based on meta-analytic support for practice elements. Finally, our team reviewed anonymous and objective summaries of the recommendations from experts and systematic reviews together. This approach was designed to produce more informed conclusions than we could reach as individuals, while addressing some limitations of group decision making (e.g., influence of relative authority figures, tendency towards group conformity or polarization) (Hsu & Sandford, 2007; Surowiecki, 2004).

Survey 1

Participants

We identified experts whose research focuses on child and/or adolescent mental health in at least one of the following categories of relevance: developmental psychopathology research focused on mechanisms linking adversity with psychopathology; prevention science; intervention development or treatment-focused research; and dissemination and implementation science. Potential participants were identified through four sources: (a) the principal investigator (PI), coinvestigators (CIs), and established consultants of the intervention research project; (b) scholars nominated by these members of the research team as experts in the domains of interest; (c) members of relevant professional task forces and executive boards; and (d) PIs of relevant preventive intervention trials currently active on clinicaltrials.gov. Participants were eligible to participate if the majority of their work was in one of the categories of interest (as determined by their professional website or CV) and they were still active in their career (e.g., not retired, had published research in one of the categories of interest within the past 2 years). We focused on identifying experts across a range of populations served (e.g., racial and ethnic minorities, rural and urban populations, and so on). A list of potential participants was iteratively compiled and revised to limit redundancy and broaden representative coverage of expertise areas. Recruitment was conducted by the PI/CIs through email invitations that explained the purpose of the study, the rationale for expert selection, and

Table 2. Demographics of participants in Survey 1

	N	%
Age		
26–35 years	3	9.4
36–45 years	9	28.1
46–55 years	9	28.1
Over 55 years	11	34.4
Gender identity		
Female	18	56.3
Male	14	43.8
Other	0	0.0
Race ^a		
Asian	3	9.3
Black or African American	3	9.3
Native American or Alaska Native	1	3.1
White	24	75.0
Middle Eastern	2	6.3
Multiple racial groups	1	3.1
Hispanic or Latinx		
Yes	2	6.3
Highest education completed ^a		
Doctoral degree (e.g., PhD, PsyD)	28	88.0
Doctor of medicine (e.g., MD, DO)	5	15.6
Master's degree (e.g., MS, MSW) alone	1	3.1
Self-reported expertise area ^a		
Dissemination and implementation science	14	43.8
Prevention science	14	43.8
Mechanisms-focused research	18	56.3
Research on childhood adversity	18	56.3
Treatment-focused research	17	53.1

^aNote: Respondents could select multiples, so percentages may exceed 100%.

anonymity of participation. The study PI/CIs and previously hired consultants did not receive financial compensation, while all remaining invited experts were offered a payment of \$100.

In total, 46 experts, including the four PI/CIs, were invited to participate in the expert recommendations survey, of which 32 (69.6%) provided consent and completed the survey. Demographic data for the participating sample ($n = 32$) are provided in Table 2.

Procedure and measures

The participants responded to an online survey that assessed demographic information and asked them to recommend intervention strategies that would be most effective in targeting our four developmental mechanisms of interest. The survey also asked participants to recommend intervention strategies for four types of psychopathology symptoms that commonly occur in children exposed to adversity, including symptoms of disruptive behavior problems, depression, anxiety, and PTSD. As such,

intervention recommendations were provided for a total of eight targets – four about developmental mechanisms linking adversity and psychopathology, and four about mental health symptoms. In each subsection, the survey provided examples of adversity experiences that have been commonly associated with each developmental mechanism or symptom type, and included examples of how each developmental mechanism or symptom type might present clinically or subclinically (e.g., heightened emotional reactivity may present as rapid mood changes or being quickly consumed by intense emotions; depressive symptoms may present as sad mood or loss of pleasure in activities that used to be enjoyable). Examples of each developmental mechanism were taken from standardized measures or published instruments, including: the Behavioral Inhibition System/Behavioral Activation System scales for blunted reward processing (Carver & White, 1994); the Emotion Reactivity Scale (Nock, Wedig, Holmberg, & Hooley, 2008) for heightened emotional reactivity; the Children's Response Styles Questionnaire (Abela, Vanderbilt, & Rochon, 2004) and Zeman and Garber's (2016) emotion regulation vignette coding system for emotion regulation difficulties; and the Social Information Processing Application (Kupersmidt, Stelter, & Dodge, 2011) for social information processing biases. Examples for each of the mental health symptoms were adapted from the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) (American Psychiatric Association, 2013) to remove "clinical-level" qualifying language. Anxiety symptom examples were informed by the symptom criteria in the generalized anxiety disorder, panic disorder, social anxiety disorder, specific phobia, and separation anxiety disorder sections of the DSM-5. Depression symptom examples were informed by the symptom criteria in the major depressive disorder and persistent depressive disorder sections. Posttraumatic stress symptom examples were informed by the symptom criteria in the PTSD section. Disruptive behavior examples were informed by the symptom criteria in the conduct disorder and oppositional defiant disorder sections.

The survey instructed participants to recommend specific intervention strategies that could be used for children and adolescents who have experienced adversity and were exhibiting either a specific developmental mechanism or symptom pattern. Participants were told that all youth would have received case management prior to intervention and that the goal was to develop a preventive intervention rather than treatment. Participants were told to select a single answer they felt best addressed the specific presentation and were reminded that there were no correct answers. For all recommendations, participants were asked to select their primary recommendations and secondary recommendations (optional) for the most relevant intervention principles from FIRST. Participants were provided with a written description and a visual of FIRST and its five core treatment principles – Feeling calm, Increasing motivation, Repairing thoughts, Solving problems, and Trying the opposite. Participants were asked to make separate recommendations for younger (i.e., children aged 8–12 years) and older (i.e., adolescents aged 13–15 years) presentations. Participants were also asked if they would make additional recommendations beyond the FIRST principles and if their recommendations would change for children and adolescents of certain demographics. Finally, participants were asked to separately rank the three parenting practices – Increasing positive reinforcement, Decreasing caregiver criticism and hostility, and Coaching of child emotional competence – in order of importance for a brief preventive

intervention for adversity-exposed children and adolescents. Although training in specific discipline strategies is also a core element of PCIT, we did not include this in our list of parenting practices as these techniques are subsumed within the Increasing motivation principle of FIRST. Throughout the survey, participants were allowed to provide qualitative feedback or explanations for their responses.

Survey 2

Participants

Members of our research team (i.e., PI/CIs, staff, graduate students) were invited to participate via an email sent from the study lead (LRN). The recruitment email emphasized that participation was voluntary and no compensation was provided for completing the survey. Of the eight individuals who were invited to participate, six provided consent and completed the full survey. This survey intentionally took place after Survey 1 had closed to ensure that the study investigators provided their prior responses before systematically reviewing the relevant meta-analyses that were incorporated in Survey 2.

Literature search for relevant systematic reviews

Members of the study team identified meta-analyses through an informal literature search and article nomination process (i.e., no specified databases or mandatory key terms). The search aimed to identify recently published systematic reviews and meta-analyses on “common elements,” “practice elements,” or “dismantling studies” of clinical trials, with a particular focus on identifying articles that examined prevention programs, a range of internalizing/externalizing problems, a range of ages (i.e., early childhood through adolescence), and (when possible) samples of adversity-exposed youth, or interventions conducted through primary care gatekeeping. Included studies met the following criteria: (a) they examined the effects of prevention programs (at the universal, selective, or indicated level) and (b) they reported effect sizes by common elements or intervention strategies. Meta-analyses that examined both prevention and intervention programs were included if they reported effect sizes for prevention programs separately from treatment programs.

Round 1: Meta-analysis selection

The informal literature search yielded 15 potential meta-analyses. The study team members reviewed the identified articles and voted to select the final group of meta-analyses. They were instructed to vote for the meta-analyses that they felt were “most useful in informing” the design of a preventive intervention. They were told to consider relevance (i.e., prevention programs, special samples, age ranges), breadth of targets covered within and across meta-analyses, and the quality of analysis (e.g., meta-regression was considered highest quality). Meta-analyses were included if more than half of the reviewers ($n = 6$) voted for their inclusion, resulting in a final sample of seven meta-analyses (Caldwell et al., 2019; Filene, Kaminski, Valle, & Cachat, 2013; Leijten et al., 2019; Sanchez et al., 2018; Schleider & Weisz, 2017; Singla et al., 2020; van der Put, Assink, Gubbels, & Boekhout van Solinge, 2018).

Round 2: Survey data collection

Participants were provided with links and were instructed to read the final set of meta-analyses. In addition, the survey instructions summarized key information from each meta-analysis, including

a list of which age groups, settings, and clinical targets were reviewed, as well as key tables and figures. For each meta-analysis, participants were asked to rate the clinical practice elements/intervention components (e.g., psychoeducation, contingency management, relaxation) that had the strongest support indicated for treatment and prevention programs. Participants were also asked to indicate which of the five FIRST principles (Feeling calm, Increasing motivation, Repairing thoughts, Solving problems, and Trying the opposite) and three parenting practices (Increasing positive reinforcement, Decreasing caregiver criticism and hostility, and Coaching of child emotional competence) were supported by each meta-analysis. After responding to questions for each individual meta-analysis, participants were provided with a summary of their responses and a spreadsheet that aggregated all the effect sizes reported across the meta-analyses. Participants were asked to rank the top practice elements/intervention components that they felt had the “strongest evidence” (a) across all targets and treatment/prevention programs, (b) across all prevention programs only, and (c) and that were relevant for a “brief (two to six sessions) prevention-level service offered to adversity-exposed youth (age 8–15) and their families who were screened in via primary care.” Finally, participants were instructed to rate their perceptions of support for the FIRST principles and parenting practices on a 5-point Likert-type scale, ranging from “*extremely effective*” to “*not effective at all*.” For all questions there was an open-text field in which participants could write notes elaborating on their decision making. Those with a bachelor’s degree were both currently enrolled in a clinical psychology doctoral program.

Survey 3

Participants

The same eight individuals invited for Study 2 were invited in the same way for Study 3, all of whom provided consent and completed the full survey. These individuals were selected to participate because of their knowledge of the ultimate goal of the project – to develop a brief intervention for adversity-exposed youth. All participants had completed one or both of the prior surveys. This sample was 75% female, 88% white with one respondent identifying as biracial, and the highest educational degree levels were doctoral (50%), master’s (25%), and bachelor’s (25%) degrees. Those with a bachelor’s degree were both currently enrolled in a clinical psychology doctoral program.

Summarizing the findings from Survey 1 and Survey 2

The study lead (LRN) consulted with KM and JW, then compiled a summary of the results and qualitative feedback from Surveys 1 and 2. Based on findings from the prior two rounds, some recommendation results were collapsed across age groups rather than presented separately. The study lead distributed the results summary and instructed participants to focus their review on the quantitative results summaries from both surveys as well as overall qualitative feedback from Surveys 1 and 2. Participants were encouraged to read additional qualitative feedback on each target from Survey 1 as desired or needed.

Data collection

Survey 3 instructed participants to rate the FIRST principles and parenting practices in a similar manner to Survey 2. Based on the findings from Survey 1 that suggested recommendations did not differ significantly by age, participants provided recommendations

irrespective of child age. Based on qualitative data from Survey 2, participants were also instructed to provide their recommendations in three different ways: first, by using responses from the aggregated evidence and recommendations from the results of the two prior surveys; second, by also considering their own knowledge and experiences; and third, by ranking the skills they would prioritize including in the planned intervention. Participants were also invited to provide open-ended responses explaining any changes in their recommendations for this final survey.

Data analyses

We used a similar statistical approach for analyzing the data from all three surveys. To evaluate recommendations, rankings, and ratings across respondents, it is important to differentiate between stability, agreement, and consensus (Gisev, Bell, & Chen, 2013; Trevelyan & Robinson, 2015). For Delphi studies, the *stability* of responses can be assessed by measurements of inter-rater reliability, such as intra-class correlation coefficients (ICCs) (Trevelyan & Robinson, 2015), whereas *consensus* of responses can be assessed with measurements of variation, such as interquartile range (Trevelyan & Robinson, 2015), or measurements such as Kendall's coefficient of concordance (Gisev et al., 2013). As such, we first assessed the stability of responses by calculating Krippendorff's alphas (α) for nominal data and ICCs for ordinal data. We selected Krippendorff's α , as opposed to Fleiss' K , because they yield similar results but α is more robust to missing data (Zapf, Castell, Morawietz, & Karch, 2016). We selected ICC(3, k) to test two-way mixed effects of average measures for multiple raters (Koo & Li, 2016) for assessing consistency ($<.4 = \text{poor}$, $.4 = \text{fair}$, $.6 = \text{good}$, $.75 = \text{excellent}$; Cicchetti, 1994). Next, for assessing consensus of responses, we calculated variation ratios for nominal data and Kendall's coefficient of concordance for ordinal data. The variation ratio represents spread of responses ($VR = 1 - \text{proportion of cases in the mode}$; Weisberg, 1992), with scores from 0 to 1 indicating less to more dispersion. Kendall's coefficient of concordance (W , or W_t when corrected for ties) can assess the relative strength of agreement ($.3 = \text{low}$, $.5 = \text{fair}$, $.7 = \text{high}$, $.9 = \text{very high}$) by 3+ raters for rank order data (Schmidt, 1997). Given the number of participants and questions asked in Survey 1, multiple values of VR and α were calculated for assessing the recommendations overall, by subtype (e.g., primary vs. secondary), and by participant type (e.g., self-reported expertise). In addition, for Survey 1, we only ran chi-squared analyses to determine if the final presentation of recommendations should be collapsed or separated by age group. Finally, for all data types, we calculated frequencies to assess which responses were most commonly endorsed. All VR s were calculated in Microsoft Excel and all remaining analyses were conducted with RStudio version 3.6.0 (R Core Team, 2019) using the irr package (Gamer, Lemon, Fellows, & Singh, 2019).

Results

Survey 1

Frequency of FIRST skills recommendations

Examining recommendations for the eight targets individually, chi-squared analyses (or Fisher's exact tests in instances of low cell count) compared the frequency of skills recommendations for children (age 8–12 years) versus adolescents (age 13–15 years). These analyses showed that none of the recommendations significantly differed by age (p values from .28 to .99), so we collapsed the recommendations across both age groups. The

frequencies of skills recommendations by target are presented graphically in Figure 1.

Stability of FIRST skills recommendations

The overall Krippendorff's α for primary skills recommendations was .27, suggesting little stability across the intervention skills recommended across all developmental mechanisms and symptom types. There was a similar α value for recommendations on developmental mechanisms overall ($\alpha = .27$) and symptom types overall ($\alpha = .25$). Examining α values for skill recommendations for individual targets (e.g., heightened emotion reactivity, anxiety), no stability was observed regardless of collapsing or separating recommendation by age groups (α values from $-.03$ to $.00$). The α values varied slightly between participant groups, with the highest stability of responses by study investigators ($n = 4$, $\alpha = .36$), followed by external experts ($n = 25$, $\alpha = .26$), and then study consultants ($n = 3$, $\alpha = .23$). The α values also varied slightly by self-reported expertise, with the highest stability of responses by dissemination and implementation researchers ($n = 14$, $\alpha = .29$), followed by treatment researchers ($n = 17$, $\alpha = .28$), researchers focused on developmental mechanisms ($n = 18$, $\alpha = .22$), childhood adversity researchers ($n = 18$, $\alpha = .22$), and prevention researchers ($n = 14$, $\alpha = .18$).

Consensus of FIRST skills recommendations

The overall variation ratio (VR) from primary skills recommendations was .39, indicating that 39% of recommendations fell outside each respective modal recommendation. For reference, the "at chance" VR for items in this survey would be .80, with lower values reflecting higher consensus. There was slightly better consensus for primary skills recommendations for developmental mechanisms (overall $VR = .36$) than for symptom types (overall $VR = .42$).

There was quite a range in VR values when examining primary skill recommendations by individual targets, with the highest consensus for disruptive behavior (collapsed age $VR = .19$) and the lowest consensus for PTSD symptoms (collapsed age $VR = .58$) (see Figure 1 for recommendations for each FIRST skill). Compared with the primary skills recommendations, secondary skill recommendations had worse consensus overall ($VR = .72$) and for individual targets (VR range from .67 to .75). As such, we discontinued the evaluation of secondary skills recommendations. The VR values varied slightly between participant groups, with the highest consensus of responses by study investigators ($n = 4$, $VR = .27$), followed by study consultants ($n = 3$, $VR = .33$), and then external experts ($n = 25$, $VR = .26$). The VR values also varied slightly by self-reported expertise, with the highest consensus of responses by dissemination and implementation researchers ($n = 14$, $VR = .36$), followed by treatment researchers ($n = 17$, $VR = .39$), childhood adversity researchers ($n = 18$, $VR = .40$), developmental mechanisms-focused researchers ($n = 18$, $VR = .41$), and prevention researchers ($n = 14$, $VR = .43$).

Ranking of parenting practices

Examining the ranking of recommended parenting practices, stability was excellent both for children ($ICC(C,29) = .83$) and adolescents ($ICC(C,29) = .90$). However, consensus was very low both for children ($W_t = .17$) and adolescents ($W_t = .27$). A chi-squared analysis revealed that the first ranked parenting practice did not differ by age group ($p = .4$); the rankings of parenting practices for each age group are presented in Figure 2.

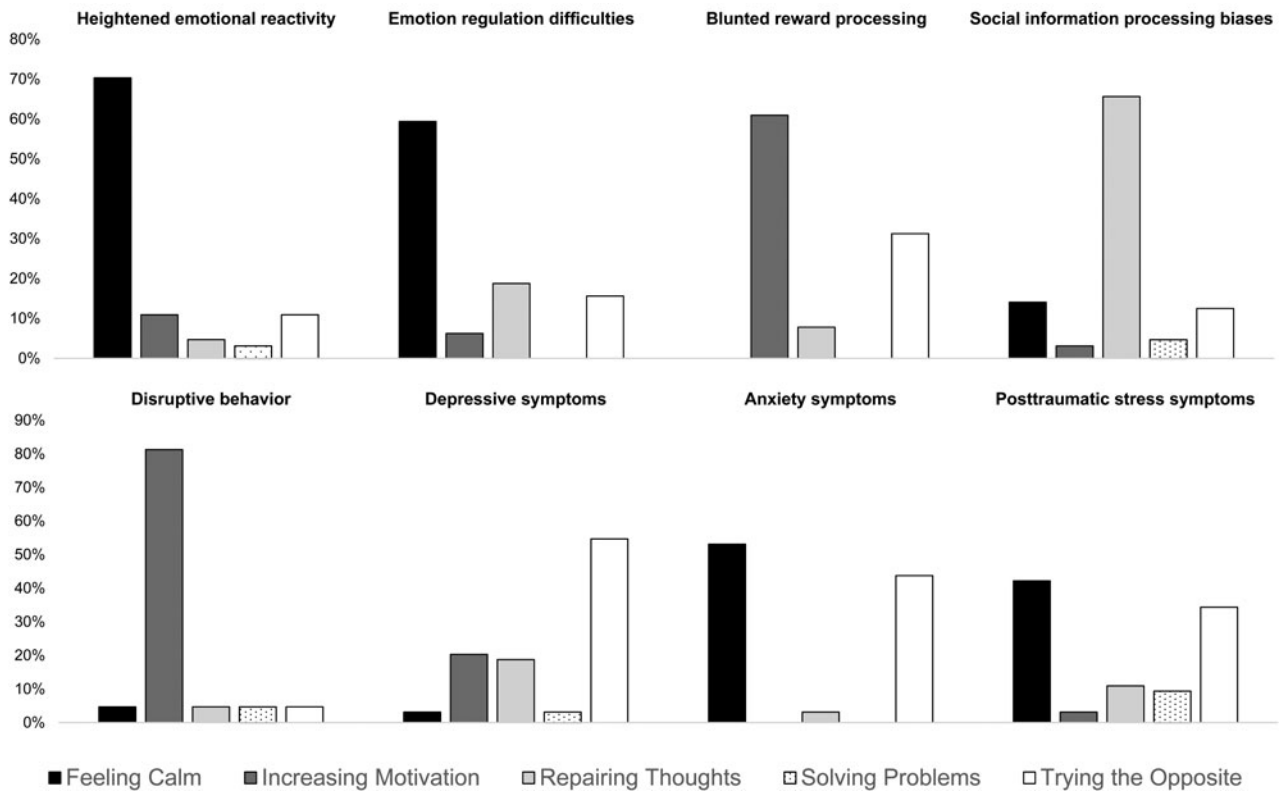


Figure 1. Frequencies of recommended FIRST principles by targets (Survey 1).

Survey 2

Intervention skills broadly

From a compiled list of 33 intervention components/practice elements that were reviewed in the seven screened meta-analyses, respondents were asked which they would select as the top five with “strongest evidence” for different purposes. Stability across respondents ($n = 6$) was low for top five skills selected overall (α values from .29 to .37). However, consensus was very good for the selection of the top five skills overall ($VR = .11$), considering prevention programs only ($VR = .11$), and considering programs that were most similar to our planned early adversity intervention ($VR = .09$). For reference, the “at chance” VR for these cases would be .83. Moreover, about 20 options were never ranked by any respondent, showing some collective discrimination of elements that were clearly not in the top five. Figure 3 shows those skills ranked as the top five by respondents for our prevention program purposes.

FIRST skills and parenting practices

Examining respondents’ strength of evidence ratings for the five FIRST skills, stability was excellent ($ICC(C,6) = .88$) and consensus was high ($Wt = .85$). The five intervention strategies ranked as having the strongest evidence were discipline/behavior management, parenting skills for positive reinforcement, improving parent–child relationship/interactions, problem-solving for parents, and relaxation skills (see Figure 4). Examining respondents’ strength of evidence ratings for the three parenting practices of interest, stability was excellent ($ICC(C,6) = .96$) and consensus was high ($Wt = .81$).

Survey 3

FIRST skills ratings and rankings

When respondents ($n = 8$) were instructed to rate each FIRST skill according to the evidence they reviewed from the results of Survey 1 (expert recommendations) and Survey 2 (meta-analysis review), stability was excellent ($ICC(C,8) = .94$) and consensus was high ($Wt = .75$) (see Figure 5 for average ratings). When respondents were instructed that they could also consider knowledge from their own experiences, stability was still excellent ($ICC(C,8) = .85$) but consensus was moderate ($Wt = .60$). Relative to the ratings presented in Figure 5, these second ratings based also on respondents’ individual knowledge produced the following changes (as measure by standardized mean difference; SMD): Feeling calm had a small decrease ($SMD = -0.43$); Increasing motivation had a small increase ($SMD = 0.23$); Repairing thoughts, Solving problems, and Trying the opposite had a moderate increase ($SMD = 0.57$ to 0.66). Finally, when respondents rank-ordered the skills to prioritize for our planned intervention, stability was still excellent ($ICC(C,8) = .97$) and consensus was high ($Wt = .82$), with the rank order of FIRST skills exactly following the strength of evidence ratings order as per Figure 5.

Parenting practice ratings and rankings

When respondents were instructed to rate each parenting practice based on evidence from Surveys 1 and 2, stability was excellent ($ICC(C,8) = .98$) and consensus was very high ($Wt = .97$) (see Figure 5 for average ratings). When respondents were instructed to also consider their own past knowledge, stability was again excellent ($ICC(C,8) = .97$), while consensus decreased but was still high ($Wt = .88$), with the following changes: Increasing

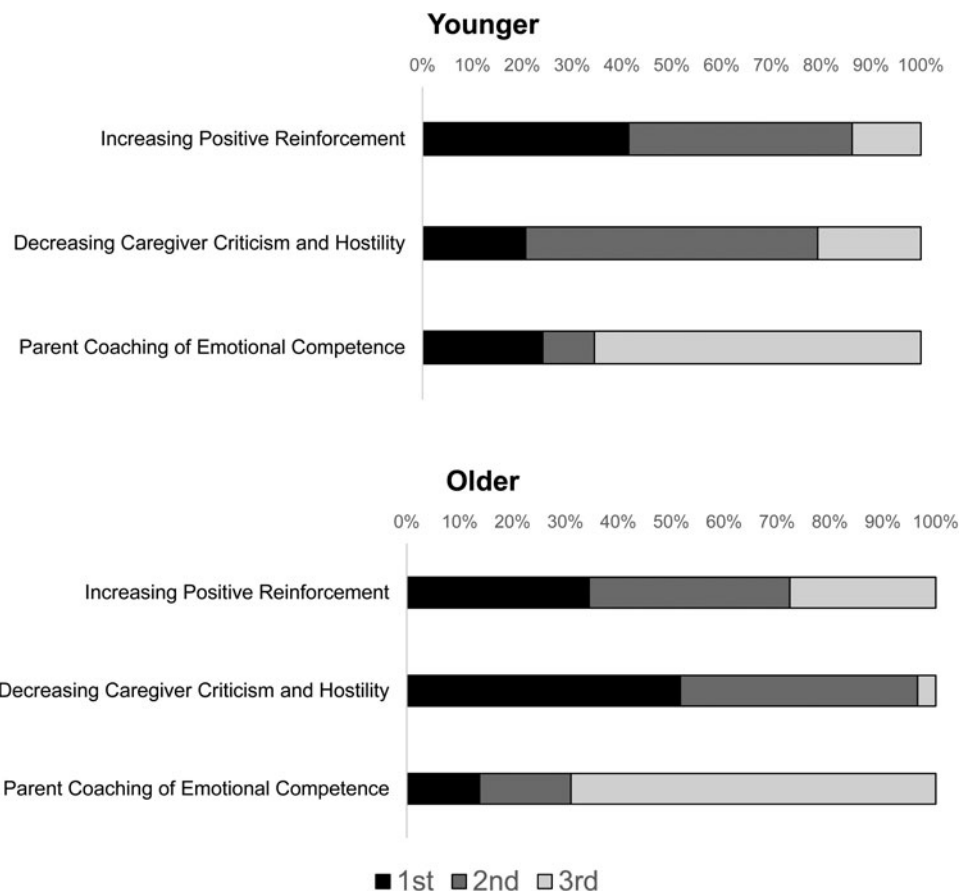


Figure 2. Ranking of parenting practices for younger (ages 8–12) and older (ages 13–15) presentations (Survey 1).

positive reinforcement stayed at the maximum possible rating, Decreasing caregiver criticism and hostility had a small increase ($SMD = 0.40$), and Coaching of child emotional competence had a large increase ($SMD = 0.86$). Finally, when respondents rank-ordered the parenting practices, stability was still excellent ($ICC(C,8) = .98$) and consensus was still high ($Wt = .89$), with the rank order again exactly following the strength of evidence ratings order as per Figure 5.

Final stability and consensus of FIRST skills recommendations

The overall α value for skills recommendations was .60, suggesting moderate stability. There was also moderate stability for recommendations by developmental mechanisms overall ($\alpha = .58$) and symptom types overall ($\alpha = .52$). As for consensus, the overall VR was .22, an improvement over the consensus of Survey 1. The “at chance” VR for these items is .80. The consensus was identical for skills recommendations for developmental mechanisms (overall $VR = .22$) and symptom types (overall $VR = .22$).

Regarding skill recommendations for individual targets, we reached absolute stability ($ICC(C,8) = 1.00$) and consensus ($VR = .00$) for four targets – Feeling calm was recommended by all participants for targeting heightened emotional reactivity; Increasing motivation for targeting disruptive behavior; Repairing thoughts for targeting social information processing biases; and Trying the opposite for targeting depressive symptoms. For blunted reward processing, consensus was split ($VR = .5$) and thus stability was very low ($\alpha = -.12$), with 50% ($n = 4$) of respondents recommending Increasing motivation and 50%

recommending Trying the opposite. The α value was $-.12$ and the VR was .38 for the remaining three targets of emotion regulation difficulties, anxiety symptoms, and posttraumatic stress symptoms (see Figure 6 for details).

Qualitative feedback across surveys

Although a comprehensive qualitative analysis is beyond the purview of this paper, selected responses may be helpful to contextualize quantitative findings. We have provided some exemplar quotes from each survey in Table 3 to demonstrate how respondents explained their decision making process or elaborated on rationales, qualifiers, or additions to their recommendations.

Discussion

In the service of developing an intervention to prevent the mental health consequences of childhood adversity, we conducted a modified Delphi study to assess intervention strategy recommendations and ratings based on expert feedback and review of the existing intervention evidence base. Over three survey rounds, we found increasing consensus on intervention strategies regarding recommendations for individual targets, strength of evidence ratings, and ranking order. Certain intervention strategies for youth skills and parenting practices emerged as stronger candidates, some with consensus for targeting specific developmental mechanisms and symptoms. In addition, our findings produced

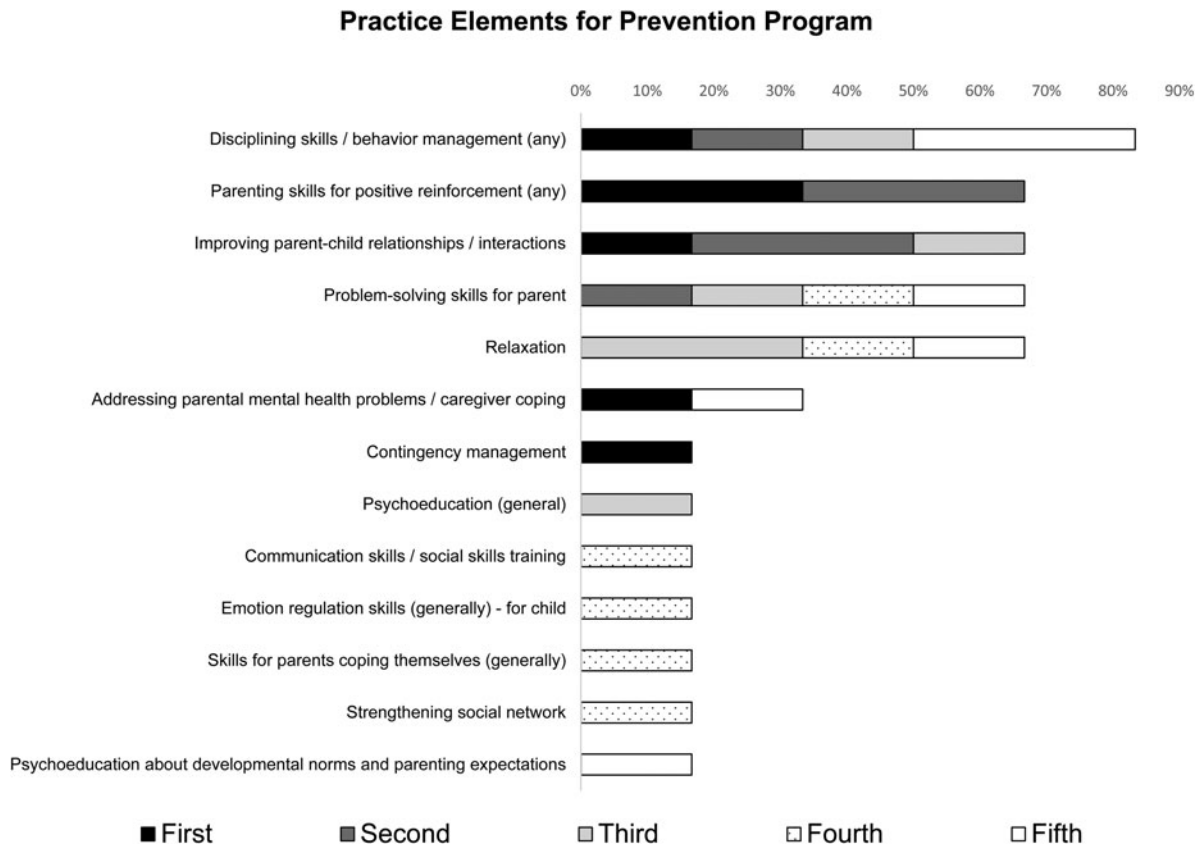


Figure 3. Skills ranked as top five based on review of meta-analyses (Survey 2).

useful information about the process of evidence-informed intervention decision making.

First, we were interested in investigating which candidate intervention strategies – specifically, youth skills drawn from FIRST (Weisz & Bearman, 2020) and parenting practices drawn from PCIT-ED (Luby et al., 2018) – were recommended for the developmental mechanisms and symptoms we intend to target. To do so, we first asked a group of experts (Survey 1). Consensus in these recommendations can be interpreted from the lower *VR* values and higher *W* values within the survey rounds, and by relative decreases or increases between rounds (Gisev et al., 2013; Holey, Feeley, Dixon, & Whittaker, 2007). For the youth skills recommendations in Survey 1, we found very low stability and low consensus overall, though one or two modal recommendations emerged for all eight targets. Overall, the most recommended youth skills in Survey 1 were Feeling calm, Increasing motivation, and Trying the opposite; Solving problems was rarely recommended. Regarding parenting practices recommended for our intervention overall, the experts in Survey 1 had excellent stability of rankings, but very low consensus. Low consensus reflects that respondents are divergent from each other, whereas high stability reflects that they are diverging in a consistent way. The parenting practice Coaching of child emotional competence was clearly ranked lowest of the three. Of note, consensus of the recommendations in Survey 1 did not appear to depend on youth age, respondent's affiliation with our study team, or self-reported expertise type. In Survey 3, we invited only study team members to complete these recommendations a second time after they had reviewed summaries of the recommendations from the broader group of experts. This time, stability and consensus substantially

improved for recommendations of both youth skills and parenting practices. The final most frequently recommended skills were again Feeling calm, Increasing motivation, and Trying the opposite; Solving problems was never recommended. Youth skills recommendations reached 100% consensus in Survey 3 for four of the eight targets – Feeling calm for heightened emotional reactivity, Increasing motivation for disruptive behavior symptoms, Repairing thoughts for social information processing biases, and Trying the opposite for depressive symptoms – and was split 50/50 between Increasing motivation and Trying the opposite for blunted reward processing. For the final parenting practice recommendations in Survey 3, Increasing positive reinforcement emerged as the unanimous first parenting practice to prioritize and Coaching of child emotional competence remained the last.

We were also interested in investigating which intervention strategies were most strongly supported by existing empirical evidence. To do so, we identified and reviewed relevant meta-analyses and rated which intervention elements had the strongest evidence related to our intended prevention program. In Survey 2, although there were more than 30 practice elements and intervention strategies presented across the reviewed meta-analyses, the respondents showed good consensus for identifying the intervention elements that had the strongest effects. When asked which intervention strategies had the most evidence for our prevention program, at least two thirds of our study team selected the following within their top five: discipline skills/behavior management; parenting skills for positive reinforcement; improving parent-child relationships/interactions; problem-solving skills for parents; and relaxation strategies. Respondents were then asked to translate their

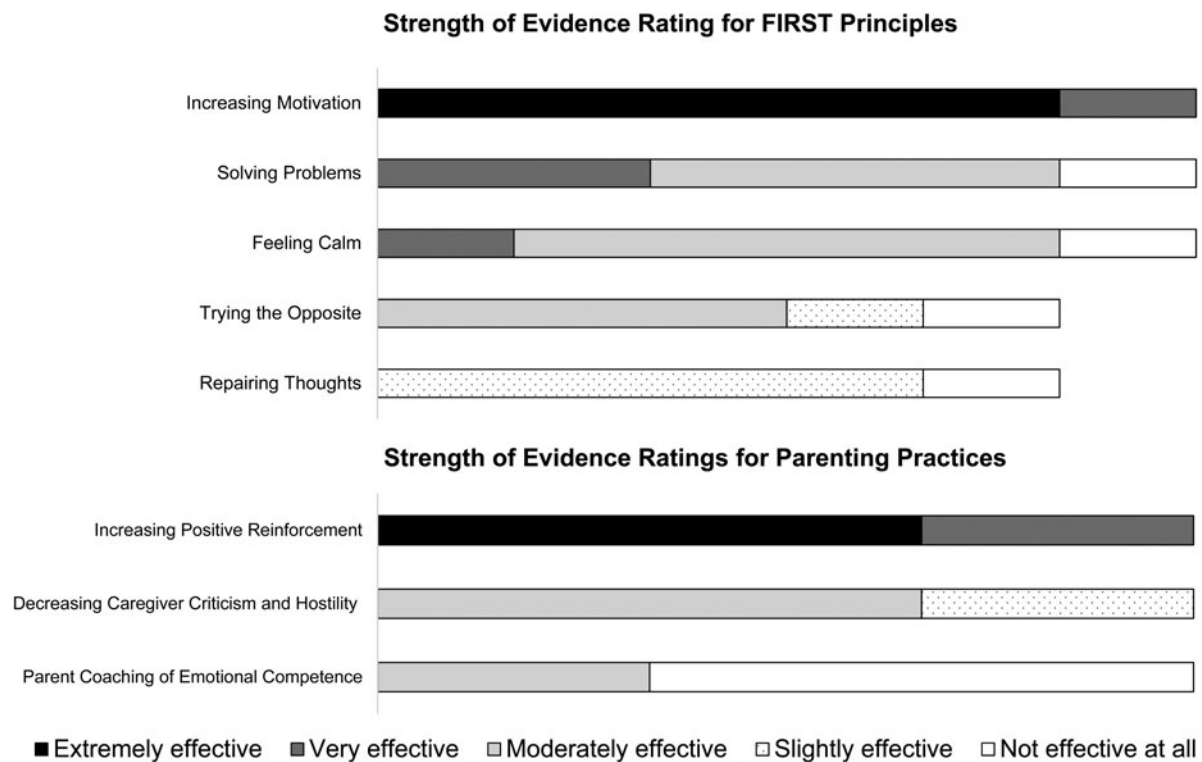


Figure 4. Strength of evidence ratings based on review of meta-analyses (Survey 2).

selections into strength of evidence for our candidate FIRST youth skills and parenting practices, and the resultant ratings had excellent stability and high consensus. According to Survey 2, strength of evidence ratings were “extremely effective” for Increasing motivation (a skill that targets both youth and parent behaviors; see Table 1) and Increasing positive reinforcement (parenting practice), “moderately effective” for Feeling calm (youth skill), Solving problems (youth skill), and Decreasing caregiver criticism and hostility (parenting practice), and “slightly effective” for Trying the opposite (youth skill), Repairing thoughts (youth skill), and Coaching of child emotional competence (parenting practice). Strength of evidence ratings in Survey 3 again had excellent stability, while consensus for youth skills was still high but somewhat decreased, and consensus for parenting practices was very high, a relative increase. According to Survey 3, the final strength of evidence ratings stayed the same for Increasing motivation, Increasing positive reinforcement, Repairing thoughts, and Coaching of child emotional competence. The changes in strength of evidence ratings from Survey 2 to Survey 3 were an increase to “very effective” for Feeling calm and Decreasing caregiver criticism and hostility, an increase to “moderately effective” for Trying the opposite, and a decrease to “slightly effective” for Solving problems.

Implications for preventing mental health consequences of childhood adversity

Our primary aim in investigating these research questions was to evaluate intervention strategy consensus so that findings could inform the development of our prevention program for youth with varying presentations. As stability and consensus improved across surveys, there was some convergence with Feeling calm (from FIRST), Increasing motivation (from FIRST), Trying the

opposite (from FIRST), and Increasing positive reinforcement (from PCIT) receiving the most frequent recommendations, higher rankings, and stronger strength of evidence ratings. The fact that these same intervention strategies were the most highly recommended/rated across all three surveys suggests that consensus is generalizable beyond our study team. These intervention strategies are also conceptually similar to the practice elements with strong effects in relevant meta-analyses. As examples, relaxation (which was included in two of the meta-analyses) is the central intervention strategy of the Feeling calm principle in FIRST, and positive attending, improving parent–child relationships, positive reinforcement strategies, and other behavior management skills (at least one of which was included in all seven meta-analyses) are core skills in the Increasing motivation principle of FIRST and the Increasing positive reinforcement element of PCIT. Although it is important to highlight that intervention strategies with high consensus do not guarantee that those strategies will necessarily be most effective (Trevelyan & Robinson, 2015), the current findings certainly can help us prioritize our best candidates, given the currently existing evidence. In particular, given that our planned prevention program will be brief (up to six sessions), it may be useful to focus on Feeling calm, Increasing motivation, and Trying the opposite as the most relevant intervention elements from FIRST, while deprioritizing Repairing thoughts and Solving problems. Similarly, Increasing positive reinforcement is the parenting practice that could be prioritized above both Decreasing caregiver criticism and hostility and Coaching of child emotional competence.

One particularly encouraging aspect of our findings was reaching consensus on intervention strategies for some of our target developmental mechanisms. Many psychosocial interventions are designed for targeting symptoms, which necessarily involves more remedial intervention strategies rather than preventive

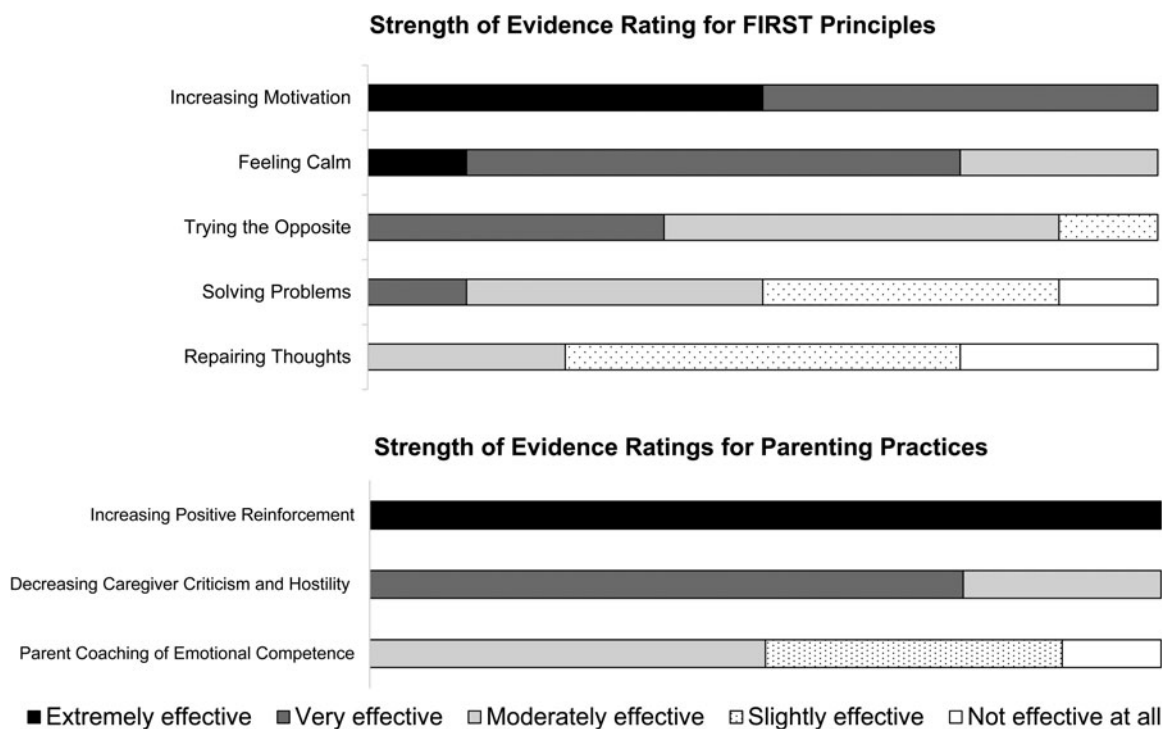


Figure 5. Strength of evidence ratings based on respondents' review of prior rounds (Survey 3).

strategies that reduce risks and promote strengths (American Psychological Association, 2014). Moreover, many clinical psychologists engage in intervention decision making in a symptom-driven way (Bakker, 2019) – part of the motivation for developing the APA prevention guidelines was that, despite the increased focus on prevention within clinical psychology, there is little understanding of the best preventive practices (American Psychological Association, 2014). Considering how little guidance and experience clinical psychologists have regarding prevention strategies, it was notable that we achieved 100% or 50/50 consensus for skill recommendations for three of our four developmental mechanisms (i.e., heightened emotional reactivity, social information processing biases, and blunted reward processing). Moreover, even before we reached consensus, these mechanisms already had modal skill recommendations in Survey 1, which indicates that the experts were not making recommendations randomly. It is not the case that each mechanism was simply treated as a proxy for some perceived parallel symptom (e.g., reward processing vs. depression); otherwise, the same four recommendations for symptom targets would have emerged for developmental mechanism targets. Qualitative data also supported this idea, as exemplified by one respondent describing how their intervention recommendation for depressive symptoms was informed by both the meta-analysis review (which covered depression targets) and recommendations “for blunted reward processing by experts – a mechanism closely related to depression.” Conceptually, the skills with strongest strength of evidence ratings largely involved risk reduction strategies (e.g., addressing positive parent-child relationship through Increasing positive reinforcement and Increasing motivation) and strengths promotion (e.g., Feeling calm as a coping strategy), as opposed to remedial skills that require some symptoms to be already present (e.g., exposure, cognitive restructuring). Thus, it appears

possible for mental health professionals to apply their intervention decision making to planning a mechanisms-based preventive intervention.

Insights into evidence-informed clinical decision making

We had an additional exploratory aim, which was to gain insight into how evidence-informed decision making in intervention development unfolds. Our findings illuminate understanding of intervention decision making processes in four main ways.

First, the use of meta-analysis to inform intervention decisions is frequently recommended (Powell et al., 2015; Weisz et al., 2005) yet rarely studied as a practice. Although meta-analyses and systematic reviews can, in theory, provide meaningful summaries of intervention evidence, there is little guidance on how to actually use them in practice when designing interventions. Our practice element ranking findings from Survey 2 demonstrate that even experts reviewing the same meta-analyses come to different conclusions about which intervention strategies emerged with the strongest evidence. Relatedly, respondents noted the difficulty of cognitively condensing and interpreting results across meta-analyses, especially for those with less familiarity with meta-analysis statistical methods. The difficulty in translating meta-analysis findings into intervention selections might explain some variation in the recommendations. For example, one respondent rated Solving problems as “moderately effective” based on some meta-analytic support of *problem solving for caregivers*, while another respondent acknowledged that same evidence but concluded low support for Solving problems because of other meta-analytic null results of *problem solving for youth*. Thus, although it is common practice to point to meta-analyses to decide best clinical practices, interpretation of meta-analytic findings is not actually a clear-cut task, even for a study team working closely together.

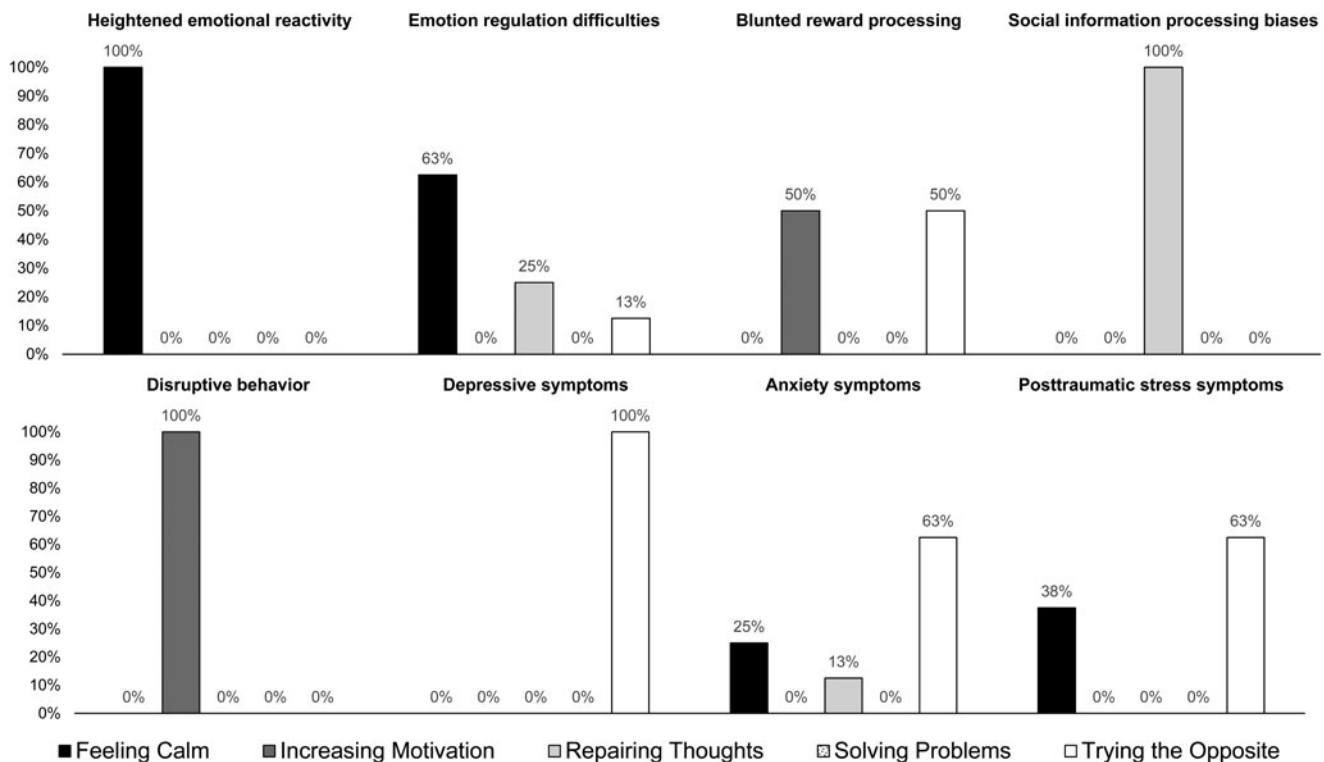


Figure 6. Survey 3: Frequencies of recommended FIRST principles by target.

Second, our findings provide evidence that a Delphi study can help compensate for biases in decisions made by individuals alone or by groups without structured consensus building. The use of a Delphi study is intended to enhance group decision making by mitigating issues of bias and group dynamics (Hsu & Sandford, 2007; Surowiecki, 2004). Moreover, much research has found that clinical decisions by individuals are affected by cognitive errors such as the availability heuristic, overconfidence, and hindsight bias (Lilienfeld & Lynn, 2014; Witteman, 2020). To overcome these limitations, it has been recommended that we develop debiasing methods, or at least ways to systematically compensate for them (Lilienfeld, 2019; Lilienfeld & Lynn, 2014), which we believe could include the Delphi approach. The fact that the study investigators already had relatively high consensus on specific intervention recommendations in Survey 1, even before reviewing the meta-analyses in Survey 2, might indicate that some groupthink had already been occurring. The improved consensus reached in Survey 3 demonstrates how reviewing objective summaries of evidence and others' ideas from prior survey rounds ultimately influenced the final decisions. One respondent confirmed this experience, saying "the aggregated results very much informed my thinking," even allowing their review of prior results to override personal opinions when caught between two recommendations. Finally, when respondents provided their strength of evidence ratings in Survey 3, again based also on their own knowledge and experiences (i.e., beyond the information aggregated from experts and the meta-analyses), there was a decrease in consensus for both youth skills and parenting practices. This decrease was accompanied by a trend of respondents giving higher evidence ratings across intervention strategies, which makes sense given that our team selected these candidate skills in the first place. Together, our findings increase our confidence that a Delphi study can increase consensus in decisions

about intervention selection while decreasing bias in group decisions.

Third, our findings are a reminder that *aiming* for consensus does not necessitate *achieving* consensus. Any Delphi study can be continuously iterative, so there must be a determination by researchers about when more rounds would no longer be useful (Hsu & Sandford, 2007). Compared with more traditional Delphi methods, our modified Delphi study used a larger initial sample and then had only our smaller study team continue. Through this approach, our Survey 3 had the highest consensus levels and we reached maximum consensus for recommendations on half of our targets. However, our consensus on strength of evidence ratings slightly decreased from Survey 2 to Survey 3, along with the growing amount of information for respondents to review and consider. Given that we still achieved high consensus in Survey 3, it seemed that we may have reached a ceiling. This finding highlights how a consensus process may reach saturation, at which point perpetually synthesizing group feedback may provide diminishing clarity. Finally, consensus is not always the goal when there may not be a single answer (Hall, 2009), with a good example being our 50/50 split on intervention recommendations for blunted reward processing.

Fourth, despite attempts at debiasing our clinical decisions, objective decision making is quite challenging. Qualitative feedback across surveys emphasizes this fact. Even though respondents were capable of providing intervention recommendations for the prevention mechanism targets, they still had some difficulty avoiding symptom-driven thinking. Some comments revealed that respondents were still considering clinical experiences or citing evidence related to treatment programs. For example, one respondent in Survey 2 admitted "biased interpretation from my own clinical understanding of working with [children]" after reviewing meta-analyses on *prevention* programs. Even by

Table 3. Selected qualitative data across surveys

<i>Survey 1</i>
After recommending Trying the opposite for PTSD symptoms: “If there were not enough time, I would prioritize feeling calm interventions. Also, if we are talking about nonclinical symptoms of PTSD then I would probably focus on feeling calm strategies and cognitive strategies to think about the traumatic event in a way that is more helpful and less distressing.”
After recommending Trying the opposite for anxiety symptoms: “[I] want to clarify that I see relaxation skills and exposure being part of the same set here (used jointly).”
“My judgment of how to rank order the parenting skills would actually depend on the kind of behavior the child was showing, not on what adversity the child had experienced, because similar patterns of diversity can be followed by very diverse patterns of child behavior – each requiring an appropriate intervention.”
“I think there are important considerations about caregiver preparation to enact these skills and dosing considerations are important given different caregiver baseline characteristics which may covary with nature of family adversity as well as cultural and educational background.”
<i>Survey 2</i>
“This was a lot of information to absorb, and my stats background is not that strong. I’m also probably biased from my own clinical understanding of working with early/middle childhood... I’m not very confident in my responses.”
“Problem-solving skills for parents was a second rated skill, and I rated Solving problems as moderately effective.”
“For ‘Solving problems,’ child problem solving was tested in 3 of 7 meta-analyses, and none showed it being significant. The only time I saw problem solving with support it was as a caregiver skill, so I’m not counting that. Moreover, caregiver problem solving was also only significant in 1 of 2 meta-analyses that tested it.”
“It’s a bit difficult to do this in a broad, generic way, since the skills most needed and likely to be helpful may differ widely depending on what concerns and problems the parents want to have addressed – e.g., very different skills if dealing with child conduct problems than if dealing with child anxiety or depression.”
<i>Survey 3</i>
“The aggregated results very much informed my thinking. I had my own opinions going in which informed my recommendations somewhat, but I probably weighed the aggregated results more highly than my independent opinions... in a few places where there were two strategies that were close in my mind, there were situations where reviewing the results of the two surveys pushed my top choice from one strategy to another.”
“The aggregated results did lead me to lower ratings on problem solving than my experience in our intervention trials would suggest, so I did raise that particular rating when invited to use my own knowledge and experience.”
After recommending Feeling calm for emotion regulation difficulties: “[I would say] repairing thoughts for older youth maybe, but my main hesitation after reading the first-round feedback is feasibility of teaching higher level skills without all the time necessary for psychoeducation, rapport building, etc.”
After recommending Trying the opposite for depressive symptoms: “Increasing motivation would be my second choice, particularly given the weight of evidence for this strategy in the meta-analysis review, and that it was most highly rated for blunted reward processing by experts – a mechanism closely related to depression.”
“I rated Trying the opposite as the 3rd most effective skill for a range of reasons. First, in the [meta-analysis of brief interventions], which is similar to our approach, behavioral interventions targeting the child were most effective. Second, Trying the opposite was [recommended first or second] for depression, anxiety, and PTSD by experts. I am also including my personal experiences here of the remarkably large (and fast) effects that behavioral activation and exposure can have for children with mood and anxiety problems.”
When asked about PTSD symptoms: “Ideally, we would want to provide something like TF-CBT, to help kids really address their PTS and their avoidance. However, in six sessions – and for many families, fewer – and with clinicians who don’t have TF-CBT training, we may be safer focusing on calming strategies.”

Survey 3, respondents were still referring to treatment recommendations (e.g., TF-CBT) and effective interventions for symptoms (e.g., “behavioral activation and exposure [for] mood and anxiety problems”) without making the distinction between treatment and prevention programs. Finally, the most notable example comes from some conflicting quantitative findings in Survey 3. On the one hand, we had 100% consensus for the recommendation of Repairing thoughts for the social information processing biases, despite also having high consensus in rating this skill with the lowest strength of evidence overall. On the other hand, none of the respondents recommended Solving problems for a single target, despite having rated Solving problems as “moderately effective” on average. In other words, within the same survey, we sometimes recommended intervention strategies for our prevention targets that contradicted our own rating of their strength of evidence for our overall prevention program. Though such recommendations may still hold some merit, these findings highlight inconsistency in how we apply clinical knowledge. Other research has found that that clinician intervention recommendations can be more reflective of personally held theories rather than current scientific evidence (de

Kwaadsteniet & Hagmayer, 2018). In summary, consensus does not solve all decision making problems and updating new knowledge into our pre-existing – and sometimes entrenched – frameworks remains a challenge for evidence-informed intervention decision making. These findings are somewhat reminiscent of sentiments from Zigler that it is difficult “to bring together politics, knowledge [...], and common sense” when developing standards for prevention content, and even “the best standards in the world are meaningless if they are not enforced” (Zigler, 1996, p. 38; Zigler & Styfco, 2001, p. 10)

Limitations

Our study was subject to some limitations. First, interpretation of the findings was constrained by which intervention strategies we considered. As an example, we did not query respondents about computerized attention bias modification strategies, which are less commonly used in clinical settings, yet theoretically might be the most effective for threat-related social information processing biases (e.g., attention bias to threat; Bar-Haim, 2010). The

selected meta-analyses added to this constraint, as we likely missed some relevant reviews through our informal search procedures; furthermore, various intervention strategies are not equally represented in the published literature.

Second, the five FIRST principles, which served as our candidate intervention strategies for youth, vary in the breadth of skills contained within each module. Thus, while a respondent's recommendation of Feeling calm clearly involves relaxation strategies, a recommendation of Increasing motivation could be focused on one, some, or all of the encapsulated parent management skills (e.g., reinforcement strategies to promote positive behaviors, specific discipline strategies).

The third limitation was that the specific respondents varied across the three samples. However, we decided on this methodological approach to balance (a) representation – compared with most Delphi studies, we invited a larger group with greater diversity of research expertise and target populations (Cairns, Yap, Reavley, & Jorm, 2015), (b) feasibility – it seemed unrealistic to expect a group of experts to commit time to multiple surveys (Cairns et al., 2015; Hsu & Sandford, 2007, and (c) quality of responses – our study team members had intrinsic motivation to thoughtfully review materials and could more readily provide recommendations with our specific intervention purposes in mind.

Fourth, not only does *consensus* on intervention recommendations from experts differ from *effectiveness* (as previously discussed), it also does not ensure *acceptability*. Client preferences predict their satisfaction with, completion of, and clinical response to intervention (Lindhiem, Bennett, Trentacosta, & McLearn, 2014). Accordingly, it ultimately does not matter how evidence-based a preventive intervention is if primary care clinics and families who have experienced adversity do not want to participate. We are thus currently collecting stakeholder feedback from pediatric clinic providers and staff as well as families who have encountered adversity. Relatedly, to increase multicultural appropriateness as per APA prevention guidelines (American Psychological Association, 2014), we will (a) more thoroughly review expert qualitative feedback regarding multicultural recommendations and (b) specifically ask stakeholder families in our feasibility trial about cultural acceptability of materials and language adaptations.

Conclusion

Although many interventions have been developed to ameliorate the mental health consequences of childhood adversity, evidence of effectiveness for preventive interventions is scarce – particularly those that are brief and could be feasibly integrated into pediatric primary care. As such, we have yet to discover which prevention strategies are most effective for families with adversity-exposed youth. Simply borrowing and repackaging the most common treatment intervention strategies seems unlikely to advance effectiveness. Instead, our study team systematically considered evidence from experts and meta-analyses to determine which intervention strategies were the most relevant and most potent for our prevention purposes. Moreover, we invited recommendations on targeting four mechanisms related to the neurodevelopmental adaptations of childhood adversity exposure that may place youth at risk of later psychopathology. Given that not all adversity-exposed youth will be symptomatic at the time of screening, it is crucial that our prevention approach includes skills that are not dependent on symptom presentations. For example,

behavioral exposures have shown stronger effectiveness than relaxation skills in the *treatment* of youth anxiety (Whiteside et al., 2020). However, exposures are essentially meaningless in the absence of impairing behavioral avoidance, whereas relaxation strategies, in contrast, can be used by anyone. Our findings thus give credence to that idea that strong intervention candidates for prevention programs will likely differ from those of treatment programs, in particular by including strategies that are conceptually more focused on risk reduction and strength promotion. Finally, our findings provided initial support for some candidate intervention strategies that could hypothetically target these developmental mechanisms. By developing our preventive program accordingly, we will now be able to empirically test the actual effects of the intervention strategies recommended through the current study. In the words of Zigler, once we “know what quality components are necessary to build successful interventions” we must then “put this wisdom to use” (Zigler, 1996, p. 42).

Supplementary Material. The supplementary material for this article can be found at <https://doi.org/10.1017/S0954579420002059>.

Acknowledgments. We would like to thank our study team members – Shannon Dorsey, Arianna Riccio, Olivia Fitzpatrick, and Katherine Venturo-Conerly – for their contributions in the preparation of this study. We also thank Joan Luby for her ongoing consultation for this intervention development. Finally, we thank Beth Erickson and Deneen Vojta for their support of the ongoing implementation studies associated with this larger intervention research project.

Financial Statement. The presented research was supported by the Harvard University Dean's Competitive Fund for Promising Scholarship (KM). The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the funding sponsor.

Conflicts of Interest. John Weisz is a coauthor of a treatment manual cited in this paper that will be used in the research described herein (*Principle-guided psychotherapy for children and adolescents: The FIRST treatment program for behavioral and emotional problems*, published in 2020 by Guilford Press (Weisz & Bearman, 2020)). He receives royalties for its sales.

References

- Abela, J. R. Z., Vanderbilt, E., & Rochon, A. (2004). A test of the integration of the response styles and social support theories of depression in third and seventh grade children. *Journal of Social and Clinical Psychology, 23*, 653–674. doi:10.1521/jscp.23.5.653.50752
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.
- American Psychological Association. (2014). Guidelines for prevention in psychology. *American Psychologist, 69*, 285–296. doi:10.1037/a0034569
- Asarnow, J. R., Rozenman, M., Wiblin, J., & Zeltzer, L. (2015). Integrated medical-behavioral care compared with usual primary care for child and adolescent behavioral health: A metaanalysis. *JAMA Pediatrics, 169*, 929–937. doi:10.1001/jamapediatrics.2015.1141
- Aupont, O., Doerfler, L., Connor, D. F., Stille, C., Tisminetzky, M., & McLaughlin, T. J. (2013). A collaborative care model to improve access to pediatric mental health services. *Administration and Policy in Mental Health and Mental Health Services Research, 40*, 264–273. doi:10.1007/s10488-012-0413-0
- Bakker, G. M. (2019). A new conception and subsequent taxonomy of clinical psychological problems. *BMC Psychology, 31*, 653–673. doi:10.1186/s40359-019-0318-8
- Bar-Haim, Y. (2010). Research review: Attention bias modification (ABM): A novel treatment for anxiety disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 51*, 859–870. doi:10.1111/j.1469-7610.2010.02251.x
- Barch, D. M., Whalen, D., Gilbert, K., Kelly, D., Kappenman, E. S., Hajcak, G., ... Luby, J. L. (2019). Neural indicators of anhedonia: Predictors and mechanisms of treatment change in a randomized clinical trial in early

- childhood depression. *Biological Psychiatry*, 85, 863–871. doi:10.1016/j.biopsych.2018.11.021
- Bellis, M. A., Hughes, K., Ford, K., Ramos Rodriguez, G., Sethi, D., & Passmore, J. (2019). Life course health consequences and associated annual costs of adverse childhood experiences across Europe and North America: a systematic review and meta-analysis. *The Lancet Public Health*, 4, E517–E528. doi:10.1016/S2468-2667(19)30145-8
- Benjet, C., Borges, G., Medina-Mora, M. E., Zambrano, J., Cruz, C., & Méndez, E. (2009). Descriptive epidemiology of chronic childhood adversity in Mexican adolescents. *Journal of Adolescent Health*, 45, 483–489. doi:10.1016/j.jadohealth.2009.03.002
- Bentovim, A., Chorpita, B. F., Daleiden, E. L., Gray, J., Pizzey, S., & Vizard, E. (2020). The value of a modular, multi-focal, therapeutic approach to addressing child maltreatment: Hope for children and families intervention resources – a discussion article. *Child Abuse & Neglect*, 1–13. Advance online publication. doi:10.1016/j.chiabu.2020.104703
- Berkowitz, S. J., Stover, C. S., & Marans, S. R. (2011). The child and family traumatic stress intervention: Secondary prevention for youth at risk of developing PTSD. *Journal of Child Psychology and Psychiatry*, 52(6), 676–685. doi:10.1111/j.1469-7610.2010.02321.x
- Brestan, E. V., & Eyberg, S. M. (1998). Effective psychosocial treatments of conduct-disordered children and adolescents: 29 years, 82 studies, and 5,272 kids. *Journal of Clinical Child Psychology*, 27, 180–189. doi:10.1207/s15374424jccp2702_5
- Brown, D. W., Anda, R. F., Tiemeier, H., Felitti, V. J., Edwards, V. J., Croft, J. B., ... Giles, W. H. (2009). Adverse childhood experiences and the risk of premature mortality. *American Journal of Preventive Medicine*, 37, 389–396. doi:10.1016/j.amepre.2009.06.021
- Cairns, K. E., Yap, M. B. H., Reavley, N. J., & Jorm, A. F. (2015). Identifying prevention strategies for adolescents to reduce their risk of depression: A Delphi consensus study. *Journal of Affective Disorders*, 183, 229–238. doi:10.1016/j.jad.2015.05.019
- Caldwell, D. M., Davies, S. R., Hetrick, S. E., Palmer, J. C., Caro, P., López-López, J. A., ... Welton, N. J. (2019). School-based interventions to prevent anxiety and depression in children and young people: a systematic review and network meta-analysis. *The Lancet Psychiatry*, 6, 1011–1020. doi:10.1016/S2215-0366(19)30403-1
- Campbell, T. L. (2020). Screening for adverse childhood experiences (ACEs) in primary care: A cautionary note. *JAMA*, 323, 2379–2380.
- Campbell, S. M., Roland, M. O., & Buetow, S. A. (2000). Defining quality of care. *Social Science and Medicine*, 51, 1611–1625. doi:10.1016/S0277-9536(00)00057-5
- Carpenter, A. L., Puliafico, A. C., Kurtz, S. M. S., Pincus, D. B., & Comer, J. S. (2014). Extending parent–child interaction therapy for early childhood internalizing problems: New advances for an overlooked population. *Clinical Child and Family Psychology Review*, 17, 340–356. doi:10.1007/s10567-014-0172-4
- Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, 67, 319–333. doi:10.1037/0022-3514.67.2.319
- Chaffin, M., Silovsky, J. F., Funderburk, B., Valle, L. A., Brestan, E. V., Balachova, T., ... Bonner, B. L. (2004). Parent-child interaction therapy with physically abusive parents: Efficacy for reducing future abuse reports. *Journal of Consulting and Clinical Psychology*, 72, 500–510. doi:10.1037/0022-006X.72.3.500
- Chen, E., Turiano, N. A., Mroczek, D. K., & Miller, G. E. (2016). Association of reports of childhood abuse and all-cause mortality rates in women. *JAMA Psychiatry*, 73, 920–927. doi:10.1001/jamapsychiatry.2016.1786
- Chevarley, F. M. (2003). *Statistical brief #12: Children's access to necessary health care, fall 2001*. Retrieved from http://meps.ahrq.gov/mepsweb/data_files/publications/st12/stat12.shtml
- Cho, E., Bearman, S. K., Woo, R., Weisz, J. R., & Hawley, K. M. (2020). A second and third look at FIRST: Testing adaptations of a principle-guided youth psychotherapy. *Journal of Clinical Child & Adolescent Psychology*, 1–14. Advance online publication. doi:10.1080/15374416.2020.1796678
- Churchill, R. C. (2016). The effectiveness, acceptability and cost-effectiveness of psychosocial interventions for maltreated children and adolescents: an evidence synthesis. *Health Technology Assessment*, 20, 1–508.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6, 284–290. doi:10.1037//1040-3590.6.4.284
- Cicchetti, D., & Toth, S. L. (1995). A developmental psychopathology perspective on child abuse and neglect. *Journal of the American Academy of Child & Adolescent Psychiatry*, 34, 541–565. doi:10.1097/00004583-199505000-00008
- Cicchetti, D., & Toth, S. L. (2009). A developmental psychopathology perspective on adolescent depression. In S. Nolen-Hoeksema, & L. M. Hilt (Eds.), *Handbook of Depression in Adolescents* (pp. 3–32). Routledge/Taylor & Francis Group.
- Cohen, P., Brown, J., & Smailes, E. (2001). Child abuse and neglect and the development of mental disorders in the general population. *Development and Psychopathology*, 13, 981–999. doi:10.1017/s0954579401004126
- Cohen, J. A., Deblinger, E., Mannarino, A. P., & Steer, R. A. (2004). A multi-site, randomized controlled trial for children with sexual abuse-related PTSD symptoms. *Journal of the American Academy of Child & Adolescent Psychiatry*, 42, 393–402. doi:10.1097/00004583-200404000-00005
- Cohen, J. A., Mannarino, A. P., & Iyengar, S. (2011). Community treatment of posttraumatic stress disorder for children exposed to intimate partner violence: A randomized controlled trial. *Archives of Pediatrics and Adolescent Medicine*, 165, 16–21. doi:10.1001/archpediatrics.2010.247
- Danese, A., & McEwen, B. S. (2012). Adverse childhood experiences, allostasis, allostatic load, and age-related disease. *Physiology and Behavior*, 106, 29–39. doi:10.1016/j.physbeh.2011.08.019
- de Kwaadsteniet, L., & Hagmayer, Y. (2018). Clinicians' personal theories of developmental disorders explain their judgments of effectiveness of interventions. *Clinical Psychological Science*, 6, 228–242. doi:10.1177/2167702617712270
- Dorsey, S., McLaughlin, K. A., Kerns, S. E. U., Harrison, J. P., Lambert, H. K., Briggs, E. C., ... Amaya-Jackson, L. (2017). Evidence base update for psychosocial treatments for children and adolescents exposed to traumatic events. *Journal of Clinical Child & Adolescent Psychology*, 46, 303–330.
- Dubowitz, H. (2014). The Safe Environment for Every Kid model: promotion of children's health, development, and safety, and prevention of child neglect. *Pediatric Annals*, 43, e271–e277. doi:10.3928/00904481-20141022-11
- Dubowitz, H., Feigelman, S., Lane, W., & Kim, J. (2009). Pediatric primary care to help prevent child maltreatment: The Safe Environment for Every Kid (SEEK) model. *Pediatrics*, 123, 858–864. doi:10.1542/peds.2008-1376
- Dubowitz, H., Feigelman, S., Lane, W., Prescott, L., Blackman, K., Grube, L., ... Tracy, J. K. (2007). Screening for depression in an urban pediatric primary care clinic. *Pediatrics*, 119, 435–443. doi:10.1542/peds.2006-2010
- Dubowitz, H., Lane, W. G., Semiatin, J. N., & Magder, L. S. (2012). The seek model of pediatric primary care: Can child maltreatment be prevented in a low-risk population? *Academic Pediatrics*, 12, 259–268. doi:10.1016/j.acap.2012.03.005
- Dubowitz, H., Prescott, L., Feigelman, S., Lane, W., & Kim, J. (2008). Screening for intimate partner violence in a pediatric primary care clinic. *Pediatrics*, 121, e85–e91. doi:10.1542/peds.2007-0904
- Eismann, E. A., Theuerling, J., Maguire, S., Hente, E. A., & Shapiro, R. A. (2019). Integration of the Safe Environment for Every Kid (SEEK) model across primary care settings. *Clinical Pediatrics*, 58, 166–176. doi:10.1177/0009922818809481
- Feigelman, S., Dubowitz, H., Lane, W., Prescott, L., Meyer, W., Tracy, J. K., ... Kim, J. (2009). Screening for harsh punishment in a pediatric primary care clinic. *Child Abuse & Neglect*, 33, 269–277. doi:10.1016/j.chiabu.2008.09.011
- Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., ... Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE) study. *American Journal of Preventive Medicine*, 14, 245–258. doi:10.1016/S0749-3797(98)00017-8
- Filele, J. H., Kaminski, J. W., Valle, L. A., & Cachat, P. (2013). Components associated with home visiting program outcomes: A meta-analysis. *Pediatrics*. doi:10.1542/peds.2013-1021H
- Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2019). Package 'irr': Various coefficients of interrater reliability and agreement author. Retrieved from <http://cran.r-project.org/package=irr>
- Garner, A. S., Shonkoff, J. P., Siegel, B. S., Dobbins, M. I., Earls, M. F., McGuinn, L., ... Wegner, L. M. (2012). Early childhood adversity, toxic

- stress, and the role of the pediatrician: translating developmental science into lifelong health. *Pediatrics*, 129, e224–e231. doi:10.1542/peds.2011-2662
- Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability: Key concepts, approaches, and applications. *Research in Social & Administrative Pharmacy*, 9, 330–338.
- Godoy, L., Long, M., Marschall, D., Hodgkinson, S., Bokor, B., Rhodes, H., ... Beers, L. (2017). Behavioral health integration in health care settings: Lessons learned from a pediatric hospital primary care system. *Journal of Clinical Psychology in Medical Settings*, 24, 245–258. doi:10.1007/s10880-017-9509-8
- Green, J. G., McLaughlin, K. A., Berglund, P. A., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., ... Kessler, R. C. (2010). Childhood adversities and adult psychiatric disorders in the national comorbidity survey replication I: Associations with first onset of DSM-IV disorders. *Archives of General Psychiatry*, 67, 113–123. doi:10.1001/archgenpsychiatry.2009.186
- Hall, E. (2009). The Delphi primer: Doing real-world or academic research using a mixed-method approach. In C. A. Lentz (Ed.), *The refractive thinker: Vol. 2: Research methodology* (pp. 3–27). Las Vegas, NV: Refractive Thinker Press.
- Hankin, B. L. (2020). Screening for and personalizing prevention of adolescent depression. *Current Directions in Psychological Science*, 29, 327–332. doi:10.1177/0963721420920231
- Holey, E. A., Feeley, J. L., Dixon, J., & Whittaker, V. J. (2007). An exploration of the use of simple statistics to measure consensus and stability in Delphi studies. *BMC Medical Research Methodology*, 7, 52. doi:10.1186/1471-2288-7-52
- Hsu, C. C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research & Evaluation*, 12, 10. doi:10.7275/PDZ9-TH90
- Jorm, A. F. (2015). Using the Delphi expert consensus method in mental health research. *Australian and New Zealand Journal of Psychiatry*, 49, 887–897. doi:10.1177/0004867415600891
- Kessler, R. C., McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., ... Williams, D. R. (2010). Childhood adversities and adult psychopathology in the WHO world mental health surveys. *British Journal of Psychiatry*, 197, 378–385. doi:10.1192/bjp.bp.110.080499
- Keyes, K. M., Eaton, N. R., Krueger, R. F., McLaughlin, K. A., Wall, M. M., Grant, B. F., ... Hasin, D. S. (2012). Childhood maltreatment and the structure of common psychiatric disorders. *British Journal of Psychiatry*, 200, 107–115. doi:10.1192/bjp.bp.111.093062
- Kolko, D. J., & Perrin, E. (2014). The Integration of Behavioral Health Interventions in Children's Health Care: Services, Science, and Suggestions. *Journal of Clinical Child & Adolescent Psychology*, 43, 216–228. doi:10.1080/15374416.2013.862804
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15, 155–163. doi:10.1016/j.jcm.2016.02.012
- Kupersmidt, J. B., Stelter, R., & Dodge, K. A. (2011). Development and validation of the social information processing patterns in elementary school-age boys. *Psychological Assessment*, 23, 834–847. doi:10.1037/a0023621
- Lane, W. G., Dubowitz, H., Feigelman, S., Kim, J., Prescott, L., Meyer, W., ... Tracy, J. K. (2007). Screening for parental substance abuse in pediatric primary care. *Ambulatory Pediatrics*, 7, 458–462. doi:10.1016/j.ambp.2007.07.007
- Leijten, P., Gardner, F., Melendez-Torres, G. J., van Aar, J., Hutchings, J., Schulz, S., ... Overbeek, G. (2019). Meta-analyses: Key parenting program components for disruptive child behavior. *Journal of the American Academy of Child & Adolescent Psychiatry*, 58, 180–190. doi:10.1016/j.jaac.2018.07.900
- Lenze, S. N., Pautsch, J., & Luby, J. L. (2011). Parent-child interaction therapy emotion development: A novel treatment for depression in preschool children. *Depression & Anxiety*, 28, 153–159. doi:10.1002/da.20770
- Lilienfeld, S. O. (2019). Reflections on clinical judgment and the dimensional-categorical distinction in the study of personality disorders: Comment on Bornstein (2019). *Personality Disorders: Theory, Research, and Treatment*, 10, 210–214. doi:10.1037/per0000318
- Lilienfeld, S. O., & Lynn, S. J. (2014). Errors/biases in clinical decision making. In R.L. Cautin, & S.O. Lilienfeld (Eds.), *The Encyclopedia of Clinical Psychology* (pp. 1–9). New York, NY: John Wiley & Sons, Inc. doi:10.1002/9781118625392.wbecp567
- Lindhiem, O., Bennett, C. B., Trentacosta, C. J., & McLearn, C. (2014). Client preferences affect treatment satisfaction, completion, and clinical outcome: A meta-analysis. *Clinical Psychology Review*, 34, 506–517. doi:10.1016/j.cpr.2014.06.002
- Luby, J. L., Barch, D. M., Whalen, D., Tillman, R., & Freedland, K. E. (2018). A randomized controlled trial of parent-child psychotherapy targeting emotion development for early childhood depression. *American Journal of Psychiatry*, 175, 1102–1110. doi:10.1176/appi.ajp.2018.18030321
- Luby, J. L., Lenze, S., & Tillman, R. (2012). A novel early intervention for preschool depression: Findings from a pilot randomized controlled trial. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 53, 313–322. doi:10.1111/j.1469-7610.2011.02483.x
- MacMillan, H. L., Fleming, J. E., Streiner, D. L., Lin, E., Boyle, M. H., Jamieson, E., ... Beardslee, W. R. (2001). Childhood abuse and lifetime psychopathology in a community sample. *American Journal of Psychiatry*, 158, 1878–1883. doi:10.1176/appi.ajp.158.11.1878
- Manly, J. T., Kim, J. E., Rogosch, F. A., & Cicchetti, D. (2001). Dimensions of child maltreatment and children's adjustment: Contributions of developmental timing and subtype. *Development and Psychopathology*, 13(4), 759–782. doi:10.1017/S0954579401004023
- Martini, R., Hilt, R., Marx, L., Chenven, M., Naylor, M., Sarvet, B., ... Ptakowski, K. K. (2012). *Best principles for integration of child psychiatry into the pediatric health home*. Washington, DC: American Academy of Child & Adolescent Psychiatry.
- Masten, A. S., & Cicchetti, D. (2010). Developmental cascades. *Development and Psychopathology*, 22(3), 491–495. doi:10.1017/S0954579410000222
- McLaughlin, K. A. (2016). Future Directions in Childhood Adversity and Youth Psychopathology. *Journal of Clinical Child & Adolescent Psychology*, 45, 361–382. doi:10.1080/15374416.2015.1110823
- McLaughlin, K. A., Colich, N. L., Rodman, A. M., & Weissman, D. G. (2020). Mechanisms linking childhood trauma exposure and psychopathology: A transdiagnostic model of risk and resilience. *BMC Medicine*, 18, 96. doi:10.1186/s12916-020-01561-6
- McLaughlin, K. A., Conron, K. J., Koenen, K. C., & Gilman, S. E. (2010). Childhood adversity, adult stressful life events, and risk of past-year psychiatric disorder: A test of the stress sensitization hypothesis in a population-based sample of adults. *Psychological Medicine*, 40, 1647–1658. doi:10.1017/S0033291709992121
- McLaughlin, K. A., DeCross, S. N., Jovanovic, T., & Tottenham, N. (2019). Mechanisms linking childhood adversity with psychopathology: Learning as an intervention target. *Behaviour Research and Therapy*, 118, 101–109. doi:10.1016/j.brat.2019.04.008
- McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., & Kessler, R. C. (2012). Childhood adversities and first onset of psychiatric disorders in a national sample of US adolescents. *Archives of General Psychiatry*, 69, 1151–1160. doi:10.1001/archgenpsychiatry.2011.2277
- McLaughlin, K. A., & Lambert, H. K. (2017). Child trauma exposure and psychopathology: mechanisms of risk and resilience. *Current Opinion in Psychology*, 14, 29–34. doi:10.1016/j.copsyc.2016.10.004
- McNeil, C. B., & Hembree-Kigin, T. L. (2010). *Parent-child interaction therapy*. New York, NY: Springer Science+Business Media.
- Nock, M. K., Wedig, M. M., Holmberg, E. B., & Hooley, J. M. (2008). The Emotion Reactivity Scale: Development, evaluation, and relation to self-injurious thoughts and behaviors. *Behavior Therapy*, 39, 107–116. doi:10.1016/j.beth.2007.05.005
- Okamura, K. H., Orimoto, T. E., Nakamura, B. J., Chang, B., Chorpita, B. F., & Beidas, R. S. (2020). A history of child and adolescent treatment through a distillation lens: Looking back to move forward. *Journal of Behavioral Health Services & Research*, 47, 70–85. doi:10.1007/s11414-019-09659-3
- Pardee, M., Kuzma, E., Dahlem, C. H. G. Y., Boucher, N., & Darling-Fisher, C. S. (2017). Current state of screening high-ACE youth and emerging adults in primary care. *Journal of the American Association of Nurse Practitioners*, 29, 716–724. doi:10.1002/2327-6924.12531
- Powell, B. J., Bosk, E. A., Wilen, J. S., Danko, C. M., Van Scoyoc, A., & Banman, A. (2015). Evidence-based programs in “real world” settings: Finding the best fit. In D. Daro, A. C. Donnelly, L. A. Huang, & B. J.

- Powell (Eds.), *Advances in child abuse prevention knowledge* (pp. 145–177). Cham, Switzerland: Springer.
- R Core Team. (2019). R: A language and environment for statistical computing Version 3.5.3. Vienna: R Foundation for Statistical Computing.
- Sanchez, A. L., Cornacchio, D., Poznanski, B., Golik, A. M., Chou, T., & Comer, J. S. (2018). The effectiveness of school-based mental health services for elementary-aged children: A meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, *57*, 153–165. doi:10.1016/j.jaac.2017.11.022
- Schleider, J. L., & Weisz, J. R. (2017). Little treatments, promising effects? Meta-analysis of single-session interventions for youth psychiatric problems. *Journal of the American Academy of Child & Adolescent Psychiatry*, *56*, 107–115. doi:10.1016/j.jaac.2016.11.007
- Schmidt, R. C. (1997). Managing Delphi surveys using nonparametric statistical techniques. *Decision Sciences*, *28*, 763–774. doi:10.1111/j.1540-5915.1997.tb01330.x
- Singla, D. R., Waqas, A., Hamdani, S. U., Suleman, N., Zafar, S. W., Zill-e-Huma, ., ... Rahman, A. (2020). Implementation and effectiveness of adolescent life skills programs in low- and middle-income countries: A critical review and meta-analysis. *Behaviour Research and Therapy*, *130*, 103402. doi:10.1016/j.brat.2019.04.010
- Sun, M., Rith-Najarian, L. R., Williamson, T. J., & Chorpita, B. F. (2019). Treatment features associated with youth cognitive behavioral therapy follow-up effects for internalizing disorders: A meta-analysis. *Journal of Clinical Child & Adolescent Psychology*, *48*, S269–S283. doi:10.1080/15374416.2018.1443459
- Surowiecki, J. (2004). *The Wisdom of Crowds: Why the Many are Smarter Than the Few and how Collective Wisdom Shapes Business, Economies, Societies, and Nations*. New York, NY: Doubleday.
- Thomas, R., & Zimmer-Gembeck, M. J. (2007). Behavioral outcomes of parent-child interaction therapy and triple P-positive parenting program: A review and meta-analysis. *Journal of Abnormal Child Psychology*, *35*, 475–495. doi:10.1007/s10802-007-9104-9
- Timmer, S. G., Urquiza, A. J., Zebell, N. M., & McGrath, J. M. (2005). Parent-child interaction therapy: Application to maltreating parent-child dyads. *Child Abuse & Neglect*, *29*, 825–842. doi:10.1016/j.chiabu.2005.01.003
- Trevelyan, E. G., & Robinson, N. (2015). Delphi methodology in health research: How to do it? *European Journal of Integrative Medicine*, *7*, 423–428. doi:10.1016/j.eujim.2015.07.002
- van der Put, C. E., Assink, M., Gubbels, J., & Boekhout van Solinge, N. F. (2018). Identifying effective components of child maltreatment interventions: A meta-analysis. *Clinical Child and Family Psychology Review*, *21*, 171–202. doi:10.1007/s10567-017-0250-5
- van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., Coughlan, B., & Reijman, S. (2020). Annual research review: Umbrella synthesis of meta-analyses on child maltreatment antecedents and interventions: differential susceptibility perspective on risk and resilience. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *61*, 272–290. doi:10.1111/jcpp.13147
- Wegman, H. L., & Stetler, C. (2009). A meta-analytic review of the effects of childhood abuse on medical outcomes in adulthood. *Psychosomatic Medicine*, *71*, 805–812. doi:10.1097/PSY.0b013e3181bb2b46
- Weich, S., Patterson, J., Shaw, R., & Stewart-Brown, S. (2009). Family relationships in childhood and common psychiatric disorders in later life: systematic review of prospective studies. *British Journal of Psychiatry*, *194*, 392–398. doi:10.1192/bjp.bp.107.042515
- Weisberg, H. F. (1992). Measures of Spread. In H. F. Weisberg (Ed.), *Quantitative applications in the social sciences: Central tendency and variability* (pp. 46–75). Thousand Oaks, CA: Sage.
- Weisz, J. R., & Bearman, S. (2020). *Principle-guided psychotherapy for children and adolescents: The FIRST treatment program for behavioral and emotional problems*. New York, NY: Guilford Press.
- Weisz, J., Bearman, S. K., Santucci, L. C., & Jensen-Doss, A. (2017). Initial test of a principle-guided approach to transdiagnostic psychotherapy with children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, *46*, 44–58. doi:10.1080/15374416.2016.1163708
- Weisz, J. R., Kuppens, S., Ng, M. Y., Eckshtain, D., Ugueto, A. M., Vaughn-Coaxum, R., ... Fordwood, S. R. (2017). What five decades of research tells us about the effects of youth psychological therapy: A multi-level meta-analysis and implications for science and practice. *American Psychologist*, *72*, 79–117.
- Weisz, J. R., Sandler, I. N., Durlak, J. A., & Anton, B. S. (2005). Promoting and protecting youth mental health through evidence-based prevention and treatment. *American Psychologist*, *60*, 628–648. doi:10.1037/0003-066X.60.6.628
- Whiteside, S. P. H., Sim, L. A., Morrow, A. S., Farah, W. H., Hilliker, D. R., Murad, M. H., ... Wang, Z. (2020). A Meta-analysis to guide the enhancement of CBT for childhood anxiety: Exposure over anxiety management. *Clinical Child and Family Psychology Review*, *23*, 102–121. doi:10.1007/s10567-019-00303-2
- Witteman, C. L. M. (2018). Concluding commentary: Clinical decision making. *Clinical Psychological Science*, *6*, 266–270. doi:10.1177/2167702617745642
- Zapf, A., Castell, S., Morawietz, L., & Karch, A. (2016). Measuring inter-rater reliability for nominal data: Which coefficients and confidence intervals are appropriate? *BMC Medical Research Methodology*, *16*, 93. doi:10.1186/s12874-016-0200-9
- Zeman, J., & Garber, J. (1996). Display rules for anger, sadness, and pain: It depends on who is watching. *Child Development*, *67*, 957–973. doi:10.1111/j.1467-8624.1996.tb01776.x
- Zigler, E. (1994). Reshaping early childhood intervention to be a more effective weapon against poverty. *American Journal of Community Psychology*, *22*, 37–47. doi:10.1007/BF02506816
- Zigler, E. (2000). Foreword. In J. P. Shonkoff & S. J. Meisels (Eds.), *Handbook of early childhood intervention* (2nd ed., pp. xi–xv). Cambridge: Cambridge University Press.
- Zigler, E., Pfannenstiel, J. C., & Seitz, V. (2008). The parents as teachers program and school success: A replication and extension. *Journal of Primary Prevention*, *29*, 103–120. doi:10.1007/s10935-008-0132-1
- Zigler, E., & Styfco, S. J. (2001). Extended childhood intervention prepares children for school and beyond. *Journal of the American Medical Association*, *285*, 2378–2380. doi:10.1001/jama.285.18.2378
- Zigler, E., & Styfco, S. J. (2002). A life lived at the crossroads of knowledge and children's policy. *New Directions for Child and Adolescent Development*, *98*, 5–16.