Clinical Child and Family Psychology Review (2021) 24:684–706
https://doi.org/10.1007/s10567-021-00362-4

A Meta-Analytic Review of Emotion Regulation Focused Psychosocial Interventions for Adolescents

Hana-May Eadeh[1] · Rosanna Breaux[2] · Molly A. Nikolas[1]

Accepted: 5 July 2021 / Published online: 18 July 2021

Abstract

Emotion regulation (ER) is the ability to monitor, evaluate, and modify one's emotional responses to be appropriate for environmental demands. Poor ER has been considered a transdiagnostic risk factor for a range of internalizing and externalizing disorders and overall decreased well-being in adolescents. A range of evidence-based interventions exist which may improve ER. However, much of the intervention research to date does not include a measure of ER to assess change pre and post treatment, with limited information about the efficacy of these interventions in youth across a range of sample types. There is a clear need for a comprehensive review of the literature examining ER-focused interventions in adolescents with a wide range of presenting disorders. A literature search was originally conducted in January 2020 and an updated search was conducted in February 2021 which elicited 1245 articles, of which 605 were duplicates and were removed. Abstracts of the remaining 640 articles were screened with 121 articles being reviewed in full. Of note, 16 additional articles were identified through references and other sources during this process and were also included in the full review. Of the 137 articles reviewed in full, 41 studies were ultimately included in the present review. The present paper provides a descriptive review of intervention approaches and findings from community prevention programs, programs for war-affected youth, programs for clinical populations, and programs for incarcerated and delinquent adolescents. The overall pooled effect was significantly different from zero based on the pre/post effects [Hedge's $g = 0.29$, 95% CI (0.22, 0.36)] and the intervention/control effects [Hedge's $g = 0.19$, 95% CI (0.06–0.32)]. Although neither sex nor age significantly accounted for heterogeneity in effect sizes, there were significant findings for population type (clinical vs. community), with community samples having significantly lower effect sizes on average. Impacts of the different ER measures used and significant methodological variability (e.g., use of control groups, length of intervention) across included studies are discussed. Implications and suggestions for future research are reviewed, specifically, that additional understanding of moderators of effects are needed and that measures used to assess change in ER, both dysregulation and adaptive skill use, may need to more directly align with the intervention's focus and the strategies taught as part of the intervention.

Keywords Emotion regulation · Meta-analysis · Adolescent · Intervention · Review

Introduction

There has been a proliferation of research over the past 20 years regarding the processes that contribute to successful emotion regulation (ER) and the importance of adaptive ER skills for overall well-being. ER is defined as the ability to monitor, evaluate, and modify one's emotional responses that appropriately suits environmental demands. A widely used process model of ER (Gross, 1998) describes five types of ER strategies: (1) situation selection, avoiding or approaching a potentially emotionally ladened situation; (2) situation modification, changing aspects of the situation or environment; (3) attentional deployment, influencing perception of the situation or environment; (4) cognitive change, changing cognitive representations; and (5) response modulation, changing emotion-related behaviors. This work has been extended to describe three stages of ER, including (1) identification, which involves distinguishing emotional responses and deciding whether to regulate emotion;

✉ Hana-May Eadeh
hana-may-eadeh@uiowa.edu

[1] Department of Psychological and Brain Sciences, University of Iowa, 340 Iowa Ave, Iowa City, IA 52242, USA

[2] Department of Psychology, Virginia Tech, Blacksburg, VA, USA

(2) selection, which involves choosing which strategies to use; and (3) implementation, which involves applying the appropriate strategy (Gross, 2015). Thus, conceptually, these models of ER lay the groundwork for multi-faceted intervention approaches that can improve adaptive ER strategies and/or reduce maladaptive ER strategies.

Much of the empirical work to date has sought to understand how youth learn these ER strategies (e.g., Kopp, 1989; Kopp & Neufeld, 2003), as well as the short and long-term consequences of deficits in ER during adolescence. Recently, poor ER (either lack of adaptive ER strategies or increased use of maladaptive ER strategies) has been considered as a transdiagnostic risk factor for a range of internalizing and externalizing disorders, academic failure, and overall decreased well-being in adolescents (Aldao et al., 2010; Ehring et al., 2010; Shapero et al., 2016). Additionally, ER has been supported as a mediator between psychopathology and both functional outcomes and comorbid conditions (Bunford et al., 2018; Fogleman et al., 2019; Graziano et al., 2007; Walerius et al., 2018). Further, youth displaying more adaptive ER show improved well-being and more positive outcomes than those who display maladaptive ER (McLaughlin et al., 2011).

Although ER abilities begin to develop during the preschool years and continue into childhood (Zeman et al., 2006), adolescence may be a specific period during which youth use less adaptive and more maladaptive ER strategies compared with both younger and older individuals (Cracco et al., 2017). Adolescents are undergoing a range of developmental changes (i.e., physical development during puberty, increased interpersonal salience of peers and sensitivity to social context, increased expectation and demand for autonomy; Blakemore & Mills, 2014; Steinberg, 2014), and adaptive navigation of this time relies on learning and implementing adaptive ER strategies (and also minimizing the use of maladaptive ER strategies). Adolescents also begin to differentiate their use of ER strategies depending upon socio-emotional context as well as motivation and emotion type (Zeman et al., 2006). Further, research in social neuroscience has largely supported the notion that the brain is particularly sensitive to social context during adolescence (Schriber & Guyer, 2016), indicating that this period may be particularly salient for developing, implementing, and modifying ER strategies that can promote better social adjustment (Chervonsky & Hunt, 2019). Thus, improving adaptive ER abilities during adolescence may be particularly beneficial for helping youth navigate the challenges and opportunities they encounter during this developmental period.

A range of interventions have been developed to improve ER abilities, with ER being shown to be a malleable construct and a potentially salient target for treatment (Gross, 2002). Additionally, improvements in ER have been associated with improvements in psychopathology for a range of disorders (e.g., Wijana et al., 2018; Wolff & Ollendick, 2012). However, much of the research with these intervention methods is with adults or specific clinical populations [e.g., dialectical behavior therapy (DBT) for individuals with borderline personality disorder] and less is understood about their efficacy in community samples and/or adolescents. In this review, we provide a brief synopsis of why ER-focused interventions are needed, particularly during adolescence, a narrative overview of ER focused interventions for adolescents, estimates of effect size for ER focused interventions for improving adolescent ER abilities, and a discussion of the implications for future research.

A Need for ER-Focused Interventions

Although a range of evidence-based interventions exist which may help improve ER skills, most do not explicitly target ER. One exception is the recent Unified Protocol, which has a child and adolescent version (Ehrenreich-May et al., 2017) and is based in cognitive behavioral therapy (CBT). The Unified Protocol takes a transdiagnostic approach to treating all emotional disorders, in part, by targeting the underlying emotional difficulties and emotional dysregulation (e.g., Ellard et al., 2010). Additionally, although traditional CBT (see Beck & Beck, 2011) or DBT (see Linehan, 1993) interventions may indirectly improve ER by challenging negative cognitions or training competing behaviors, teaching ER skills is not always the primary focus of the treatment nor is it a primary measure used in efficacy or effectiveness evaluation (e.g., Hadley et al., 2017; Smith et al., 2018). In fact, it has been posited that the effectiveness for CBT in treating some disorders, such as anxiety, is limited because it does not target ER, when those are the very skills that need building (Hannesdottir & Ollendick, 2007). In line with this, ER deficits have been found to be a central component of psychopathology in youth (e.g., Compas et al., 2014, 2017) and thus treatments that specifically target improving these skills, regardless of diagnostic status (i.e., Unified Protocol) may be especially relevant for reducing psychological distress and improving overall well-being. This may be due to the fact that some ER strategies work transdiagnostically regardless of the context or diagnosis (e.g., deep breathing), whereas others may be specific to a certain context or population (e.g., distraction).

Limitations of Existing Reviews

Targeting adaptive ER development as a specific intervention component as well as assessing changes in ER pre and post treatment is an important next step to understanding the efficacy of working to improve ER and subsequently reducing psychopathology. ER interventions have been used in a wide range of clinical, non-clinical, and medical populations

from childhood through adulthood and in both individual and group formats (e.g., Smyth & Arigo, 2009). These interventions have also been applied to a wide range of psychological disorders and thus may have differing levels of effectiveness depending on the symptoms present. It is possible certain facets of ER (e.g., identification) may be important for some populations or settings, such as school-wide prevention programs where rates of emotion dysregulation are lower, whereas other facets (e.g., selection and implementation of targeted strategies, such as challenging negative thoughts for adolescents with internalizing disorders, but engaging in positive coping behaviors to calm down for adolescents with externalizing disorders) will be most integral for clinical populations who experience high levels of emotion dysregulation. However, the overall effectiveness of interventions for improving ER across heterogenous samples of adolescents has not been investigated.

A recent systematic review of the intervention literature broadly spanning adolescence and adulthood found that ER deficits decreased throughout the course of intervention regardless of the psychopathology present (the review examined depression, anxiety, eating disorders, substance use disorders, and borderline personality disorders), but with large variation in the effect sizes (Cohen's $d = 0.18–2.87$; Sloan et al., 2017). These results support the possibility that ER may be operating as a transdiagnostic factor, given that findings were consistent regardless of diagnostic status of the participants in the intervention. Notably, only five of the 67 included studies included adolescent samples and thus, the effects were largely relevant to adult samples. Moreover, the Sloan et al. (2017) review focused on clinical populations and did not include prevention programs or community samples, for which ER interventions may have differential effects. Specifically, it is likely that clinical populations and community populations have different baseline levels of dysregulation and adaptive ER skill use; therefore, there may be more or less room to move, so to speak, during and after going through an intervention program. This is an important point of consideration when understanding how well and for whom ER interventions work. Additionally, another recent narrative review focused on 12 ER interventions with an emotion socialization component for young children (ages 0–6; England-Mason & Gonzalez, 2020). Findings indicated that although the included interventions may show promising results for improving parenting behaviors related to ER, there is limited evidence for how these programs improve the child's ER abilities.

Finally, Moltrecht et al. (2020) reviewed articles and completed a meta-analysis of published studies through April of 2018 using samples of youth aged 6–24. The authors limited their search to psychiatric diagnoses with ER difficulties. The authors found moderate improvements in reducing emotion dysregulation (Hedge's $g = -0.46$) and increasing adaptive ER (Hedge's $g = 0.36$) with large heterogeneity across studies. Importantly, the authors chose to exclude samples with autism, intellectual disability, and specialized medical populations, despite that these populations, specifically autism (Mazefsky et al., 2013) and intellectual disability (McClure et al., 2009) frequently have difficulties with ER. Furthermore, studies with no control group were excluded, which likely excluded many relevant interventions in early stages of development and evaluation. Given the proliferation of interventions to improve ER during the past year (i.e., nearly as many articles were identified in the search for the past year for the present study as were in the search for anytime up until 2020) it is imperative to include work at all stages of development and evaluation. Although Moltrecht et al. (2020) reported non-significant changes in overall effect sizes across different developmental periods, the number of studies in each age subgroup was small. Therefore, the impact of age and development on the efficacy of ER interventions remains an open question.

A more focused examination of adolescents specifically is needed given the salience of this period in both the development of ER strategies (Ahmed et al., 2015; Riediger & Klipker, 2014) as well as the relevance of the adolescent period for the onset and continuation of psychopathology into adulthood (see Holmbeck et al., 2006). Specifically, adolescence involves a confluence of developmental changes that include significant biological and physical changes, increased cognitive capacity, and heightened demands on navigating interpersonal relationships (Casey et al., 2010). Prior work has suggested that adolescence may be a critical time of heightened plasticity to the social environment (Blakemore, 2008; Steinberg, 2005), given that the neural architecture underpinning the cognitive and social processes required for successful ER continues to develop during this period (Casey & Caudle, 2013). Successful peer relationships and psychosocial functioning may be, in part, dependent on employing adaptive ER strategies (Cronin et al., 2018). As such, this may be a salient time to teach adolescents adaptive ways to manage their emotions, possibly having a positive impact on their overall development. Thus, adolescence may represent a critical time for both reducing maladaptive ER as well as for enhancing adaptive ER to promote a healthy transition to adulthood.

This review seeks to address that gap by providing a comprehensive review of the literature examining ER-focused interventions in adolescents with a wide range of presenting disorders and including a meta-analysis of effect sizes. This review and meta-analysis builds off prior work by (1) including studies that examined both clinical and non-clinical samples that implemented ER-specific interventions, (2) systematically describing the interventions, target, and measures included in interventions in addition to providing meta-analytic estimates of effect sizes, and (3) focusing on

adolescence specifically to determine the specific effectiveness of these approaches during this developmental period. This synthesis is crucial for understanding the current effectiveness of the rapidly growing ER intervention literature.

Methods

Since the focus of the present review is psychosocial interventions for adolescents addressing ER, included studies must have an ER component, consistent with Gross (1998, 2015) process model of ER, taught as part of the intervention, with an adolescent sample whose *mean* age[1] falls within the World Health Organization's definition of adolescence (i.e., 10–19 years old). Additionally, a measure of ER must have been included at pre- and post-intervention. Much of the research is based upon small pilot studies, which use a wide range of techniques and have varying levels of methodological rigor (e.g., presence of comparison groups and type, length of follow-up periods, type of ER measure). Given the literature focusing on transdiagnostic, ER-based interventions, is still in the early stages, studies were included regardless of these methodological factors. These methodological differences will be accounted for in the meta-analysis and their potential impact on the results will be discussed.

Given the range of study populations and the number and diversity of outcomes examined in this literature, this review specifically focused on whether studies showed improvement in different component processes of ER rather than the overall effect sizes on the psychopathology or behavioral outcomes of interest. This focus was established for several reasons. First, although ER is related to multiple psychopathology outcomes investigated in these studies, the effects are likely not uniform across psychopathology dimensions, complicating interpretation of findings. For example, variations in effect sizes in outcomes could reflect both less change in ER associated with that intervention or could be due to a relatively smaller association between ER and the psychopathology outcome examined in any given study. Second, prior reviews have addressed efficacy for many interventions in targeting specific behaviors, even those related to ER (e.g., self-harm, interpersonal conflict, substance use; Ougrin et al., 2015). As such, the present review examined change in ER measures specifically to maintain its focus on this transdiagnostic process.

[1] Studies may still be included if the minimum and maximum ages included in the sample are outside of 10–19 years, if the *mean* age is within the 10–19-year range.

Literature Search

A literature search was initially conducted in January 2020 and again prior to resubmission in February 2021; a second search was conducted due to the accelerated rate at which this work has been published and to ensure all recent and relevant studies were eligible for inclusion (see Fig. 1 for a flow diagram). The following search combinations were used in PubMed and PsychInfo: "emotion reg*", AND "intervention" OR "treatment" OR "training" OR "therapy" AND "adolescen*" OR "youth" contained in any field. The search elicited 1245 articles, of which 605 were duplicates and were subsequently removed. Abstracts of the remaining 640 articles were screened for inclusion in the review based on the following criteria: (1) published in English in a peer-reviewed journal, (2) paper reports outcomes from a psychosocial intervention, (3) manuscript authors described their intervention as targeting ER and/or coping skills, and upon review by the first and second author, it fit within a broad conceptualization of ER as defined by the ER process model (Gross, 1998, 2015), (4) a pre and post measure of ER is reported, and (5) target population is adolescents per World Health Organization's definition (M_{age} 10–19). Exclusion criteria included: qualitative case studies and samples comprised of those with substance use disorders, eating disorders, and mania or psychosis. These clinical populations were excluded for this review, given that individuals with these disorders are frequently treated in a medical milieu/inpatient treatment setting due to comorbid physical health dangers and/or primary use of psychotropic medications. Additionally, these criteria likely excluded some studies and samples whose interventions may have a significant impact on ER (e.g., interventions targeting specific problem behaviors). This work has been extensively reviewed and as such is not the focus of the present review (e.g., Mingebach et al., 2018; Reyno et al., 2006; Sukhodolsky et al., 2004).

Of the 640 abstracts screened, 121 articles appeared to pass inclusion criteria and were thus reviewed in full. Of note, 16 additional articles were identified through references and other sources during this process and were also include in the full review. Of the 137 articles reviewed in full, 46 studies met all the inclusion criteria; 4 of these studies met the exclusion criteria and were not included in the review. One additional study was secondary data analysis from the same sample of an already included study; as such, only the primary outcomes paper (published first) was included. This yielded a total of 41 studies meeting all inclusion criteria and no exclusion criteria and being included in the present review. See Table 1 for all included articles and their intervention and sample characteristics; see Fig. 1 for the PRISMA diagram.

Studies of ER interventions in adolescents included multiple treatment approaches, teaching both specific skills as

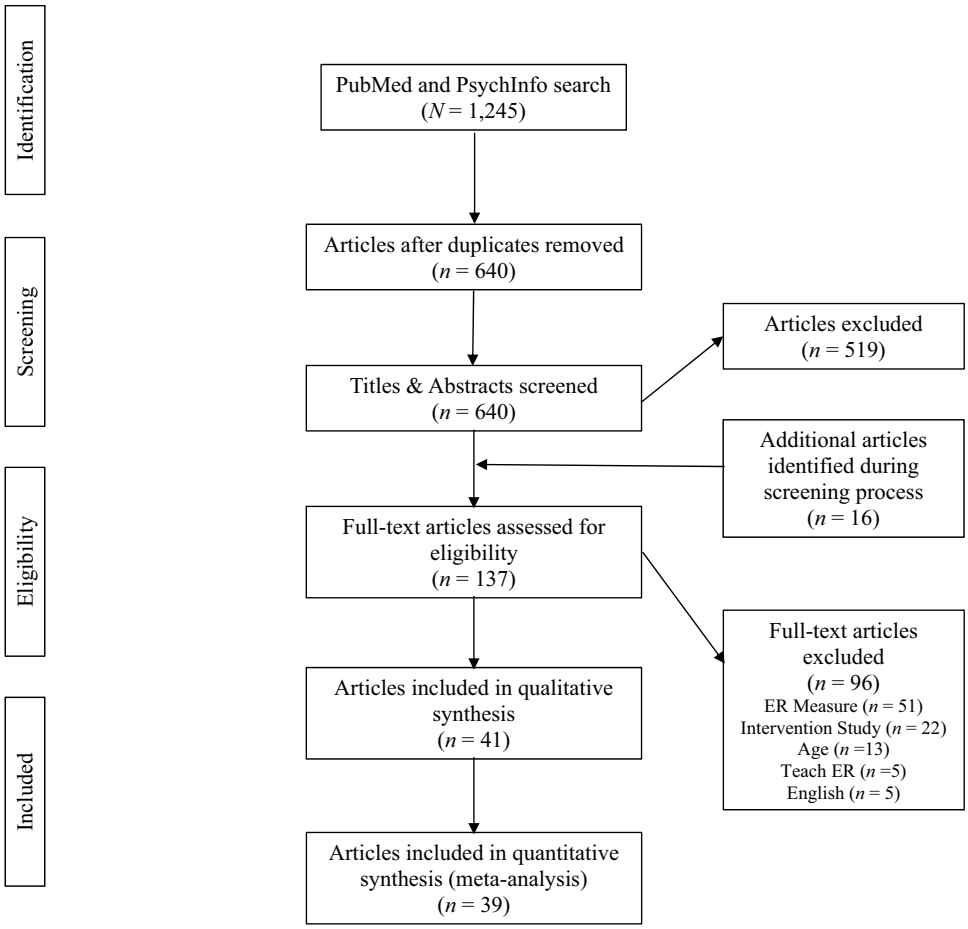


Fig. 1 PRISMA Flow Diagram of Literature Review. Inclusion criteria were as follows: (1) Published in English in peer-reviewed journal; (2) Empirical study evaluating an intervention; (3) Intervention teaches ER skills; (4) ER measure at pre and post; 5) $M_{age} = 10–19$

well as aiming to reduce dysregulation, varied populations of interest, and numerous measures of ER (see Table 2). To facilitate interpretation of findings, we grouped the studies based upon target sample and describe the overall approach and main findings for each study within each of the following sample groups: (1) community and prevention programs or populations, (2) programs for war-affected youth, (3) programs for clinical populations (i.e., externalizing, internalizing, borderline personality, inpatient, and specialized medical/health), and (4) programs for incarcerated or delinquent adolescents.

Pooled Effect Size Analysis

Standardized mean differences were calculated for all measures of ER and included in the narrative review for each study. Pre- and post-intervention means and standard deviations were used to calculate standardized mean differences for all ER measures in line with recommendations by Morris (2008) that minimize bias and enhance precision. Effect sizes were calculated so that improvements in ER (i.e., increases in adaptive ER strategies or ER regulation, declines in maladaptive ER strategies or emotion dysregulation) were always scaled positively, whether calculated from pre- to post-intervention or between intervention and control groups. As such, any negative effect sizes reflect decreases in ER abilities or increases in emotion dysregulation. For studies with follow-up data beyond post-intervention, we included effect sizes in the narrative review, but limited analysis of pooled effect sizes to post-intervention scores only. Further, we elected to convert all mean differences to Hedge's *g* due to the variation in sample size and the significant number of pilot studies with relatively smaller samples. Two studies were not included in the effect size analysis due to insufficient data presented in the original manuscript and the inability to contact the authors. As such, a total of 38 studies were included in the effect size analyses, with some papers having multiple effects included due to more than one measure of ER. Effects from the same study were weighted to account for non-independence so that no single study or effect disproportionately accounted for the results (Borenstein et al., 2011). Given the large range in methodological rigor of the included studies, a random effects model was used and pooled effects were examined separately for all studies using pre and post data and again only for those studies with an intervention and control group.

Table 1 Study and sample characteristics of included articles ($N=41$)

Authors	Intervention	Sample characteristics				Study characteristics	
		N	Age	% female	Sample type	Design	Control group
Afshari et al. (2014)	ECBT	30	10.57	52.5	Anxiety	RCT	CBT and no treatment
Alampay et al. (2020)	M-BCT	186	11.88	58.6	Low SES Filipino	RCT	Handicrafts
Betancourt et al. (2014)	YRI	436	17.90	45.6	War-affected	RCT	Control[b]
Bjureberg et al. (2017)	ERITA	17	15.31	100.0	NSSI	Pilot	N/A
Bjureberg et al. (2018)	ERITA (online)	25	15.70	76.0	NSSI	Pilot	N/A
Breaux and Langberg (2020)	RELAX	18	13.50	30.0	ADHD	Pilot	N/A
Broderick and Metz (2009)	L2B	137	16.9	100.0	Community	Pilot	IAU
Claro et al. (2015)	CERTIFY	28	14.20	29.0	School failure risk	Pilot	Usual school supports
Conner et al. (2019)	EASE	20	14.94	11.8	ASD	Pilot	N/A
Cotter et al. (2020)	L2B	11	14.36	73.0	Obesity	Pilot	N/A
Cotton et al. (2016)	M-BCT	10	13.20	80.0	Anxiety	Pilot	N/A
Cotton et al. (2020)	M-BCT	33	13.7	75.0	Anxiety w/BD risk	Pilot	Waitlist
Dingle et al. (2016)	Tuned In	41; 216	14.80; 13.60	24.0	At-risk/Mainstream	Pilot	N/A
Ford et al. (2012)	TARGET	59	14.70	100.0	Delinquent females	RCT	Enhanced TAU
Ford-Paz et al. (2019)	WOW	960	N/A[a]	100.0	Community females	Pilot	N/A
Fung et al. (2016)	L2B	19	12.7	57.9	Ethnic minority w/ elevated mood symptoms	Pilot	Waitlist
Hadley et al. (2020)	HealthTRAC	38	14.71	60.0	Obesity	RCT	Std Behavioral Weight Control
Hafeman et al. (2020)	Mindfulness	17	12.1	60.0	At-risk for BD	Pilot	N/A[c]
Houck et al. (2016)	ERI	420	13.00	47.0	Community	RCT	Health Promotion
Johnstone et al. (2020)	ER Program	295	11.04	52.5	Community	RCT	BA and no treatment control
Keiley (2007)	MFGI	73	15.60	41.1	Incarcerated	Pilot	N/A
Keiley et al. (2015)	MFGI	115	15.70	0.0	Incarcerated males	Pilot	N/A
Lam and Seiden (2020)	L2B	115	12.4	34.8	Lowest academic tier; Hong Kong	RCT	IAU
Lau et al. (2020)	L2B	111	15.18	73.0	Elevated depression	RCT	Interpersonal Therapy Skills
Lindqvist et al. (2020)	Affect-focused IPDT	76	16.60	80.0	Depression	RCT	Supportive Therapy
Metz et al. (2013)	L2B	216	16.5	34.1	Community	Pilot	IAU
Newnham et al. (2015)	YRI	32	18.20	50.0	War-affected	Pilot	N/A
Punamäki et al. (2014)	TRT	482	11.29	49.4	War-affected	RCT	Waitlist
Ramdhonee-Dowlot et al. (2021)	SSL	100	11.75	76.0	Mauritian in RCIs	RCT	Waitlist
Riggs Romaine et al. (2018)	JJAM	57	17.45	100.0	Incarcerated females	RCT	TAU
Rizzo et al. (2020)	STRONG	119	13.04	0.0	Community males	RCT	Waitlist
Santomauro et al. (2016)	CBT with ER	20	15.75	40.0	ASD and Depression	RCT	Waitlist
Schoeps et al. (2018)	EEP	360	12.63	57.0	Community	RCT	Usual school supports
Schuppert et al. (2009)	ERT	43	16.14	88.4	Borderline traits	Pilot RCT	TAU
Schuppert et al. (2012)	ERT	109	15.98	96.0	Borderline traits	RCT	TAU
Suveg et al. (2009)	CBT	37	10.47	40.5	Anxiety	Pilot	N/A
Thomson et al. (2015)	SAS-OR	13	10.40	100.0	ASD (males)	Pilot	N/A
Wijana et al. (2018)	ICT	49	14.60	85.7	NSSI and SI	Pilot	N/A
Wolff and Ollendick (2012)	CBT	5	12.60	60.0	ODD or CD	Pilot	N/A

Table 1 (continued)

Authors	Intervention	Sample characteristics				Study characteristics	
		N	Age	% female	Sample type	Design	Control group
Wolff et al. (2020)	CBT	30	11.55	66.0	Depression and conduct problems	Pilot RCT	N/A[d]
Yeo et al. (2020)	DBT-A	51	14.73	69.3	Ethnic minority w/ NSSI	Pilot	N/A

Sample type and study design are as stated in the manuscript

YRI Youth Readiness Intervention, *ERITA* Emotion Regulation Individual Therapy for Adolescents, *RELAX* Regulating Emotions Like An eXpert, *L2B* Learning to BREATHE, *CERTIFY* Cognitive Emotion Regulation Training Intended for Youth, *EASE* Emotional Awareness and Skills Enhancement, *M-CBT* Mindfulness Cognitive Behavioral Therapy, *TARGET* Trauma Affect Regulation Guide for Education and Therapy, *WOW* Working on Womanhood, *ERI* Emotion Regulation Intervention, *ER* emotion regulation, *MFGI* Multiple Family Group Intervention, *IPDT* internet based psychodynamic therapy, *TRT* Teaching Recovery Techniques, *SSL* super skills for life, *JJAM* Juvenile Justice Anger Management, *STRONG* domestic violence prevention program, *EEP* Emotion Education Program, *ERT* Emotion Regulation Training, *SAS-OR* Secret Agent Society-Operation Regulation, *ICT* Intensive Contextual Treatment, *NSSI* non suicidal self-injury, *ADHD* attention-deficit/hyperactivity disorder, *BD* bipolar disorder, *RCI* residential care institution, *ASD* autism spectrum disorder, *SI* suicidal ideation, *ODD* oppositional defiance disorder, *CD* conduct disorder, *RCT* randomized controlled trial, *BA* behavioral activation, *IAU* instruction as usual, *TAU* treatment as usual

[a]Ford-Paz et al. (2019) reported their participants being in grades 7 through 12

[b]Betancourt et al. (2014) did not describe their control condition

[c]Hafeman et al. (2020) compared at-risk youth to a healthy control group that did not participate in the intervention no complete the ER measure at post, as such, only the at-risk group is included here

[d]Wolff et al. (2020) compared two methods of intervention implementation (decision rule vs. sequential treatment) for the same intervention and did not have an intervention comparison group

Subgroup analyses were also computed to compare adaptive skill use or positive and dysregulation or negative measures of ER. Given reporting pooled effect sizes from pre-post within only one group can yield bias (Cuijpers et al., 2016) a second pooled effect was computed for studies with both an intervention and control group using a pre to post change score. Twenty-one studies with both an intervention and control group were included and again subgroup analyses were completed. To address issues of non-independence of effects given that several studies reported more than one measure of ER, a meta-regression was completed for both pooled effect analyses to determine if the individual studies were accounting for heterogeneity in effect sizes. Lastly, sample characteristics (age, sex, and sample composition) were evaluated as contributors to the heterogeneity among effect sizes. Analyses were run in R version 3.6.2 using the meta package (version 4.11).

ER Interventions for Adolescents

Of the 41 included studies, 17 were randomized controlled trials (RCTs) and 24 were pilot studies (including four effectiveness studies without randomization). Target populations ranged from community samples to prevention programs for at-risk adolescents to clinical populations (e.g., adolescents with borderline personality disorder traits, comorbid conduct problems and depression). The interventions and an overview of their effects on ER are described next, organized by population type to ease organization in description of the interventions and their findings.

Community and Prevention Programs

Thirteen community and prevention programs were identified. Four of nine studies investigated the effects of an ER intervention within school settings broadly (Alampay et al., 2020; Ford-Paz et al., 2012; Johnstone et al., 2020; Schoeps et al., 2018) with five of those eight studies investigating the effects of the specific prevention program, Learning to BREATHE within the school setting (Broderick & Metz, 2009; Fung et al., 2016; Lam & Seiden, 2020; Lau et al., 2020; Metz et al., 2013), one study used a mix of at-risk youth and typical youth (Dingle et al., 2016), two studies used at-risk school samples (Claro et al., 2015; Houck et al., 2016), and one study aimed to reduce dating violence for early adolescent males (Rizzo et al., 2020).

The Schoeps et al. (2018) RCT took a preventative approach by providing an emotion education program to 148 seventh and eighth graders (ages 12–15), to help develop the adolescents' emotional skills, to improve interpersonal relationships, reduce conflicts between peers, and improve coexistence and well-being. The emotion education program took place in 11 sessions administered over 3 months during the school day. Schools were randomly assigned to the intervention ($n=72$) or control ($n=76$) group. The control group did not participate in the intervention and had access to typical resources provided at the schools. There were no significant

Table 2 Emotion regulation measures from included articles

Measure Name	Reporter	Content	Evidence of validity and reliability	Study
ADS	Self	Six items; self-reported measure of frequency of difficulties with affect regulation	1—minimal evidence of reliability or validity	Hadley et al. (2020)
Affective instability	Clinician	Subscale from a structured interview assessing the 9 criteria for borderline personality disorder	3—strong evidence of reliability or validity	Schuppert et al. (2009), Schuppert et al. (2012)
ASRI	Self; Parent	Twenty seven items; short- and long-term ER abilities	1—minimal evidence of reliability or validity	Rizzo et al. (2020)
CEMS	Self	Twelve Likert items; sadness, anger, and worry regulation	2—moderate evidence of reliability or validity	Suveg et al. (2009)
CERQ	Self	Thirty six Likert items; cognitive ER strategies	3—strong evidence of reliability or validity	Afshari et al. (2014), Claro et al. (2015), Riggs Romaine et al. (2018), Ramdhonee-Dowlot et al. (2021)
CISS	Self; Parent	Forty eight Likert items; emotion coping	1—minimal evidence of reliability or validity	Keiley et al. (2007)
DERS	Self; Parent	Thirty six Likert items[a,b] or 16 and 18-item short versions; capacity for emotion regulation	3—strong evidence of reliability or validity	Alampay et al. (2020), Betancourt et al. (2014), Bjureberg et al. (2017), Bjureberg et al. (2018), Breaux and Langberg (2020), Broderick and Metz (2009), Cotter et al. (2020), Ford-Paz et al. (2019), Hadley et al. (2020), Houck et al. (2016), Lindqvist et al. (2020), Metz et al. (2013), Newnham et al. (2015), Lam and Seidon (2020), Wolff et al. (2020), Yeo et al. (2020)
EDI	Parent	Thirty items; ER associated with autism	2—moderate evidence of reliability or validity	Conner et al. (2019)
EESC	Self	Sixteen Likert items; deficient emotional expression	2—moderate evidence of reliability or validity	Suveg et al. (2009)
ERQ	Self	Ten Likert items[c]; cognitive reappraisal and expressive suppression	3—strong evidence of reliability or validity	Dingle et al. (2016), Fung et al. (2016), Hafeman et al. (2020), Johnstone et al. (2020), Lau et al. (2020), Santomauro et al. (2016), Wijana et al. (2018)
ERBS	Self	Eight Likert items; assesses ER strategies taught in intervention	1—minimal evidence of reliability or validity	Hadley et al. (2020), Houck et al. (2016), Rizzo et al. (2020)
ERC	Parent	Twenty four Likert items; Parental perceptions of child's management of emotions and lability	2—moderate evidence of reliability or validity	Cotton et al., (2016, 2020), Hadley et al. (2020), Keiley et al. (2015), Thomson et al. (2015), Wolff and Ollendick (2012)
ER Questionnaire for Children	Self	Twenty two vignettes; Level of regulation of fear, anger, and sadness	1—minimal evidence of reliability or validity	Punamäki et al. (2014)
ESCQ	Self	Forty five Likert items; emotional perception, understanding, labeling, expressing, and regulation	2—moderate evidence of reliability or validity	Schoeps et al. (2018)
MERLC	Self	Six items; internal locus of control related to emotions subscale	1—minimal evidence of reliability or validity	Schuppert et al. (2009)

Table 2 (continued)

Measure Name	Reporter	Content	Evidence of validity and reliability	Study
NMR	Self	Thirty Likert items; ability to identify, manage, and utilize adaptively a variety of negative emotion states	3—strong evidence of reliability or validity	Ford et al. (2012)
Weekly Ratings	Clinician; Parent	Weekly ratings of adolescent ER; 3 items on 7-point Likert scale	1—minimal evidence of reliability or validity	Breaux and Langberg (2020)

Evidence of validity and reliability refers to both validity/reliability studies conducted outside of those included in the present review as well as reliability information (e.g., internal consistency) reported in the included studies, when available

ADS affect dysregulation scale, *ASRI* adolescent self-regulatory inventory, *CEMS* Children's Emotional Management Scales, *CERQ* Cognitive Emotion Regulation Questionnaire, *CISS* Coping Inventory for Stressful Situations, *DERS* Difficulties with Emotion Regulation Scale, *EDI* Emotion Dysregulation Inventory, *ERC* Emotion Regulation Checklist, *NMR* Negative Mood Regulation, *MERLC* Multidimensional Emotion Regulation Locus of Control, *ERBS* Emotion Regulation Behaviors Scale, *ERQ* Emotion Regulation Questionnaire, *ESCQ* Emotional Skills and Competencies Questionnaire, *EESC* Emotion Expressivity Scale for Children, *CEMS* Children's Emotion Management Scales

[a]The Betancourt et al. (2014) study reported using only 23 items from the DERS and did not report which items or scales were omitted, no validity or reliability information for the reduced item set was presented

[b]The Alampay et al. (2020) study reported using only four subscales (26 items) of the DERS and reported acceptable internal consistencies

[c]The Lau et al. (2020) used only the 4-item emotion suppression scale of the ERQ and reported acceptable internal consistencies

differences at post-intervention for perception, expression, or management of emotions (Hedge's $g = -0.12, -0.07$, and -0.03, respectively) relative to the control group. At the 6-month follow-up, participants in the emotion education program had significantly better emotional perception (Hedge's $g = 0.46$) and management of emotions (Hedge's $g = 0.51$) relative to the control group with no significant differences in expression (Hedge's $g = 0.09$).

Next, Ford-Paz et al. (2012) examined the feasibility and preliminary outcomes of the Working On Womanhood (WOW) community-developed intervention for ethnic minority girls across multiple schools. WOW aims to serve the social-emotional needs to minority girls with specific modules focused on building emotional intelligence. Skills include learning to label and identify emotions and use coping skills to manage stress. The WOW program is delivered within the school setting using a group intervention format. The current study included 960 girls in grades seven through twelve (age not reported) and did not include a control condition. Results showed a significant decrease in emotion dysregulation at post-intervention, with a small effect size (Hedge's $g = 0.30$).

Several studies investigated the effects of mindfulness or mindfulness-based cognitive therapy (MBCT) programs within the school setting, with three focusing on low-income and marginalized populations. First, Alampay et al. (2020) examined MBCT for Filipino children from "low-resource" schools. Participants were ages 9–16, with 87 participants assigned to the intervention group and 99 assigned to an active control. The intervention was in line with standard MBCT and had eight modules including modules related to ER such as increasing awareness and regulating attention, affect, and behavior. The control condition included similar amounts of activity, levels of attention from adults, and peer interaction, with time spent making crafts. Results showed significant improvements in emotion awareness in the intervention group relative to the comparison (Hedge's $g = 0.41$) with no significant differences in non-acceptance of emotions, impulse control, or regulation (Hedge's $g = 0.27, -0.14$, and -0.11, respectively).

Next, Broderick and Metz (2009) investigated the Learning to BREATHE (L2B) curriculum implemented as part of a typical health class in a private high school. L2B consists of 12 sessions organized around six themes from the BREATHE acronym: Body, Reflection, Emotion, Attention, Tenderness, Habits, and Empowerment, with six modules focusing on topics such as awareness, thoughts, and feelings. The L2B program was developed to use mindfulness to help improve ER skills in adolescents within a group setting. Participants ($n = 120$) in the intervention group were females in their senior year ages 17 to 19 and those in the control group were juniors ($n = 17$; ages 16–17). The control group did not participate in the intervention. The

authors reported significant changes in levels of negative affect (Cohen's $d = 0.57$; means and standard deviations not reported) in the intervention group relative to the control. However, there were no significant changes in emotion dysregulation (Cohen's $d = 0.28$; Hedge's g could not be calculated as means and standard deviations were not reported and attempts to gain data from authors were unsuccessful). Fung et al. (2016) also examined the L2B intervention in a pilot intervention study of ethnic minority youth. Participants were randomized to either the intervention group ($n = 9$) or a waitlist control group ($n = 10$). Relative to the those in the waitlist condition, there were improvements in emotion suppression (Hedge's $g = 0.57$) and decreases in the use of cognitive reappraisal strategies (Hedge's $g = -0.98$). Using the full sample after all participants had received treatment, there were significant decreases in emotion suppression (Hedge's $g = 0.68$) from pre to post and no significant changes in the use of reappraisal (Hedge's $g = -0.26$).

Similarly, Lam and Seidon (2020) examined L2B in seventh grade students from Hong Kong public schools in the lowest academic tier. Participants ($M_{age} = 12.4$) were randomized to either the L2B intervention ($n = 53$) or an instruction as usual control ($n = 62$). There were no significant changes in ER from pre to post relative to the comparison group (Hedge's $g = -0.09$). Metz et al. (2013) evaluated L2B in a community sample of high schools students. Participants were randomized to either the L2B condition ($n = 129$; $M_{age} = 16.5$) or an instruction as usual condition ($n = 87$; $M_{age} = 16.4$). Results showed significant improvements in ER difficulties, with a small effect size (Hedge's $g = 0.24$). The last study to examine the L2B intervention aimed to evaluate the impact of intervention preference on intervention outcomes for a sample of ethnically diverse youth ($M_{age} = 15.18$; Lau et al., 2020). Participants completed either the L2B intervention ($n = 62$) or an interpersonal therapy adolescent skills training intervention ($n = 49$). Those in the L2B intervention compared to the interpersonal skills intervention showed larger improvements in the use of cognitive reappraisal strategies (Hedge's $g = 0.35$), with no changes in emotion suppression (Hedge's $g = -0.04$) at post.

Johnstone et al. (2020) examined the effects of a preventative ER program within five schools in Australia in an RCT. Participants were children ($M_{age} = 11.04$) assigned to either the ER intervention ($n = 185$), behavioral activation control condition ($n = 85$), or a usual class control condition ($n = 25$). Both the ER intervention and behavioral activation control condition were delivered over eight sessions of 50 min each but differed in their content. The ER intervention was comprised of emotion awareness and understanding, cognitive skills (e.g., challenging unhelpful thoughts), and prevention skills (e.g., promoting adaptive ER), whereas the behavioral activation control condition consisted of identifying avoidance behaviors, understanding worry, and practicing behaviors on a "coping behaviors ladder." There were significant improvements in the use of cognitive reappraisal strategies between the ER condition and the behavioral activation condition at post (Hedge's $g = 0.31$), but no significant differences in emotion suppression (Hedge's $g = -0.01$). The findings were reversed comparing the ER condition to the instruction as usual condition with non-significant differences in reappraisal (Hedge's $g = 0.05$) and significant improvements in emotion suppression (Hedge's $g = 0.49$).

Next, Rizzo et al. (2021) aimed to reduce and prevent dating violence among early adolescent males ($N = 119$; $M_{age} = 13.0$) through the Project STRONG intervention. The goal of this intervention is to teach relationship health knowledge, ER, and communication skills. Participants (and a parent) were randomized to either the intervention condition ($n = 59$) or a waitlist control condition ($n = 60$). The intervention consisted of six online modules completed across four weeks (1–2 modules per week) in the context of a game in which players must work to restore healthy relationships and reduce violence on an alien planet. ER skills taught were adapted from project TRAC (developed in Houck et al., 2016 described on page 18). There were no significant changes from pre to post on any of the ER measures (Hedge's g ranged from -0.07 to 0.07).

Dingle et al. (2016) included a mix of at-risk youth ($n = 41$; 76% male) and typical female students ($n = 216$) to validate the ER intervention Tuned In through an effectiveness trial. Tuned In focuses on improving ER and emotional identification, labelling, and awareness through music. Each session focuses on a different emotion and considers the valence (i.e., positive or negative) and arousal levels (i.e., high or low) associated with that emotion. In the at-risk sample, results showed significant increases in emotional awareness (Hedge's $g = 0.54$), emotion identification (Hedge's $g = 0.54$), and use of healthy emotion management strategies (Hedge's $g = 0.49$) from pre to post-intervention. Similar improvements were also observed in the typically developing female students from pre- to post-intervention in regard to emotional awareness (Hedge's $g = 0.75$), emotion identification (Hedge's $g = 0.39$), and management (Hedge's $g = 0.57$). Additionally, the at-risk youth exhibited a significant decrease in emotional suppression (Hedge's $g = 0.64$), but also a decrease in emotional reappraisal (Hedge's $g = -0.52$).

The first study using an at-risk school sample aimed to teach cognitive ER strategies through a group intervention (Claro et al., 2015). This sample consisted of 28 adolescents ages 12–17 who were at-risk for school failure, who took part in the CERTIFY (Cognitive Emotion Regulation Training Intended for Youth) intervention. Among the participants, 13 adolescents who were not at-risk for school failure and did not take part in the intervention served as a control group. Importantly, their design of selecting a healthier

control group with high adaptive skills at baseline may have inflated the effect sizes when comparing across groups. CERTIFY uses a two-pronged approach in that it seeks to both reduce maladaptive cognitive ER while increasing adaptive cognitive ER strategies in adolescents. Each of the eleven sessions work to teach adolescents about cognitive ER, the possible strategies one can use to regulate their emotions, and work towards increasing the likelihood of using the beneficial cognitive ER strategies. For adolescents in CERTIFY, the use of self-reported adaptive cognitive ER skills significantly increased from pre- to post-intervention (Hedge's $g = 1.07$) relative to the control; notably, however, this effect size may be inflated, given that the control group exhibited a relatively large decline in adaptive cognitive ER skills over the course of the intervention (Hedge's $g = 0.40$ for just the intervention group pre to post). In contrast, maladaptive cognitive ER skills were not successfully reduced post-treatment for those in the intervention group relative to controls (Hedge's $g = -0.61$).

Finally, the Houck et al. (2016) study sought to prevent negative outcomes for adolescents at risk for sexual activity and other health risks. The study consisted of 420 seventh graders (ages 12–14) at risk for negative health outcomes due to emotional/behavioral problems or suspected sexual behavior or substance use. Participants were randomly assigned either to an ER intervention or a health promotion intervention. The ER intervention aimed to improve ER skills, to decrease poor decision making that results in unplanned sex, substance use, or physical aggression. The intervention provided psychoeducation on emotions, how emotions and behaviors relate, and recognizing triggers. Additionally, the intervention taught ER strategies for handling decision making including getting away from the trigger, coping skills to release emotional energy in positive ways, and using cognitive reappraisal. Participants in the ER intervention were better able to identify emotion expressions at 6-month follow-up (but not 12-month follow-up) relative to the health risk intervention group, the overall impact of the intervention on ER skills was not significant at either 6-month follow-up (Hedge's g ranging from −0.06 to 0.08) or 12-month follow-up (Hedge's g ranging from 0.00 to 0.07).

Programs for War-Affected Affected Youth

Three studies investigated the use of an ER intervention on war-affected youth with two RCTs (Betancourt et al., 2014; Punamäki et al., 2014) and one pilot study (Newnham et al., 2015). Punamäki et al. (2014) used the Teaching Recovery Techniques (TRT) intervention which is based in CBT and aims to develop effective coping skills and adaptive ER across two weekly four hour sessions for four weeks. This study included a large sample size ($N = 482$, $M_{age} = 11.29$) of Palestinian adolescents that had been exposed to war. Half the sample was randomized to the intervention condition of group sessions over six weeks, while half was placed in the waitlist control group. Participants in both the TRT intervention and the control group increased in the amount of regulation for feelings of fear, anger, and sadness. However, the TRT group did not significantly outperform the control group on any outcome measure from baseline to post-intervention or 9-month follow-up. Notably, the TRT group's scores on ER did not improve substantially during the intervention or follow-up (Hedge's g ranging from −0.19 to 0.02).

The second RCT investigated the Youth Readiness Intervention (YRI) in a sample of 436 ($M_{age} = 17.9$) adolescents exposed to war-related trauma from Sierra-Leone (Betancourt et al., 2014). The aim of the YRI, developed from CBT and interpersonal therapy, is to teach ER, anger management, and problem-solving skills though 10 group sessions to address emotion dysregulation, risky behaviors, and functional impairments. The sample was randomly assigned to either the YRI condition or a control group which did not receive the intervention. Results showed significant improvements in ER post-intervention for the YRI group relative to the control group (Hedge's $g = 0.45$), but no significant differences between groups by 6-month follow-up were found (Hedge's $g = 0.03$). The pilot study investigating the feasibility of the YRI with war-affected adolescents ($N = 32$) from Sierra-Leone ($M_{age} = 18.2$; Newnham et al., 2015) also found ER to significantly increase from baseline to posttest with a moderate effect size (Hedge's $g = 0.73$). Overall, there appears to be some initial promising results for the use of ER interventions in youth affected by war, though with additional RCTs needed.

Programs for Clinical and Health Populations

Twenty studies examined the effects of an ER intervention within specific clinical and health populations.[2] Specifically, three used samples with primarily externalizing behaviors (Breaux & Langberg, 2020; Wolff & Ollendick, 2012; Wolff et al., 2020), four with internalizing behaviors (Afshari et al., 2014; Cotton et al., 2016; Lindqvist et al., 2020; Suveg et al., 2009), two with those at risk for bipolar disorder (Cotton et al., 2019; Hafeman et al., 2020), three with autism spectrum disorder (ASD; Conner et al., 2019; Santomauro et al., 2016; Thomson et al., 2015), two with borderline personality disorder/traits (Schuppert et al., 2009, 2012), four aimed

[2] Of note, despite the focus on clinical populations, only three studies in the present review (Afshari et al., 2014; Dingle et al., 2017; Schuppert et al., 2012) overlap with those included in the recent Moltrecht et al. (2021) meta-analysis of ER interventions.

at preventing and reducing non-suicidal self-injury (NSSI; Bjureberg et al., 2017, 2018; Wijana et al., 2018; Yeo et al., 2020), and one with youth in residential treatment centers (Ramdhonee-Dowlot et al., 2021). Lastly, two studies examined ER interventions in the context of obesity (Cotter et al., 2020; Hadley et al., 2020).

Externalizing Disorders

Both interventions for adolescents with externalizing behaviors were pilot studies. Breaux and Langberg (2020) examined the RELAX (Regulating Emotions Like An eXpert) intervention in a small ($N = 18$) sample of adolescents ($M_{age} = 13.5$ years) with attention deficit/hyperactivity disorder (ADHD) and their caregivers. RELAX aims to improve ER and decrease family conflict through teaching both the adolescent and the caregiver ER strategies in addition to communication and conflict management skills. RELAX is delivered via 90-min group sessions for 8 weeks with the first 60 min consisting of separate parent and adolescent skills and psychoeducation-based groups and 30 min of consisting of the parents and adolescents working together to problem solve and identify skills to use during the following week. Content includes ER and coping strategies, CBT principles, emotion awareness and identification, and sharing emotions in adaptive ways. Results showed large improvements in ER per both adolescent (Hedge's $g = 0.49$) and parent report (Hedge's $g = 0.68$). Adolescent emotion dysregulation as rated by the intervention clinicians also significantly decreased from pre to post (Hedge's $g = 1.78$).

The second study using an externalizing sample, specifically adolescents with a primary diagnosis of oppositional defiant disorder or conduct disorder, was a small pilot study including five adolescents (ages 11–14; Wolff & Ollendick, 2012). The intervention delivered was a flexibly implemented CBT consisting of 11–16 weekly 60–90-min sessions, that was developed based on the Treatment for Adolescent Depression Study (March et al., 2003), parent management techniques, and collaborative problem-solving techniques (e.g., Kazdin et al., 1992; Greene, 1998). Affect identification and regulation was one of 10 major components of the intervention. A significant reduction in parent-reported ER difficulties was found post-intervention (Hedge's $g = 1.46$) and these reductions were maintained at a 1-month follow-up.

Lastly, Wolff et al. (2020) examined the effects of a transdiagnostic intervention for youth with comorbid conduct problems and depression. Although the aim of the RCT was to evaluate the effects of decision rule, using the transdiagnostic intervention ($n = 19$; $M_{age} = 12.6$), versus sequential treatment, using a combination of CBT and parent management training (PMT; $n = 11$; $M_{age} = 10.5$), both intervention conditions contained ER material and pre- and post-intervention ER measures and will be considered. The transdiagnostic decision rule condition consisted of six core sessions (e.g., emotion coaching, ER, problem solving) plus additional CBT content. The order of the intervention sessions was delivered based upon identified deficits or problem areas, as well as strengths, prior to and throughout the intervention. Participants in the sequential CBT/PMT treatment were randomized to either start with CBT (six sessions) or start with PMT (six sessions) with the remaining sessions being content from either intervention based, decided by the clinician. Participants assigned to the transdiagnostic decision rule-based treatment exhibited significant improvements in ER difficulties (Hedge's $g = 0.89$) and those in the sequential CBT/PMT condition did not show significant changes in ER difficulties at post (Hedge's $g = 0.11$).

Internalizing Disorders

Four studies used samples of youth with primarily internalizing disorders. First, Afshari et al. (2014) completed a small RCT evaluating emotion-focused CBT compared to CBT and a no treatment control with 10 youth in each condition (total $N = 30$; $M_{age} = 10.57$) diagnosed with separation anxiety disorder. While broad CBT skills were used in both CBT interventions, the CBT only condition focused on managing anxiety only whereas the emotion-focused condition included greater use of emotion identification and regulation skills as well as focused on emotions outside of anxiety. Participants took part in 10 1-h weekly session in the CBT condition and the same in the emotion-focused condition, with two additional parent sessions. Results showed significant improvements at post-intervention in both dysregulation and adaptive skill use (Hedge's $g = 0.54$ and 0.56, respectively) for the emotion-focused CBT condition relative to the CBT condition and between the emotion-focused CBT condition and no treatment control condition (Hedge's $g = 0.90$ and 1.04 for dysregulation and adaptive ER, respectively). Cotton et al. (2016) completed a pilot study of a mindfulness-based CBT with 10 adolescents ($M_{age} = 13.2$) from an anxiety clinic. This intervention was a 12-week group intervention during which participants learned to identify their thoughts, emotions, and body sensations. Parents rated their child's ER at pre and post. Results showed an improvement in adolescent ER post intervention, with specific decreases in emotional lability (Hedge's $g = 0.52$) and increases in the use of regulation strategies (Hedge's $g = 0.60$).

The next study (Suveg et al., 2009) also utilized young adolescents from an anxiety clinic ($N = 37$, $M_{age} = 10.47$) and implemented a small pilot study of CBT for anxious youth. Participants received either individual or family CBT, depending on the location of services, for 16 one-hour sessions delivered on a weekly basis. Results showed significant improvement in emotional awareness and identification at

post (Hedge's $g = 0.71$). There were also some improvements of ER of worry and sadness: emotional suppression decreased for both worry (Hedge's $g = 0.63$) and sadness (Hedge's $g = 0.58$), and worry dysregulation and use of emotional coping both improved (Hedge's $g = 0.81$ and 0.58, respectively). There were no significant improvements for ER of anger (Hedge's g ranging from 0.16 to 0.24).

The last study using a sample with internalizing disorders investigated the effects of an affect-focused internet-based psychodynamic therapy for youth with major depression (Lindqvist et al., 2020). Participants had to meet criteria for major depression and were randomized to the intervention condition ($n = 34$) or to a supportive control condition ($n = 38$). The intervention consisted of eight modules based in the psychodynamic orientation and including ER skills (e.g., mindfulness skills, evaluating long-term and short-term consequences, avoidance and acceptance, communication) with the goal of decreasing emotional avoidance and increasing openness to and acceptance of emotions. The intervention and the supportive control condition were delivered online with weekly contact from their assigned therapist and tracking of symptoms and well-being. There were significant improvements in ER difficulties in the intervention condition relative to the comparison condition (Hedge's $g = 0.91$).

Bipolar Disorder

Two additional studies examined the effects of MBCT for youth at risk for bipolar disorder (e.g., one parent with a bipolar disorder diagnosis). In Cotton et al. (2020) participants were those diagnosed with anxiety and at increased risk for bipolar disorder. Participants were randomized to MBCT ($n = 24$; $M_{age} = 13.6$) or a waitlist control period ($n = 19$; $M_{age} = 13.8$). Results post intervention showed non-significant changes in ER (Hedge's $g = -0.14$ for lability and -0.21 for regulation) relative to the waitlist control. The second study, Hafeman et al. (2020) used an intervention based in MBCT and mindfulness-based stress reduction to examine changes in resting state functional connectivity in youth at risk for bipolar disorder. Youth took part in an 8-week MBCT intervention and completed measures pre- and post- intervention. Results showed significant improvements in the use of cognitive reappraisal strategies (Hedge's $g = 0.46$) and no change in emotion suppression (Hedge's $g = 0.14$) at post.

Autism Spectrum Disorder (ASD)

Two pilot studies and one RCT investigated how ER interventions work for adolescents diagnosed with ASD. The first study included 20 participants diagnosed with ASD ($M_{age} = 14.94$) who piloted the Emotional Awareness and Skills Enhancement Program (Conner et al., 2019). This was a 16-week one-on-one therapy aiming to build ER skills through decreasing problem behaviors, increasing adaptive skills, and using mindfulness techniques. It is considered a transdiagnostic approach and can be delivered flexibly in 45–50-min sessions each week. There was no control group. Per parent report, there were significant decreases in emotional reactivity at post-intervention.

The second study included 13 males with ASD ($M_{age} = 10.40$ years) who took part in the Secret Agent Society, Operation Regulation intervention (SAS-OR; Thomson et al., 2015). SAS-OR was developed from the well-studied Jr. Detective Program (Beaumont & Sofronoff, 2008) with a focus on ER. Content of the 10 h-long sessions with the adolescent and parent included identifying stressful situations and triggers, building emotional awareness and emotion recognition, and developing alternative adaptive responses. Results showed significant decreases in emotional lability (Hedge's $g = 0.39$) and decreases in total dysregulation for worry, sadness, and anger (Hedge's $g = 0.55$). There was no significant change in ER broadly (Hedges $g = -0.13$).

The third study using a sample with ASD was specifically investigating the effect of a CBT based intervention with a focus on ER within adolescents with ASD and comorbid depression (Santomauro et al., 2016). Twenty adolescents ($M_{age} = 15.75$) took part in the intervention with 10 being randomized to the intervention condition and ten placed on a waitlist. The intervention included 3–4 participants in a group with 11 one-hour sessions focused on teaching strategies to manage symptoms of depression (e.g., identify feelings and thoughts, expressing emotions in adaptive ways). All youth (regardless of condition) exhibited significant increase in the use of cognitive reappraisal strategies when examined together at post-intervention (Hedge's $g = 0.79$). There were no significant findings for changes in expressive suppression (Hedge's $g = 0.12$).

Borderline Personality Traits

One pilot study and one RCT examined the effectiveness of ER interventions for adolescents with borderline personality disorder traits. In a pilot study, Schuppert et al. (2009) explored the effectiveness of an ER training in addition to treatment as usual for 43 adolescents (ages 14–19) with borderline personality disorder traits compared to only treatment as usual (which may have included medication, therapy, and the use of emergency services). Adolescents were randomly assigned to one of the two conditions. The ER training consisted of 17 weekly 105 min group meetings and two booster sessions at 6- and 12-weeks post-treatment. The ER training sought to provide adaptive coping strategies for affective instability, daily stressors, and psychological vulnerability, and to reduce harmful behaviors (i.e., self-harm,

harm to others), by providing psychoeducation on ER, increasing awareness of behavior and emotion patterns, and ER skills (e.g., taking space, challenge negative or distorted assumptions). This study found no significant improvement post-intervention in affect regulation in the intervention group (Hedge's $g = 0.18$) or relative to the control group (Hedge's $g = 0.03$). ER locus of control, however, significantly improved in the intervention group relative to the control group (Hedge's $g = 1.22$). In 2012, Schuppert et al. examined the same ER training in a larger RCT ($N = 109$, ages 14–19) and again compared the intervention to treatment as usual. The intervention group experienced significant improvements in affect regulation (Hedge's $g = 0.54$); however, the treatment group did not improve significantly more than the control group (Hedges $g = 0.12$).

Non-Suicidal Self-Injury

In an effort to reduce or prevent NSSI, Bjureberg et al. (2017) conducted an uncontrolled open trial with $N = 17$ adolescent females ($M_{age} = 15.31$) who took part in Emotion Regulation Individual Therapy for Adolescents (ERITA), a 12-week intervention including modules such as increasing emotional awareness, using ER strategies, and how to control impulses. Additionally, parents followed along with the intervention sessions via an online program. Self-reported difficulties regulating emotion was reported weekly throughout treatment and had significantly decreased by post-treatment (Hedge's $g = 0.59$) and these improvements were largely maintained by 6-month follow-up (Hedge's $g = 1.34$ from pre to 6-month follow-up). Bjureberg et al. (2018) built upon this work with a second pilot study which used a completely online version of ERITA in which 25 adolescents ($M_{age} = 15.7$) took part. Not only was this version of ERITA further adapted based upon feedback from participants in the first sample, but Bjureberg et al. (2018) again included three and six-month follow-up periods to examine long-term effects of the online version of ERITA. Through the online version, treatment providers would give feedback on progress and homework assignments via the online platform or phone app and would provide phone guidance as necessary. Emotion dysregulation consistently declined from pre-treatment through to six-month follow-up with significant differences from pre-treatment to post-treatment, post-treatment to 3-month follow-up, and pre-treatment to 6-month follow-up (Hedge's $g = 0.62$, 0.38, 1.20, respectively).

Wijana et al. (2018) included a sample of 49 adolescents ($M_{age} = 14.60$) who have self-harmed or are suicidal. The Intensive Contextual Treatment for Self-harm intervention was piloted which focuses on improving effective ER and includes a plan for maintenance once treatment ends. At post treatment, there was a significant increase in cognitive reappraisal (Hedge's $g = 0.46$) and no significant change in expressive suppression (Hedge's $g = 0.19$). From the post-treatment to 6- and 12-month follow-ups, the improvement in cognitive reappraisal was maintained (as indicated by a non-significant Hedge's $g = -0.19$ and -0.01 for 6 and 12 month, respectively). Lastly, Yeo et al. (2020) examined the effects of DBT for ethnic minority adolescents who engage in self-harm and have poor self-regulation. Participants were 51 youth ($M_{age} = 14.73$) who were engaging in outpatient treatment and referred to a DBT outpatient program due to their history of self-harm. Results showed significant effects from pre to post intervention in reducing emotion dysregulation (Hedge's $g = 0.84$).

Residential Treatment Centers

Ramdhonee-Dowlot et al. (2021) evaluated the effects of the Super Skills for Life (SSL) program for youth in residential treatment centers in Mauritius using an RCT. Participants ($N = 100$; $M_{age} = 11.75$) all had been identified as having moderate to high levels of emotional or behavioral problems and were assigned to either the intervention condition or a waitlist control. Those assigned to the SSL intervention ($n = 50$) participated in eight CBT-based sessions using a transdiagnostic approach to treat emotional and behavioral disorders and includes behavioral activation, communication skills, and feedback on their skill use in a group setting. Participants in the SSL condition showed significant decreases in the use of maladaptive cognitive ER strategies, with large effects (Hedge's $g = 1.58$), as well as significant increases in the use of adaptive ER strategies (Hedge's $g = 0.78$) relative to the waitlist group.

Obesity

Cotter et al. (2020) and Hadley et al., (2020) examined the effects of mindfulness-based interventions within individuals with obesity. Cotter et al. (2020) investigated the previously described L2B intervention in a small sample ($N = 11$) of adolescents ($M_{age} = 14.36$). The authors describe the importance of using ER strategies within interventions for these populations given the connections between emotions and eating. There were no significant changes in ER difficulties from pre to post (Hedge's $g = 0.04$). The final study (Hadley et al., 2020) examined the HealthTRAC intervention. HealthTRAC was developed specifically for this population and is a combination of a standard weight management intervention and an ER intervention (the same ER intervention used in Houck et al., 2016). Each session included a weight control component (e.g., physical activity) and an ER component (e.g., recognizing triggers for feelings) for a total of 14 sessions. The sessions were divided into "skill training" for the first half and "skill application" for the second half. Participants randomized to the intervention condition

($n = 19$; $M_{age} = 14.89$) or the control condition (a standard weight management intervention; $n = 19$, $M_{age} = 14.53$). Those assigned to the HealthTRAC condition showed significant improvement on ER strategies specifically taught in the intervention compared to the comparison group (Hedge's $g = 0.84$). Relative to the comparison group, participants also reported significant increases in ER difficulties (Hedge's $g = -0.28$) and no significant changes in affect regulation (Hedge's $g = -0.08$). There were also significant changes in parent-reported ER from pre to post, the effect sizes were small when compared to the comparison group (Hedge's $g = -0.15$ for lability and -0.09 for regulation).

Programs for Incarcerated or Delinquent Adolescents

The final four studies investigated ER interventions in the context of incarcerated (Keiley, 2007; Keiley et al., 2015; Riggs Romaine et al., 2018) or delinquent (Ford et al., 2012) adolescents. First, an effectiveness trial was conducted for the Multiple Family Group Intervention (MFGI) for incarcerated adolescent males ($N = 115$; $M_{age} = 15.70$) who had sexually offended (Keiley et al., 2015). This intervention helps participants manage tense and difficult situations through learning to work through high arousal and respond differently when those types of emotions arise. Parents meet with the adolescent and group leaders to help build positive and supportive relationships. MFGI was originally piloted by Keiley (2007) in a smaller trial of 73 incarcerated adolescents ($M_{age} = 15.6$) which showed small but significant improvements in affect regulation from pre to post intervention (all $ps < 0.05$; effect sizes and means and standard deviations not reported). The larger effectiveness trial indicated that according to adolescent-, mother-, and father-report MFGI significantly decreased maladaptive ER from pre to post-intervention (Hedge's $g = 0.16$, 0.36, 0.04, for adolescent, mother, and father, respectively). Importantly, according to adolescent's own report maladaptive ER remained decreased through the 1-year follow-up period (Hedge's $g = 0.38$). The third study used a sample of incarcerated girls ($N = 57$, $M_{age} = 17.45$) to determine the effectiveness of the Juvenile Justice Anger Management (JJAM) which aims to build problem solving, coping, and ER skills through 16 group sessions, twice weekly, for 90-min (Riggs-Romaine et al., 2012). Participants were randomized to either the JJAM condition (and were allowed to continue their treatment as usual) or a treatment as usual only condition. JJAM did not significantly increase positive cognitive ER (Hedge's $g = 0.11$) nor did it significantly decrease negative cognitive ER (Hedge's $g = 0.43$), although the change in negative cognitive ER was marginally significant.

The final study by Ford et al. (2012) used an adolescent sample of delinquent girls ($N = 59$; $M_{age} = 14.7$) to investigate the effectiveness of the TARGET (Trauma Affect Regulation: Guide for Education and Therapy) intervention. Participants were randomized to either the TARGET condition or an enhanced treatment as usual group which included supportive therapeutic services to help the adolescent identify goals and ways of achieving them. TARGET including 12 weekly individual sessions. Results showed an improvement in ER from pre to post, for the TARGET group relative to the enhanced treatment as usual condition (Hedge's $g = 0.34$).

Results of Pooled Effect Size Analysis

Thirty-nine studies were included in the pooled effect size analysis of pre to post treatment change in ER, given data was unable to be obtained from two studies included in the narrative review (Broderick & Metz, 2009; Keiley, 2007). There were 84 total effects (i.e., measures of ER) included (see Fig. 2) with 52 negative or dysregulation measures and 32 positive or adaptive skill measures. As noted earlier, all effects were scaled so that positive effects sizes indicated either an increase in adaptive skills or a decrease in dysregulation whereas a negative effect indicates a decrease in adaptive skills or an increase in dysregulation such that effective interventions would always yield positive effect sizes, regardless of the measures used in the study.

Effect sizes from pre to post treatment ranged from a Hedge's g of -0.19 (Lam & Seiden, 2020) to 1.07 (Ramdhonee-Dowlot et al., 2020) for changes in emotion dysregulation and -0.52 (Dingle et al., 2016) to 1.46 (Wolf et al., 2012) for changes in adaptive regulation or skill use. The overall pooled effect from pre to post treatment on ER was significant [Hedge's $g = 0.29$, 95% CI (0.22, 0.36), $p < 0.001$] with moderate heterogeneity related to study ($I^2 = 54\%$, $\tau^2 = 0.07$) and a prediction interval of Hedge's g −0.24–0.82. Subgroup analyses showed significant effects regarding reduction in maladaptive or negative ER [Hedge's $g = 0.35$, 95% CI (0.26–0.43), $p < 0.05$, $I^2 = 53\%$, $\tau^2 = 0.06$] and improvement in positive or adaptive ER strategies [Hedge's $g = 0.20$, 95% CI (0.08–0.32), $p < 0.05$, $I^2 = 48\%$, $\tau^2 = 0.08$]. Given the moderate heterogeneity in effects, a meta-regression was conducted to assess the effects of the individual studies. There was a significant effect of study on the heterogeneity of effect sizes, $F(40, 43) = 2.09$, $p = 0.009$, with 31.25% of the heterogeneity in effect sizes accounted for by study.

Twenty-two studies with both an intervention and control group were included in the pooled effect size analysis with 49 total measures of ER (see Fig. 3). There were 27 effects capturing change in dysregulation and 22 effects capturing change in adaptive regulation. Effect sizes of the intervention relative to the control group ranged from a Hedge's g of -0.61 (Claro et al., 2015) to 1.58 (Ramdhonee-Dowlot et al., 2021) for the dysregulation effects and from -0.98

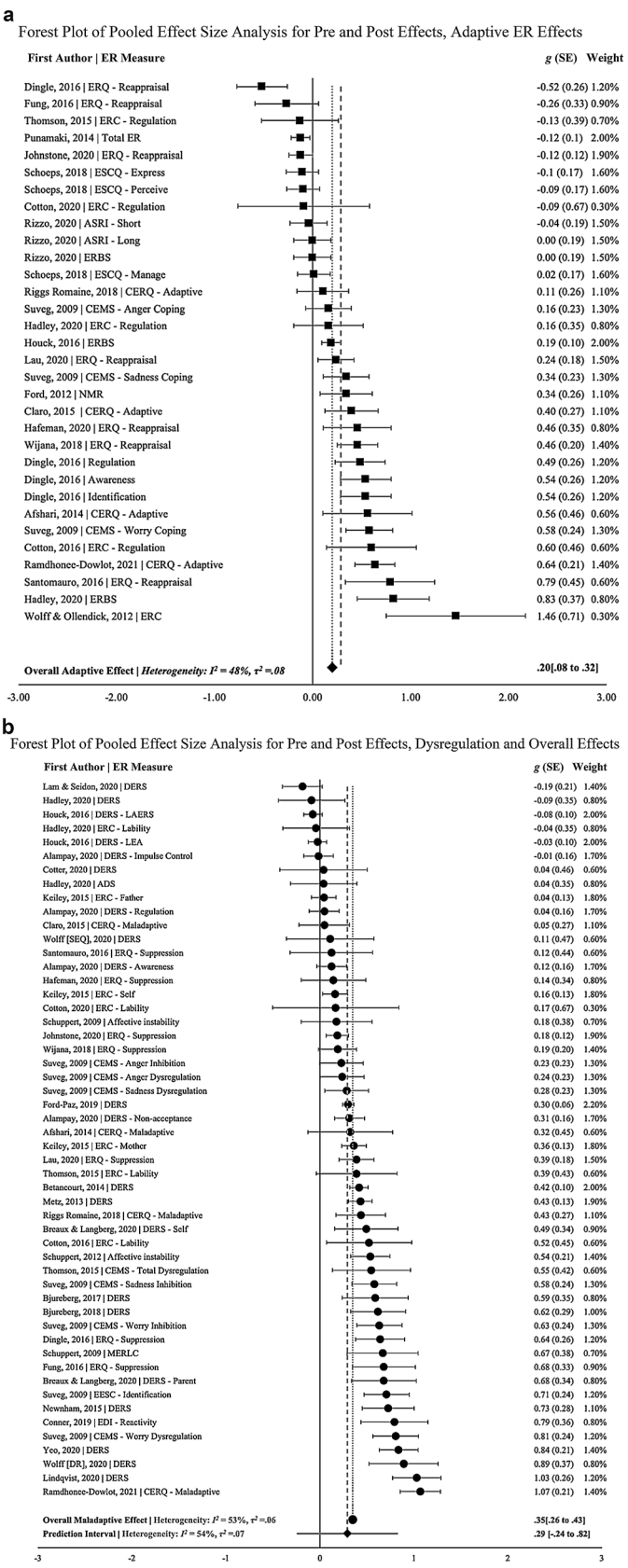


Fig. 2 **a** Forest Plot of Pooled Effect Size Analysis for Pre and Post Effects, Adaptive ER Effects. **b** Forest Plot of Pooled Effect Size Analysis for Pre and Post Effects, Dysregulation and Overall Effects

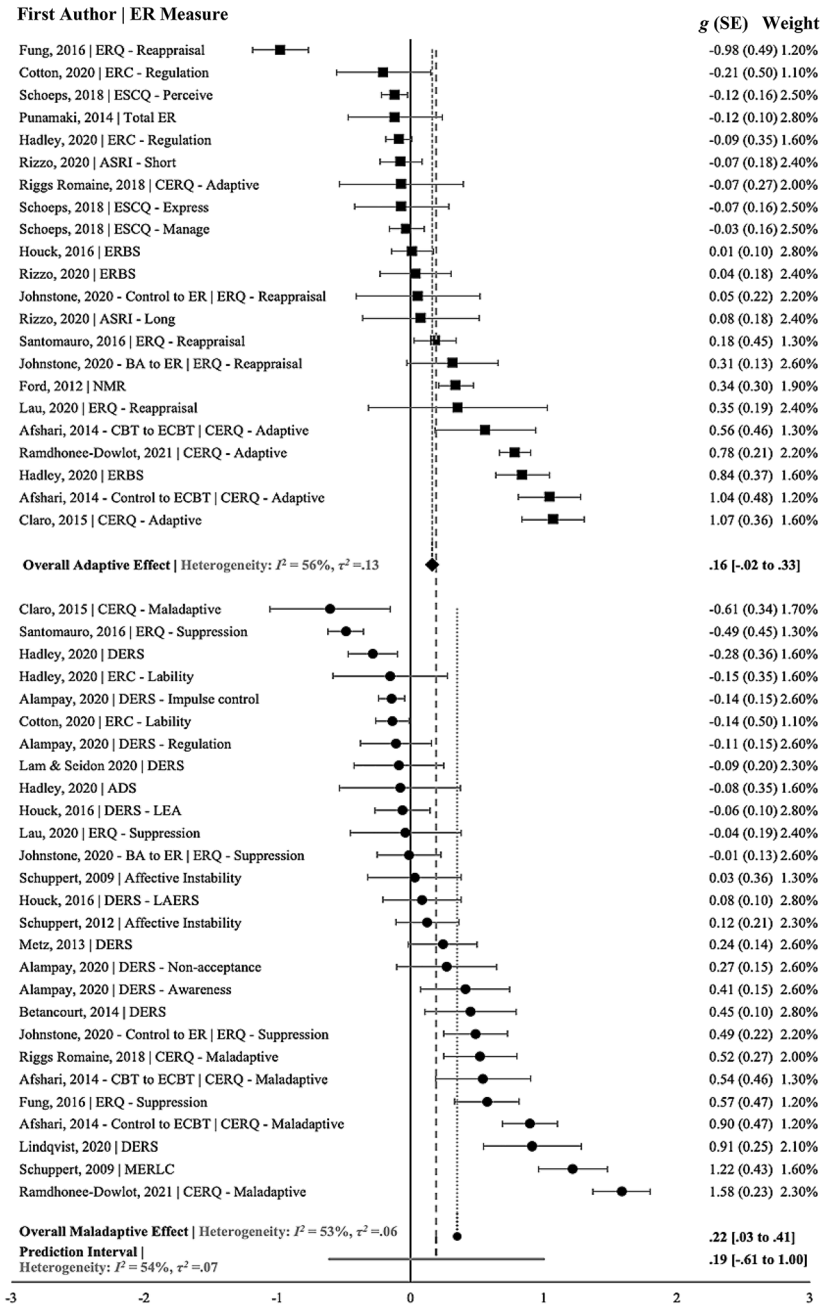


Fig. 3 Forest Plot of Pooled Effect Size Analysis for Studies with an Intervention and Control Group for Both the Dysregulation and Adaptive ER Effects

(Fung et al., 2016) to 1.07 (Claro et al., 2015) for the adaptive effects. The overall pooled effect was significantly different from zero [Hedge's $g = 0.19$, 95% CI (0.06–0.32)] with a moderate to large amount of heterogeneity related to study ($I^2 = 69\%$, $\tau^2 = 0.17$, $p < 0.01$) and a prediction interval of Hedge's $g - 0.61–1.00$. Subgroup analyses for the dysregulation effects showed an overall small effect of ER interventions ability to decrease dysregulation [Hedge's $g = 0.22$, 95% CI (0.03–0.41), $I^2 = 74\%$, $\tau^2 = 0.18$, $p < 0.01$]. By contrast, ER interventions did not show a significant effect relative to control groups in improving adaptive ER skills [Hedge's $g = 0.16$, 95% CI (-0.02–0.33), $I^2 = 56\%$, $\tau^2 = 0.13$, $p < 0.01$]. There was no significant effect of study on the heterogeneity of effect sizes, $F(23, 25) = 1.35$, $p = 0.233$, with 12.41% of the variability being accounted for by study.

Lastly, given the heterogeneity in effects found above, several additional meta-regressions were examined assessing the effects of age, given the wide age range in the present

sample (M_{age} = 10–19), sex, given the ranges of proportions of males to females, and sample types, dichotomized to be community vs. clinical. Meta-regressions were completed on the intervention effects for both the pre-post studies and those with a control and intervention groups. For the pre-post studies, there was no significant effect of age[3] ($F(1, 80) = 1.31$, $p = 0.254$), with 0.05% of the heterogeneity in effect sizes accounted for by age. Similarly, for studies with an intervention and control group, no significant proportion (0.00%) of heterogeneity in effects were accounted for by age [$F(1, 47) < 0.01$, $p = 0.957$]. Next, for the pre-post studies, sex was also not significantly associated with the heterogeneity in effect sizes [$F(1, 81) = 3.85$, $p = 0.053$], with 3.05% of the heterogeneity being accounted for by sex. For those studies with an intervention and control group, results were similar and sex was again not associated with heterogeneity in effect sizes with only 3.36% being accounted for [$F(1, 47) = 2.84$, $p = 0.099$].

By contrast, results indicated that effect sizes significantly differed with sample composition (clinical versus community). Studies were designated as clinical samples if they included sampling of individuals based on a specific diagnosis or degree of symptom severity, whereas studies were designated as community if individuals were not sampled based on youth clinical characteristics. Meta-regressions were significant when examining the effects of sample composition for all included studies with pre and post effects, [$F(1, 82) = 20.73$, $p < 0.001$ using clinical populations as the reference group]. Community samples showed significantly lower effect sizes on average (by 0.29 Hedge's g units, $p < 0.001$) compared to clinical populations. Overall, 18.3% of the heterogeneity in pre and post effect sizes was accounted for by sample composition. Similarly, sample composition also significantly predicted heterogeneity in effect sizes for intervention effects with control groups [$F(1, 47) = 5.68$, $p = 0.021$]. Again, community samples had significantly lower effect sizes, on average (by 0.29 Hedge's g units, $p = 0.021$) than clinical samples. Overall, 8.3% of the heterogeneity in effects in those intervention studies employing a control group was accounted for by sample composition.

Discussion

Across ER intervention programs for adolescents, there are a wide variety of implementation modalities (individual, group, and school-wide), variability in the rigor of research designs (comparison groups, follow-up periods), and different facets of ER taught and/or measured. Overall review of this literature revealed that interventions to improve ER skills, either by increasing adaptive strategies and/or reducing maladaptive strategies, may be helpful for a range of adolescent populations. Our meta-analytic results indicated a significant overall improvement in ER across all studies from pre to post-intervention (Hedge's $g = 0.28$). These effects appeared to be quite variable (ranging from −0.52 to 1.46), and meta-regression analyses indicated study significantly predicted the heterogeneity in the pooled effect size. However, this is consistent with prior meta-analyses of interventions for internalizing and externalizing behaviors in youth (e.g., Battagliese et al., 2015; Ishikawa et al., 2007; Maughan et al., 2005; Michael et al., 2002). Effects were also larger in pilot and non-controlled studies; however, the overall pooled effect size estimate was still significant (albeit small in magnitude) in studies that included a control group (Hedge's $g = 0.16$). Notably, effect sizes capturing reductions in maladaptive ER were larger in controlled studies relative to those capturing improvements in adaptive ER. Neither age nor sex were significant predictors of the variability in effect sizes, whereas sample composition (clinical versus community) clearly impacted effects, with larger effects observed in clinical relative to community samples.

Several factors are likely contributing to the range in effect sizes, which demonstrated moderate to substantial heterogeneity, as indicated by the I^2 values. First, effects observed in non-controlled studies may be overestimating the impact of ER interventions, given that these studies cannot disentangle improvement due to the intervention or a regression to the mean over time. Of the 41 studies reviewed, only 22 studies used a control group and the types of control groups examined varied widely, from wait-list control, treatment as usual, or other related interventions. In some cases, this likely led to an over-estimation of effects such as Claro et al. (2015) using a healthier by design control group (i.e., comparing an at-risk school sample to students not at-risk). Of note, Claro et al. (2015) demonstrated the largest overall effect from pre to post of those studies with intervention and control groups, and it is likely that this effect is inflated due to the study design. In total, 13 of the 22 studies with a control group included an active control condition (e.g., comparison treatment or instruction protocol), which may indicate that interventions with a specific ER component were not resulting in significant improvements over and above other active, evidenced-based interventions or instructional programs. Future work will benefit from using methodologically rigorous comparison groups and working to decompose whether the addition of an ER curriculum incrementally improves the efficacy of an intervention specifically beyond the traditional methods employed.

[3] One study (Ford-Paz et al., 2019) did not report age in their study and was excluded from this analysis.

Second, the types of interventions delivered varied substantially in terms of their focus on ER skills specifically, their duration and dose (e.g., once per week, twice per week, one- or two-hour sessions, total number of sessions), and the populations and contexts in which they were administered (during school, within treatment clinics, group and individual formats). Notably, the largest effect sizes were generally observed in studies measuring a reduction in emotion dysregulation or maladaptive ER strategies among youth with clinical diagnoses. The smallest effect sizes were generally observed for changes in maladaptive ER in studies implementing prevention programs for community youth. Intervention duration and dosage varied substantially among studies producing both small and large effects. Thus, although clinical samples may yield a general pattern of larger effects of ER interventions, it is important to note this may be driven by floor effects (e.g., community populations have less severe ER concerns at baseline and thus less room to improve). Still, future work is needed to determine whether intervention dosage and/or duration impacts improvement both at post-treatment and at follow-up systematically, and whether these effects vary based on diagnoses, age, and or other sample characteristics.

Samples also differed in age and sex composition, as well as which instruments and raters were used to assess ER. Although sex and age were non-significant predictors of study heterogeneity, it will be important moving forward to elucidate for whom and how these interventions are most efficacious. The lack of significant effects of age on the efficacy of ER interventions within adolescent samples may be due to several factors (e.g., insufficient power to detect small moderation effects, age as a poor proxy for developmental effects), including potentially small or absent effects on age and development on intervention efficacy *during* adolescence. It is important to note that the lack of age moderation should not be equated with a lack of age and developmental effects on the development of ER skills in general. Notably, effects of ER interventions are understudied in older adolescent samples (i.e., late high school, 17–19). This may be an important developmental period during which to teach ER skills given the increases in autonomy and additional social developments (e.g., romantic relationships) that occur in late adolescence.

In addition to age and sex, measurement characteristics should be carefully considered when interpreting these findings. There appears to be misalignment in some studies between intervention targets, which often seek to equip adolescents with adaptive coping and ER strategies and measures of ER used. That is, many of the commonly used, validated measures of ER (e.g., the Difficulties with Emotion Regulation Scale, the Emotion Regulation Checklist) assess negativity/lability or emotion dysregulation; in fact, 62% of the measures in the current review were negative or dysregulation measures, with 19 of the 41 studies *only* including a dysregulation or maladaptive measure of ER. Other positive or adaptive measures assess only cognitive ER strategies (e.g., emotion suppression, cognitive reappraisal) rather than broadly assessing engagement in positive coping and ER behaviors (e.g., engaging in a positive activity such as exercise or talking to a friend, taking space, and deep breathes to calm down), that are often a focus of these interventions. Although engagement in adaptive or positive ER behaviors likely is correlated with a reduction in emotion dysregulation, this is not typically examined directly in intervention research, with different ER outcomes being viewed as independent outcomes, often with different effects. As such, it will be critical for future intervention research to better align ER measures with specific intervention targets and to disentangle how improvements in certain emotion awareness/ER strategies learned early in the intervention may lead to improvements in more global dysregulation in daily life over time. In this vein, one of the primary reasons a study was excluded from the present review was due to lack of including any ER measure, despite the intervention having a focus on ER (occurred for 51 of the 96 articles excluded during full text review). Further, only five of the 17 ER measures used across included studies were considered to have strong evidence for reliability and validity. Including valid, reliable, and aligned measures will be integral for ER intervention work moving forward. Additionally, utilization of ecological momentary assessment and/or ecological momentary intervention could be useful to see how specific ER strategies in daily life link to negative or dysregulated emotions and behaviors, and how these associations may change over the course of the intervention.

Third, the adolescent subpopulations of focus across the included studies were also variable. Most of the included research was completed with special populations (e.g., youth with ADHD, at risk or incarcerated youth, youth with borderline personality symptoms) with only 12 of the 41 studies using a community sample. Indeed, subgroup analyses showed community samples, on average, exhibited significantly lower effect sizes compared to all other sample types. It will be important to further understand if teaching ER skills is only beneficial to those with a deficit in ER, or if interventions enhancing these skills is useful among typically developing children as well. It's notable that improvements in ER were documented across a wide range of clinical populations studied including externalizing behaviors (e.g., aggression, oppositionality, ADHD), internalizing behaviors (e.g., depression and anxiety), ASD, borderline personality traits, and war-affected youth. If ER may play a role as a transdiagnostic risk factor or process, these findings might suggest that improving ER via targeted intervention may result in improvement for a variety of youth with different clinical diagnoses. It is likely that teaching the use

of adaptive ER strategies would act as a promotive factor (i.e., help all children perform at higher levels) as opposed to only a protective factor (i.e., only help at-risk children succeed). Yet, some work has demonstrated that students at risk (e.g., Lam & Seiden, 2020) did not benefit from the ER intervention L2B, relative to the comparison group, whereas a typically developing community sample showed significant improvements in ER, relative to the comparison, from the same intervention (Metz et al., 2014). Further work is needed to elucidate for whom and how ER interventions may be most effective (e.g., increasing skills use compared to decreasing dysregulation). Additionally, it would be prudent to explore ER interventions for adolescents with depression only, as no study to date has done so, with the Wolff and Ollendick (2012) study including adolescents with comorbid depression and conduct problems, and Santomauro et al. (2016) including adolescents with ASD and depression.

Fourth, findings also suggested that efficacy and effectiveness of interventions may vary depending on whether the strategies taught focus on increasing adaptive ER skills or decreasing maladaptive ER strategies. Notably, these two dimensions, although related, appear to vary quite independently. It may be that some interventions are more efficacious for increasing adaptive ER (e.g., Claro et al., 2015; Santomauro et al., 2016; Wijana et al., 2018) whereas other intervention types are better suited for reducing maladaptive ER strategies (e.g., Bjureberg et al., 2017, 2018; Dingle et al., 2016). Additionally, many of the studies found significant improvement in related emotional constructs (e.g., emotion awareness and identification; Dingle et al., 2016; Houck et al., 2016) but not in measures of ER specifically. Although emotional awareness, naming, recognition, and identification are essential elements of regulation, it is possible that interventions targeting ER have traditionally been focused on building emotional awareness as a first step toward helping youth make changes in their ER strategies. A range of ER strategies (behavioral, cognitive) as well as different facets of ER (identification, awareness, management) should be assessed, particularly those that directly map onto the constructs being taught in the intervention as noted above.

Finally, few studies included follow-ups beyond the post-intervention assessment. Although promising there were significant improvements at post-treatment, studies that did include follow-up periods reported more mixed findings. Indeed, some work found gains to be maintained beyond the intervention period, ranging from one month to one year (e.g., Bjureberg et al., 2017, 2018; Houck et al., 2016; Schoeps et al., 2018; Wolff & Ollendick, 2012). Thus, it is possible that effects of ER interventions continue to increase over time, and it may be that once ER skills are taught, it takes time for these skills to solidify and be used in daily life. This highlights the importance of future ER intervention research doing both short and long-term follow-up visits.

Conclusion and Future Directions

Overall, results provide promise for the ability of ER-focused interventions to improve the adaptive ER skills of adolescents across a range of populations (i.e., clinical, community, and at-risk samples) with the effect of these interventions being stronger for pilot and effectiveness studies, relative to studies with a control group. Future work examining ER interventions needs to continue to employ rigorous research methodologies (e.g., compare to empirically supported treatments, include additional follow-up periods, use a multi-method assessment of change) and to examine their efficacies across a range of adolescent samples. It also remains crucial to refine ER measures used in interventions to ensure reliable and valid measurement of ER change that maps onto the specific skills being taught within the intervention frameworks. Additionally, moderators of treatment effects (e.g., duration, format delivery), as well as mediators associated with changes in ER skills during the intervention (e.g., length of practice, adolescent engagement) need to be examined to best understand for whom and how these intervention are most effective. Given these findings, additional work is needed to determine whether teaching adaptive ER strategies can serve as a promotive factor in normative samples. Understanding for whom teaching ER strategies are beneficial may have important implications for a range of domains from parenting practices to how teachers structure their school classroom environment. Overall, these findings underscore the importance of researchers continuing to explore the utility of ER interventions for additional clinical and community populations, particularly those that will be generalizable to the wider adolescent population in the United States and abroad.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest.

References

*Indicates a paper reviewed in the present manuscript.

Ahmed, S. P., Bittencourt-Hewitt, A., & Sebastian, C. L. (2015). Neurocognitive bases of emotion regulation development in adolescence. *Developmental Cognitive Neuroscience, 15*, 11–25. https://doi.org/10.1016/j.dcn.2015.07.006

*Alampay, L. P., Tan, L. J. T. G., Tuliao, A. P., Baranek, P., Ofreneo, M. A., Lopez, G. D., Fernandez, K. G., Rockman, P., Villasanta, A., Angangco, T., Freedman, M. L., Cerswell, L., & Guintu, V. (2020). A pilot randomized controlled trial of a mindfulness program for Filipino children. *Mindfulness, 11*(2), 303–316.

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review, 30*(2), 217–237.

Beck, J. S., & Beck, A. T. (2011). *Cognitive behavior therapy: Basics and beyond*. Guilford Publication.

*Betancourt, T. S., McBain, R., Newnham, E. A., Akinsulure-Smith, A. M., Brennan, R. T., Weisz, J. R., & Hansen, N. B. (2014). A behavioral intervention for war-affected youth in Sierra Leone: A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 53*(12), 1288–1297.

*Bjureberg, J., Sahlin, H., Hedman-Lagerlöf, E., Gratz, K. L., Tull, M. T., Jokinen, J., Hellner, Clara, & Ljótsson, B. (2018). Extending research on emotion regulation individual therapy for adolescents (ERITA) with nonsuicidal self-injury disorder: Open pilot trial and mediation analysis of a novel online version. *BMC Psychiatry, 18*(1), 284–296.

*Bjureberg, J., Sahlin, H., Hellner, C., Hedman-Lagerlöf, E., Gratz, K. L., Bjärehed, J., Jokinen, J., Tull, M. T., & Ljótsson, B. (2017). Emotion regulation individual therapy for adolescents with nonsuicidal self-injury disorder: A feasibility study. *BMC Psychiatry, 17*(1), 411.

Blakemore, S.-J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience, 9*(4), 267–277. https://doi.org/10.1038/nrn2353

Blakemore, S.-J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology, 65*, 187–207. https://doi.org/10.1146/annurev-psych-010213-115202

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. Wiley.

*Breaux, R., & Langberg, J. M. (2020). Development and refinement of the RELAX intervention, an intervention targeting emotion dysregulation and interpersonal conflict in adolescents with ADHD: Results from a pilot study. *Evidence-Based Practice in Child & Adolescent Mental Health, 5*, 147–163.

*Broderick, P. C., & Metz, S. (2009). Learning to BREATHE: A pilot trial of a mindfulness curriculum for adolescents. *Advances in School Mental Health Promotion, 2*(1), 35–46.

Bunford, N., Evans, S. W., & Langberg, J. M. (2018). Emotion dysregulation is associated with social impairment among young adolescents with ADHD. *Journal of Attention Disorders, 22*(1), 66–82.

Casey, B. J., & Caudle, K. (2013). The teenage brain: Self control. *Current Directions in Psychological Science, 22*(2), 82–87. https://doi.org/10.1177/0963721413480170

Casey, B. J., Duhoux, S., & Cohen, M. M. (2010). Adolescence: What do transmission, transition, and translation have to do with it? *Neuron, 67*(5), 749–760. https://doi.org/10.1016/j.neuron.2010.08.033

Chervonsky, E., & Hunt, C. (2019). Emotion regulation, mental health, and social wellbeing in a young adolescent sample: A concurrent and longitudinal investigation. *Emotion, 19*(2), 270.

*Claro, A., Boulanger, M. M., & Shaw, S. R. (2015). Targeting vulnerabilities to risky behavior: An intervention for promoting adaptive emotion regulation in adolescents. *Contemporary School Psychology, 19*(4), 330–339.

Compas, B. E., Jaser, S. S., Bettis, A. H., Watson, K. H., Gruhn, M. A., Dunbar, J. P., Williams, E., & Thigpen, J. C. (2017). Coping, emotion regulation, and psychopathology in childhood and adolescence: A meta-analysis and narrative review. *Psychological Bulletin, 143*(9), 939–991.

Compas, B. E., Jaser, S. S., Dunbar, J. P., Watson, K. H., Bettis, A. H., Gruhn, M. A., & Williams, E. K. (2014). Coping and emotion regulation from childhood to early adulthood: Points of convergence and divergence. *Australian Journal of Psychology, 66*(2), 71–81.

*Conner, C. M., White, S. W., Beck, K. B., Golt, J., Smith, I. C., & Mazefsky, C. A. (2019). Improving emotion regulation ability in autism: The emotional awareness and skills enhancement (EASE) program. *Autism, 23*(5), 1273–1287.

*Cotter, E. W., Hornack, S. E., Fotang, J. P., Pettit, E., & Mirza, N. M. (2020). A pilot open-label feasibility trial examining an adjunctive mindfulness intervention for adolescents with obesity. *Pilot and Feasibility Studies, 6*, 1–11.

*Cotton, S., Kraemer, K. M., Sears, R. W., Strawn, J. R., Wasson, R. S., McCune, N., Welge, J., Blom, T. J., Durling, M., & Delbello, M. P. (2020). Mindfulness-based cognitive therapy for children and adolescents with anxiety disorders at-risk for bipolar disorder: A psychoeducation waitlist controlled pilot trial. *Early Intervention in Psychiatry, 14*(2), 211–219.

*Cotton, S., Luberto, C. M., Sears, R. W., Strawn, J. R., Stahl, L., Wasson, R. S., Blom, T. J., & Delbello, M. P. (2016). Mindfulness-based cognitive therapy for youth with anxiety disorders at risk for bipolar disorder: A pilot trial. *Early Intervention in Psychiatry, 10*(5), 426–434.

Cracco, E., Goossens, L., & Braet, C. (2017). Emotion regulation across childhood and adolescence: Evidence for a maladaptive shift in adolescence. *European Child & Adolescent Psychiatry, 26*(8), 909–921.

Cronin, T. J., Pepping, C. A., & O'Donovan, A. (2018). Attachment to friends and psychosocial well-being: The role of emotion regulation. *Clinical Psychologist, 22*(2), 158–167.

*Dingle, G. A., Hodges, J., & Kunde, A. (2016). Tuned In emotion regulation program using music listening: Effectiveness for adolescents in educational settings. *Frontiers in Psychology, 7*, 859.

Ehrenreich-May, J., Kennedy, S. M., Sherman, J. A., Bilek, E. L., Buzzella, B. A., Bennett, S. M., & Barlow, D. H. (2017). *Unified protocols for transdiagnostic treatment of emotional disorders in children and adolescents: Therapist guide*. Oxford University Press.

Ehring, T., Tuschen-Caffier, B., Schnülle, J., Fischer, S., & Gross, J. J. (2010). Emotion regulation and vulnerability to depression: Spontaneous versus instructed use of emotion suppression and reappraisal. *Emotion, 10*(4), 563–572.

Ellard, K. K., Fairholme, C. P., Boisseau, C. L., Farchione, T. J., & Barlow, D. H. (2010). Unified protocol for the transdiagnostic treatment of emotional disorders: Protocol development and initial outcome data. *Cognitive and Behavioral Practice, 17*(1), 88–101.

*Ford, J. D., Steinberg, K. L., Hawke, J., Levine, J., & Zhang, W. (2012). Randomized trial comparison of emotion regulation and relational psychotherapies for PTSD with girls involved in delinquency. *Journal of Clinical Child & Adolescent Psychology, 41*(1), 27–37.

*Ford-Paz, R. E., Crown, L., Lawton, K., Goldenthal, H., Day, G., Coyne, C. A., Gill, T., Harris, N., Blakemore, S., & Cicchetti, C. (2019). Working on womanhood (WOW): A participatory formative evaluation of a community-developed intervention. *Evaluation and Program Planning, 72*, 237–249.

*Fung, J., Guo, S., Jin, J., Bear, L., & Lau, A. (2016). A pilot randomized trial evaluating a school-based mindfulness intervention for ethnic minority youth. *Mindfulness, 7*(4), 819–828.

Graziano, P. A., Reavis, R. D., Keane, S. P., & Calkins, S. D. (2007). The role of emotion regulation in children's early academic success. *Journal of School Psychology, 45*(1), 3–19.

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*(3), 271–299.

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology, 39*(3), 281–291.

Gross, J. J. (2015). The extended process model of emotion regulation: Elaborations, applications, and future directions. *Psychological Inquiry, 26*(1), 130–137.

*Hadley, W., Houck, C., Barker, D., Wickham, B., Bogner, J., & Jelalian, E. (2020). Preliminary impact of an adapted emotion regulation intervention for adolescents with overweight and obesity attempting to lose weight. *Journal of Developmental & Behavioral Pediatrics, 41*(9), 706–715.

*Hafeman, D. M., Ostroff, A. N., Feldman, J., Hickey, M. B., Phillips, M. L., Creswell, D., Birmaher, B., & Goldstein, T. R. (2020). Mindfulness-based intervention to decrease mood lability in at-risk youth: Preliminary evidence for changes in resting state functional connectivity. *Journal of Affective Disorders, 276*, 23–29.

Hannesdottir, D. K., & Ollendick, T. H. (2007). The role of emotion regulation in the treatment of child anxiety disorders. *Clinical Child and Family Psychology Review, 10*(3), 275–293.

*Houck, C. D., Barker, D. H., Hadley, W., Brown, L. K., Lansing, A., Almy, B., & Hancock, E. (2016). The 1-year impact of an emotion regulation intervention on early adolescent health risk behaviors. *Health Psychology, 35*(9), 1036–1045.

*Johnstone, K. M., Middleton, T., Kemps, E., & Chen, J. (2020). A pilot investigation of universal school-based prevention programs for anxiety and depression symptomology in children: A randomized controlled trial. *Journal of Clinical Psychology, 76*(7), 1193–1216.

Kazdin, A. E., Siegel, T. C., & Bass, D. (1992). Cognitive problem-solving skills training and parent management training in the treatment of antisocial behavior in children. *Journal of Consulting and Clinical Psychology, 60*(5), 733–747.

*Keiley, M. K. (2007). Multiple-family group intervention for incarcerated adolescents and their families: A pilot project. *Journal of Marital and Family Therapy, 33*(1), 106–124.

*Keiley, M. K., Zaremba-Morgan, A., Datubo-Brown, C., Pyle, R., & Cox, M. (2015). Multiple-family group intervention for incarcerated male adolescents who sexually offend and their families: Change in maladaptive emotion regulation predicts adaptive change in adolescent behaviors. *Journal of Marital and Family Therapy, 41*(3), 324–339.

Kopp, C. B., & Neufeld, S. J. (2003). Emotional development during infancy. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Series in affective science. Handbook of affective sciences* (pp. 347–374). Oxford University Press.

Kopp, C. B. (1989). Regulation of distress and negative emotions: A developmental view. *Developmental Psychology, 25*(3), 343–354.

*Lam, K., & Seiden, D. (2020). Effects of a brief mindfulness curriculum on self-reported executive functioning and emotion regulation in Hong Kong adolescents. *Mindfulness, 11*(3), 627–642.

*Lau, A. S., Kim, J. J., Nguyen, D. J., Nguyen, H. T., Kodish, T., & Weiss, B. (2020). Effects of preference on outcomes of preventive interventions among ethnically diverse adolescents at-risk of depression. *Journal of Clinical Child & Adolescent Psychology, 49*(6), 820–836.

*Lindqvist, K., Mechler, J., Carlbring, P., Lilliengren, P., Falkenström, F., Andersson, G., Johansson, R., Edbrooke-Childs, J., Dahl, H.-S.J., Bergsten, K. L., Midgley, N., Sandell, R., Thorén, A., Topooco, N., Ulberg, R., & Philips, B. (2020). Affect-focused psychodynamic internet-based therapy for adolescent depression: Randomized controlled trial. *Journal of medical internet research, 22*(3), e18047.

Linehan, M. M. (1993). *Skills training manual for treating borderline personality disorder*. Guilford press.

March, J., Silva, S., Petrycki, S., Curry, J., Wells, K., Fairbank, J., Burns, Barbara, Domino, Marisa, Vitiello, Benedetto, Severe, Joanne, Casat, Charles, Kolker, Jeanette, Feeny, Norah, Findling, Robert, Stull, Sheridan, Baab, Susan, Waslick, Bruce, Sweeney, Michael, Kentgen, Lisa, … Kennard, B. (2003). Treatment for adolescents with depression study (TADS): Rationale, design, and methods. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*(5), 531–542.

Mazefsky, C. A., Herrington, J., Siegel, M., Scarpa, A., Maddox, B. B., Scahill, L., & White, S. W. (2013). The role of emotion regulation in autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 52*(7), 679–688.

Mcclure, K. S., Halpern, J., Wolper, P. A., & Donahue, J. (2009). Emotion regulation and intellectual disability.

McLaughlin, K. A., Hatzenbuehler, M. L., Mennin, D. S., & Nolen-Hoeksema, S. (2011). Emotion dysregulation and adolescent psychopathology: A prospective study. *Behaviour Research and Therapy, 49*(9), 544–554.

*Metz, S. M., Frank, J. L., Reibel, D., Cantrell, T., Sanders, R., & Broderick, P. C. (2013). The effectiveness of the learning to BREATHE program on adolescent emotion regulation. *Research in Human Development, 10*(3), 252–272.

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*, 364–386.

*Newnham, E. A., McBain, R. K., Hann, K., Akinsulure-Smith, A. M., Weisz, J., Lilienthal, G. M., Hansen, N., & Betancourt, T. S. (2015). The youth readiness intervention for war-affected youth. *Journal of Adolescent Health, 56*(6), 606–611.

Ougrin, D., Tranah, T., Stahl, D., Moran, P., & Asarnow, J. R. (2015). Therapeutic interventions for suicide attempts and self-harm in adolescents: Systematic review and meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry, 54*(2), 97–107.

*Punamäki, R. L., Peltonen, K., Diab, M., & Qouta, S. R. (2014). Psychosocial interventions and emotion regulation among war-affected children: Randomized control trial effects. *Traumatology, 20*(4), 241–252.

*Ramdhonee-Dowlot, K., Balloo, K., & Essau, C. A. (2021). Effectiveness of the Super Skills for Life programme in enhancing the emotional wellbeing of children and adolescents in residential care institutions in a low-and middle-income country: A randomised waitlist-controlled trial. *Journal of Affective Disorders, 278*, 327–338.

Riediger, M., & Klipker, K. (2014). Emotion regulation in adolescence. *Handbook of emotion regulation* (2nd ed., pp. 187–202). The Guilford Press.

*Riggs Romaine, C. L., Kemp, K., Giallella, C. L., Goldstein, N. E., Serico, J., & Kelley, S. (2018). Can we hasten development? Effects of treatment on psychosocial maturity. *International Journal of Offender Therapy and Comparative Criminology, 62*(9), 2857–2876.

*Rizzo, C. J., Houck, C., Barker, D., Collibee, C., Hood, E., & Bala, K. (2020). Project STRONG: An online, parent–son intervention for the prevention of dating violence among early adolescent boys. *Prevention Science, 22*, 1–12.

*Santomauro, D., Sheffield, J., & Sofronoff, K. (2016). Depression in adolescents with ASD: A pilot RCT of a group intervention. *Journal of Autism and Developmental Disorders, 46*(2), 572–588.

*Schoeps, K., Villanueva, L., Prado-Gascó, V. J., & Montoya-Castilla, I. (2018). Development of emotional skills in adolescents to prevent cyberbullying and improve subjective well-being. *Frontiers in Psychology, 9*, 2050.

Schriber, R. A., & Guyer, A. E. (2016). Adolescent neurobiological susceptibility to social context. *Developmental Cognitive Neuroscience, 19*, 1–18.

*Schuppert, H. M., Giesen-Bloo, J., van Gemert, T. G., Wiersema, H. M., Minderaa, R. B., Emmelkamp, P. M., & Nauta, M. H. (2009). Effectiveness of an emotion regulation group training for adolescents—A randomized controlled pilot study. *Clinical Psychology*

& Psychotherapy: An International Journal of Theory & Practice, 16(6), 467–478.

*Schuppert, H. M., Timmerman, M. E., Bloo, J., van Gemert, T. G., Wiersema, H. M., Minderaa, R. B., Emmelkamp, P. M. G., & Nauta, M. H. (2012). Emotion regulation training for adolescents with borderline personality disorder traits: A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(12), 1314–1323.

Shapero, B. G., Abramson, L. Y., & Alloy, L. B. (2016). Emotional reactivity and internalizing symptoms: Moderating role of emotion regulation. *Cognitive Therapy and Research, 40*(3), 328–340.

Sloan, E., Hall, K., Moulding, R., Bryce, S., Mildred, H., & Staiger, P. K. (2017). Emotion regulation as a transdiagnostic treatment construct across anxiety, depression, substance, eating and borderline personality disorders: A systematic review. *Clinical Psychology Review, 57*, 141–163.

Smith, E. N., Romero, C., Donovan, B., Herter, R., Paunesku, D., Cohen, G. L., Dweck, C. S., & Gross, J. J. (2018). Emotion theories and adolescent well-being: Results of an online intervention. *Emotion, 18*(6), 781.

Smyth, J. M., & Arigo, D. (2009). Recent evidence supports emotion-regulation interventions for improving health in at-risk and clinical populations. *Current Opinion in Psychiatry, 22*(2), 205–210.

Steinberg, L. (2014). *Age of opportunity: Lessons from the new science of adolescence*. Houghton Mifflin Harcourt.

*Suveg, C., Sood, E., Comer, J. S., & Kendall, P. C. (2009). Changes in emotion regulation following cognitive-behavioral therapy for anxious youth. *Journal of Clinical Child & Adolescent Psychology, 38*(3), 390–401.

*Thomson, K., Riosa, P. B., & Weiss, J. A. (2015). Brief report of preliminary outcomes of an emotion regulation intervention for children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(11), 3487–3495.

Walerius, D. M., Reyes, R. A., Rosen, P. J., & Factor, P. I. (2018). Functional impairment variability in children with ADHD due to emotional impulsivity. *Journal of Attention Disorders, 22*(8), 724–737.

*Wijana, M. B., Enebrink, P., Liljedahl, S. I., & Ghaderi, A. (2018). Preliminary evaluation of an intensive integrated individual and family therapy model for self-harming adolescents. *BMC Psychiatry, 18*(1), 371.

*Wolff, J. C., Garcia, A., Kelly, L. M., Frazier, E. A., Jones, R. N., & Spirito, A. (2020). Feasibility of decision rule-based treatment of comorbid youth: A pilot randomized control trial. *Behaviour research and therapy, 131*, 103625.

*Wolff, J. C., & Ollendick, T. H. (2012). Treatment of comorbid conduct problems and depression in youth: A pilot study. *Child & Family Behavior Therapy, 34*(2), 141–155.

*Yeo, A. J., Germán, M., Wheeler, L. A., Camacho, K., Hirsch, E., & Miller, A. (2020). Self-harm and self-regulation in urban ethnic minority youth: A pilot application of dialectical behavior therapy for adolescents. *Child and Adolescent Mental Health, 25*(3), 127–134.

Zeman, J., Cassano, M., Perry-Parrish, C., & Stegall, S. (2006). Emotion regulation in children and adolescents. *Journal of Developmental & Behavioral Pediatrics, 27*, 155–168.

Clinical Child and Family Psychology Review (2021) 24:579–598
https://doi.org/10.1007/s10567-021-00355-3

A Systematic Review Focusing on Psychotherapeutic Interventions that Impact Parental Psychopathology, Child Psychopathology and Parenting Behavior

Yoel Everett[1] · **Christina Gamache Martin**[1] · **Maureen Zalewski**[1]

Accepted: 25 May 2021 / Published online: 12 July 2021

Abstract
Given the high rates of co-occurrence of psychopathology within families, it is important to identify and characterize interventions that simultaneously reduce both parent and child symptoms, and improve parenting quality. This is needed as intervention development is increasingly moving toward integrated interventions that target some combination of parent and child mental health, and parenting behavior. Even so, much remains unknown regarding which treatment components provide maximum benefit for parent symptoms, child symptoms, and parenting behavior. This systematic review identified and characterized psychotherapeutic interventions that report improvements in each of three outcomes: parent symptoms, child symptoms and parenting behavior. Fifty-six unique interventions were eligible for review, of which 25 reported improvements in all three outcomes. All 25 of these interventions directly intervened on parenting behavior, often as the sole target of the intervention. Few interventions improved all three outcomes in samples in which parents, children or both met clinical-level thresholds of psychopathology. Additional research is needed to better understand the bi-directional and transactional influences of treatment on family members, and to better inform the development of interventions for dually disordered parent-child dyads across a range of diagnostic profiles.

Keywords Parent mental health · Child mental health · Parenting · Intervention · Systematic review

Parent and child psychopathology co-occur at high rates (Middeldorp et al., 2016; Vidair et al., 2011). Parental psychopathology is an established risk factor for the development of child mental health problems, with an estimated 12% of child mental health disorders attributable to parental psychopathology (McLaughlin et al., 2012). While biological risk factors (i.e., genetic and prenatal) are a contributing factor, the high rates of co-occurrence are also partly due to the mechanistic pathways linking parental psychopathology to poorer parenting behaviors, that are in turn associated with negative child outcomes (Goodman & Gotlib, 1999). For example, parenting behaviors have been found to mediate the associations between parental depression, anxiety, ADHD, and borderline personality disorder and children's mental health problems (Breaux et al., 2017; Elgar et al., 2007; Kluczniok et al., 2018.). Furthermore, bidirectional effects point to the negative impact of children's mental health symptoms on parenting (Fite et al., 2006; Patterson, 2002) and on parental mental health (Pardini, 2008), and may further contribute to the high rates of co-occurring psychopathology in parents and children. For instance, children's depressive symptoms and externalizing problems have each been found to longitudinally predict subsequent maternal depression (Gross et al., 2008; Sellers et al., 2016). Despite these well-established interconnections between parent symptoms, child symptoms and poorer parenting, psychotherapeutic intervention studies have not consistently evaluated treatment improvements across all three of these outcomes (Hoagwood et al., 2012). As a result, knowledge regarding which psychotherapeutic interventions are most effective at *simultaneously* improving parent and child symptoms, as well as parenting behavior, is lacking.

The movement towards wanting to improve parent and child mental health and parenting behavior has taken two primary approaches. One approach has been to target adult symptoms *or* child symptoms *or* parenting behavior and

✉ Yoel Everett
yeverett@uoregon.edu

[1] University of Oregon, 1227 University of Oregon, Eugene, OR 97403, USA

then assess for secondary effects in the other outcomes. For example, an adult-focused treatment may target a parent's depression symptoms and assess for downstream or secondary effects on parenting quality or child mental health, even though the intervention was not designed to directly intervene on these outcomes. Much of the knowledge base on how interventions impact parent symptoms, child symptoms and parenting comes from these types of study designs. In the second approach, which incorporates a dual-generation, or integrated intervention framework (Shonkoff & Fisher, 2013), more than one treatment target is directly intervened upon within an integrated intervention. In the same scenario used above with parental depression, this approach moves beyond assessing the child's mental health symptoms or parenting quality to directly providing some level of intervention to the child or directly targeting aspects of parenting. Integrated interventions that target various combinations of either parent symptoms, child symptoms and/or parenting behavior have been developed, often depending on which family member is demonstrating symptoms at the time of the intervention. We briefly review key findings from each of the two aforementioned approaches below.

Interventions Targeting a Single Treatment Outcome

There are multiple examples of interventions that have targeted one member of the parent–child dyad while also assessing and reporting symptoms for the second member of the dyad who did not receive the intervention. For example, treatment of maternal depression, yielded on average, medium effects on parenting quality and young children's mental health (Cuijpers et al., 2015; Gunlicks & Weissman, 2008), and treatment-related improvements in parental panic disorder have been found to result in long-term improvements in children's anxiety and depression symptoms (Schneider et al., 2013). Child and adolescent treatment for anxiety or post-traumatic stress disorder (PTSD) are similarly associated with improvements in parental anxiety (Carrion et al., 2013; Cobham et al., 1998; Escovar et al., 2019) and parental post-traumatic stress (Martin et al., 2019), respectively. In contrast, the literature on parent management training (PMT) programs that primarily target parenting behavior and show robust downstream effects on child outcomes (Eyberg et al., 2008; McCart et al., 2016; Yap et al., 2016), presents mixed results with regard to its effects on parent symptoms (Barlow et al., 2012; Colalillo & Johnston, 2016). Moreover, these programs are less effective at improving children's outcomes in the context of parental psychopathology (Maliken & Katz, 2013; Reyno & McGrath, 2006).

Interventions Targeting Multiple Outcomes

Integrated interventions that directly target treatment outcomes for both the parent and child can be broadly categorized as those aimed at families in which one member of the parent–child dyad is experiencing mental health symptoms, and those aimed at families in which both parent and child are experiencing mental health symptoms. An example of the former are the various prevention programs for families in which a parent has psychopathology, that are aimed at preventing the transgenerational transmission of psychopathology to children (Beardslee et al., 2013; Ginsburg, 2009; Haggerty et al., 2008). Such interventions may integrate treatment components designed to reduce parent symptoms, build children's resiliency, and/or improve parenting behavior (Marston et al., 2016). Previous reviews of integrated prevention programs for families in which a parent has a mental disorder have shown benefits for children's psychopathology and parent–child interactions, with small effect sizes (Loechner et al., 2018; Niccols et al., 2012; Schrank et al., 2015; Siegenthaler et al., 2012; Thanhäuser et al., 2017). A review of child-focused interventions that integrate parent components also found some evidence for improvements in parent symptoms (Acri & Hoagwood, 2015). In the review, one of six interventions that directly targeted parent mental health within treatment reported improvements in parent symptoms above and beyond a comparison group. The other five showed improvements, but did not result in significant group differences.

Finally, some integrated interventions are developed to treat families in which both parents and children are experiencing mental health symptoms. Interventions using this approach have also reported improvements across a number of outcomes. Examples include, but are not limited to, interventions for mothers and children with PTSD (CPP; Lieberman et al., 2005, 2006), mothers and children with ADHD (Multimodal treatment + parent training; Jans et al., 2015), and depressed mothers with children experiencing externalizing problems (Group CBT + parent training; Verduyn et al., 2003) and internalizing disorders (IPT-MOMS; Swartz et al., 2016). Swartz et al. (2016) reported improvements in maternal and child symptoms, but these were not significantly different than in an active control group. The Verduyn et al. (2003) study reported mixed results, with symptom reductions for children which were significantly greater than those in the control group and improvements in maternal depression, which were not significantly different than in the active control.

Prior systematic reviews have addressed the effectiveness of psychotherapeutic interventions assessing more than one treatment outcome across parent and child mental health symptoms and parenting behavior (Acri & Hoagwood, 2015;

Barlow et al., 2012; Colalillo & Johnston, 2016; Cuijpers et al., 2015; Gunlicks & Weissman, 2008; Loechner et al., 2018; Martin et al., 2019; Niccols et al., 2012; Siegenthaler et al., 2012); however, many knowledge gaps remain. Previous reviews have either focused on a specific diagnostic category (e.g., parental depression, Gunlicks & Weissman, 2008; Loechner et al., 2018; maternal substance abuse, Niccols et al., 2012) or on a specific type of intervention (e.g., parent training, Barlow et al., 2012; TF-CBT, Martin et al., 2019). Furthermore, no previous reviews have systematically assessed treatment effects on all three outcomes of interest: parent symptoms, child symptoms and parenting behavior. Rather, they have typically focused on only one (e.g., child-focused intervention effects on parent mental health; Acri & Hoagwood, 2015) or two of these outcomes (e.g., effects of maternal depression treatment on children's mental health and parenting quality; Cuijpers et al., 2015). Families affected by all three may be most in need of effective intervention, and little is known regarding how best to intervene with these families. The mechanistic pathway between parent symptoms, parenting and child symptoms, and the potential bidirectional effects of child symptoms on parenting and parental mental health, make identifying interventions that can improve all three outcomes *simultaneously* all the more important.

An additional knowledge gap exists around how best to integrate treatment components aimed at targeting outcomes across parent symptoms, child symptoms, and parenting behavior, as the methods for integration vary considerably. For instance, while some interventions devote separate sessions to components aimed at parents and children (Clarke et al., 2002; Garber et al., 2009), others include conjoint parent–child sessions to provide therapeutic components to the entire family (Ginsburg, 2009; Haggerty et al., 2008). While readers can generally glean the manner in which treatment developers integrate these diverse components from the methods sections within published articles, previous reviews have not systematically addressed the manner in which they are integrated, or which types of integration may be more effective. Having a better understanding of how to most effectively impact mental health outcomes and improve parenting behavior when treating parent–child dyads, will inform the continued development of integrated interventions to better serve families in need.

The current systematic review aims to fill these gaps by taking a broad approach towards examining any psychotherapeutic intervention, for any diagnostic profile of parent or child, which has reported pre- and post-intervention outcomes for parent mental health symptoms, child mental health symptoms, and parenting behavior. We set out to accomplish two overarching aims, the first of which was to categorize eligible interventions by which and how many of these three outcomes of interest were directly targeted, and which and how many of the three outcomes showed improvements following the intervention. The goal of this aim is to provide initial information on the extent to which an intervention needs to directly intervene upon an outcome in order to improve it. In turn, the second aim focused only on those interventions reporting improvements in all three outcomes. Aim 2 is to identify and characterize those interventions which are able to simultaneously improve parent and child symptoms, as well as parenting behavior. The three goals of this second aim were: (a) to evaluate the study quality and characterize which outcomes were targeted by these interventions, (b) to describe whether integrated interventions targeted multiple outcomes by devoting distinct time to each target (e.g., targeting parent symptoms and providing parenting skills in separate sessions or separately within the same session) or rather by blending components together (e.g., targeting specific parenting behaviors associated with psychopathology), and finally, (c) to characterize the populations for whom these interventions were tested, examining the degree to which the interventions cover a wide range of age groups and diverse populations, clinical diagnostic profiles, and other key environmental stressors.

Method

Inclusion and Exclusion Criteria

Inclusion and exclusion criteria were established using the PICOS approach (Moher et al., 2009) to define the Participants, Interventions, Comparison group, Outcomes and Study design of included papers. This approach is widely used in conducting systematic and meta-analytic reviews (Cuijpers, 2016) and has been found to demonstrate greater sensitivity than other search methods (Methley et al., 2014). Studies on psychotherapeutic interventions for parents and children (ages 1–18) were included if they employed an RCT design in which all three of the following outcomes were reported pre and post-intervention: parental psychopathology symptoms, child psychopathology symptoms and parenting behavior. As part of the explicit goal to examine the need to directly target a specific outcome in order to improve it, interventions were included whether they targeted one or more of these outcomes. As study design and randomization are typically key factors in assessing the quality of individual trials included in systematic reviews, we chose to include only RCTs in an effort to ensure a set of higher quality, more rigorous studies. With regard to measures of parental psychopathology, studies had to include a measure of a DSM defined mental disorder at either the diagnostic or symptom level. To qualify as having measured child psychopathology, studies had to include a measure that assessed a DSM defined mental disorder in children (diagnosis or

symptom level), or a measure of emotional or behavioral difficulties, or for very young children, a measure of attachment style. Insecure attachment patterns are commonly regarded as emerging mental health problems in very young children and meta-analytic reviews have found significant associations between insecure attachment and internalizing symptoms (Groh et al., 2012) and externalizing symptoms (Fearon, 2010). Studies focusing on the postpartum period (children less than 1 year old) were excluded. To qualify as having measured parenting behavior, studies had to include a measure of parenting that assessed how the parent responded or behaved toward their child. Excluded on this basis were studies which measured only parenting-related constructs such as parenting stress, parental self-efficacy, family cohesion, or parent–child interactions which were measured at the dyadic (but not the parent-only) level. Finally, studies were required to be written in English.

Literature Search

The literature search was conducted according to PRISMA guidelines (Moher et al., 2009; see Fig. 1) and titles and abstracts were sourced from the PsycNET and PubMed databases in March 2021. In accordance with recommended practices for conducting literature searches (Sampson et al., 2009), the search strategy was peer-reviewed by a university librarian and used a combination of either Index or MeSH terms for psychotherapy (treatment OR therapy OR intervention OR psychotherapy) and mental health (mental health OR mental disorder OR mental illness) and parent related terms (parent* OR mother OR father OR caregiver OR offspring). We also used a methodology filter to specify study designs (clinical trial OR empirical study OR follow up study OR longitudinal study OR literature review OR systematic review OR meta-analysis OR quantitative study OR treatment outcome). Additional articles were retrieved via backward referencing, where the reference sections of the studies, reviews, and meta-analyses found

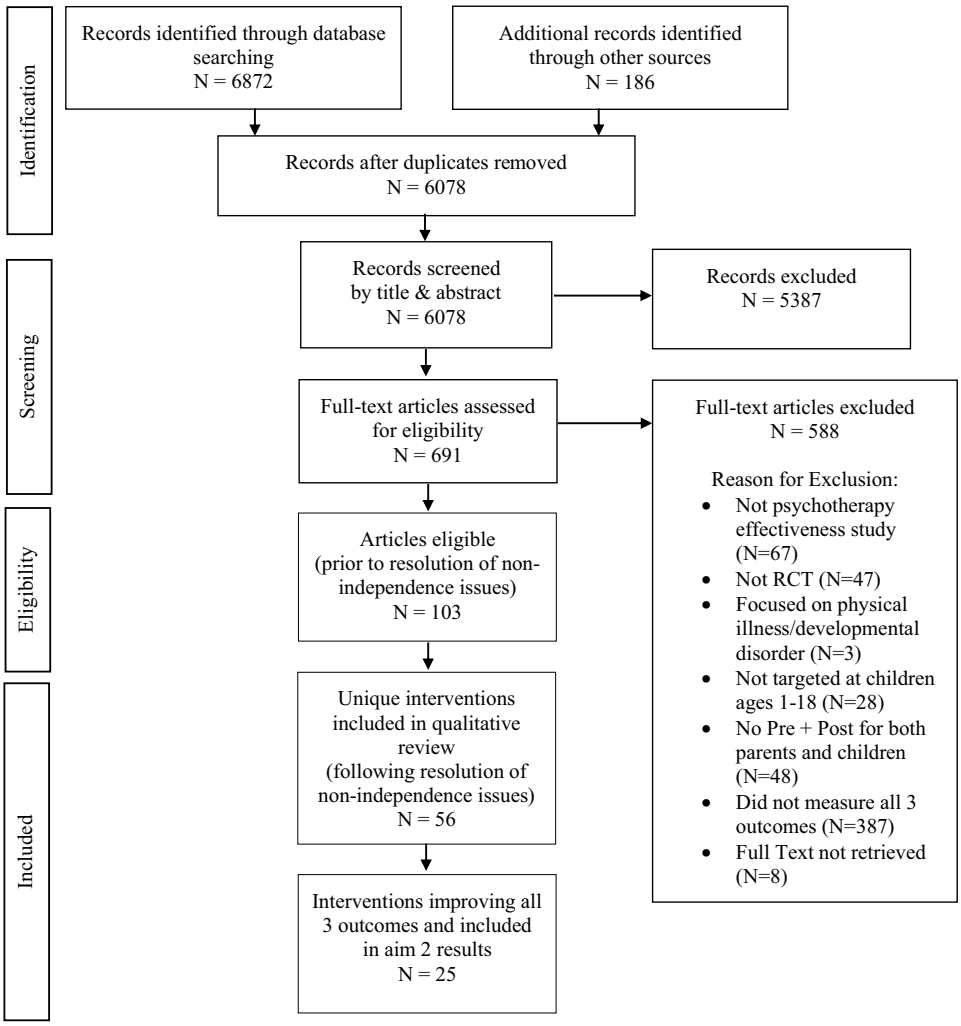


Fig. 1 PRISMA flow diagram

in the literature search were manually reviewed. The online archives of the Society for Research on Child Development (years 2015–2021) and the Association for Behavioral and Cognitive Therapies (years 2015–2020) were also examined for relevant conference papers. Given that multiple papers from a single RCT may be published without individually reporting on all three of the required outcomes, additional backward and forward referencing searches were conducted to retrieve all relevant papers from each unique RCT in order to extract the necessary data. Screening and data extraction were completed using DistillerSR (Evidence Partners, Ottawa, Canada), and the eligibility assessment was conducted by the first author.

Non-Independence Issues

Three types of non-independence issues that influenced the number of interventions reviewed were identified during the search process. A description of each issue and how it was resolved is provided below.

Single RCT Reported Across Multiple Papers

In the case of seven interventions, multiple papers had been published on a single RCT, with no single paper reporting on all three required pre and post outcomes. However, across papers, data on all three were reported, and thus extracted.

Multiple Interventions Tested in a Single RCT

Four RCTs aimed to test multiple interventions within the same study. For instance, Stattin and colleagues (2015) conducted a study that compared four different interventions to a waitlist control, and Sanders et al. (2007) compared Enhanced, Standard and Self-Directed Triple P to each other and a waitlist control. In such cases, data were extracted for each unique intervention.

Interventions with Multiple Independent RCTs

In all, eleven interventions were tested in more than one RCT which met the criteria for this review. These included: Triple P (19 RCTs), Incredible Years (9 RCTs), Parent–Child Interaction Therapy (PCIT; 3 RCTs), Tuning into Kids (TIK; 3 RCTs), the Family Bereavement Program (2 RCTs), Multisystemic Therapy (2 RCTs), the New Forest Parenting Program (2 RCTs), Parent Management Training Oregon Model (PMTO; 2 RCTs), Relational Psychotherapy Mothers' Group (RPMG; 2 RCTs), Trauma-Focused CBT (2 RCTs) and a bibliotherapy self-help behavioral parent training program (2 RCTs). As the aim of this paper was to review unique interventions, a decision was made that each intervention would only be represented once in the review, because including multiple independent trials on the same intervention would misrepresent the descriptive results presented for the systematic review codes. For inclusion in the review and to reduce bias based on results obtained, we made an a priori decision to select the RCT with the largest sample size. Based on this criterion, a particular intervention was only included in the second aim if it reported improvements in all three outcomes (parent symptoms, child symptoms, and parenting behavior).

The decision to select one exemplar study for interventions with multiple RCTs led to some interventions being included in Aim 2, while others were excluded. PCIT, Tuning into Kids, and PMTO were each included in the Aim 2 results. However, RPMG (Luthar et al., 2007), the bibliotherapeutic behavioral parent training program (Hautman et al., 2018), Triple P (Sanders et al., 2000, 2007), and Incredible Years (Leijten et al., 2018) were excluded from the Aim 2 results. The case of Incredible Years (IY) is a noteworthy one. Nine Incredible Years RCTs met criteria for review, the majority of which ($n=5$) did not report improvements in all three outcomes, specifically not finding improvements in parent symptoms. The IY study with the largest sample size ($n=1610$) that served as the exemplar IY study in our review also provided the most robust findings, as it is a meta-analysis which pooled and analyzed *individual participant data* from 14 different IY RCTs (Leijten et al., 2018). Despite the meta-analysis' advantages over other individual RCTs of IY, it is important to acknowledge that the decision to exclude other IY studies that reported improvements in all three outcomes of interest (Bywater et al., 2009; Hutchings et al., 2007, 2017; McGilloway et al., 2012) influences the review's findings. Similarly, in the case of Triple P, the study with the largest sample size ($n=305$) tested three unique versions of the intervention: Enhanced, Standard and Self-Directed, with none of the three versions reporting improvements in all three of our outcomes of interest (Sanders et al., 2000, 2007). However, as in the case of IY, it is important to acknowledge that other RCTs of Triple P that had smaller sample sizes showed improvements in parent symptoms, child symptoms and parenting behavior (Connell et al., 1997; Dittman et al., 2016; Franke et al., 2016; Mejia et al., 2015; Palmer et al., 2019; Sanders et al., 2014; Turner & Sanders, 2006). In the case of RPMG and the bibliotherapeutic behavioral parent training program, the RCT with the larger sample size (RPMG, $n=91$; bibliotherapeutic behavioral parent training, $n=110$) did not report improvements in all three outcomes (specifically, they did not report improvements in parent symptoms) and these interventions (Luthar et al., 2007; Hautman et al., 2018) were therefore not included in results of Aim 2. However, RCTs testing these interventions with smaller sample sizes did report improvements in all three outcomes (RPMG, Luthar & Suchman, 2000; bibliotherapeutic behavioral parent training, Kierfeld et al., 2013).

Across the interventions for which the selection of an exemplar RCT resulted in exclusion from Aim 2, 50% or more of the other RCTs for that intervention would also have resulted in that intervention being excluded from Aim 2. In this way, we verified that the exemplar study findings relevant to our review were aligned with findings from the majority of other RCTs for that intervention. Even so, it is important to acknowledge that for each of the interventions excluded from Aim 2 because of the decision to select a single exemplar study, there was at least one other RCT for that intervention that reported improvements across all three outcomes.

Data Extraction

A data extraction form was developed and pilot-tested on selected representative papers, with iterative adjustments made as needed. The first two authors extracted data for Aims 1 and 2. The data extraction form included codes for targeted outcomes, improved outcomes, the strategy for integrating components targeting multiple outcomes, and the diagnostic profiles of the parent and child samples, and study quality. Two trained undergraduates extracted data on the sample demographics. Demographics of interest included child age groups, the sex of parents and children, families' socioeconomic status (SES), parents' marital status, and ethnic or racial minority status.

Coding Interventions' Targeted Outcomes

For Aim 1, intervention descriptions were carefully read and coded for whether they directly targeted any of the three primary outcomes: parent mental health symptoms, child mental health symptoms, or parenting behavior. To provide some examples, intervention components to decrease parents' cognitive distortions were coded as targeting parent symptoms, intervention components in which therapists were described as working with children on developing their trauma narratives were coded as targeting child symptoms, and components teaching parents about the importance of routines for children were coded as targeting parenting behaviors.

The decision to differentiate between intervention components targeting parenting behavior and those targeting child symptoms is noteworthy, given that a commonly employed approach for improving child symptoms, especially younger children's externalizing behaviors, is to intervene upon their parents' parenting behavior. However, because an explicit aim of this review was to tease apart interventions' targeted outcomes versus the outcomes in which they report improvements, it was necessary to keep these differentiated. Coding rules were established to facilitate this challenging distinction. Specifically, interventions in which a therapist met with the child who was the primary recipient of the intervention, were coded as targeting child symptoms. This was also the case if interventions trained parents to act as a coach or therapist to help their child cope with symptoms. Interventions which focused primarily on changing parenting behaviors, and in which parents were the primary recipients of the intervention, were coded as targeting parenting behavior. Of course, it was possible for interventions to target multiple outcomes, and interventions that included components targeting child symptoms as well as components aimed at improving parenting behaviors were coded as targeting both.

Coding Improvement in Outcomes

For each intervention, data were extracted based on which of the three outcomes showed improvements. As part of this review's goal to identify any and all psychotherapeutic interventions that may simultaneously improve parent symptoms, child symptoms and parenting behavior, broad guidelines were established for coding improvement. Evidence of significant improvement ($p < 0.05$) included the following criteria: studies in which significant improvement was observed on some measures of a given outcome but not on others, studies in which improvement was observed at some time points but not at others, studies in which at least one inferential test showed significant improvement, even if other inferential tests did not (e.g., per-protocol analyses vs intention-to-treat), and studies in which an outcome improved even if not significantly differently from the control group. The rationale for this last criterion, specifically as to why we did not require interventions to outperform control conditions, warrants further explanation. This was done because studies varied in the type of control group employed, with some utilizing waitlist controls, others comparing interventions to treatment as usual, and some utilizing active control groups. The threshold for demonstrating improvement then would have been considerably different and in some cases, would have resulted in concluding that interventions that improved all three outcomes were not included in Aim 2. We therefore opted to code improvement in an outcome even if it was not above and beyond the control group. This decision resulted in four interventions being included in the results of Aim 2 (Cummings & Wittenberg, 2008; Jouriles et al., 2009; Silverman et al., 2009; Slesnick & Zhang, 2016). Interventions that showed improvements in all three outcomes were retained and coded for Aim 2a (distilled from Aim 1), 2b and 2c.

Coding Study Quality

In addition to distilling results from Aim 1 on which outcomes were targeted by interventions reporting improvement in all three outcomes of interest, Aim 2a was to evaluate the quality of the studies in which these interventions were

tested. We used four items from the Cochrane Risk of Bias assessment tool (RoB 2; Sterne et al., 2019) to code study quality, and coded each item as positive (+), negative (−), or unknown (U) if there was no information reported. The RoB 2 tool assesses possible sources of bias in randomized trials, including: (1) the adequate generation of allocation sequence which was assessed as positive when the allocation sequence was random; (2) the concealment of allocation to conditions which was assessed as positive when measures were taken to ensure that enrolling investigators and participants did not have knowledge of the forthcoming allocation; (3) the prevention of knowledge of the allocated intervention (blinding of assessors) which was assessed as positive if outcome assessors were blinded to intervention status; and (4) use of appropriate analyses to estimate the effect of assignment to intervention, which was assessed as positive when intention-to-treat analyses were conducted.

Coding the Method for Integrating Multiple Targeted Outcomes within Treatment

Aim 2b focused on how integrated interventions implemented the targeting of multiple outcomes within treatment. There were three coding options for Aim 2b: (1) devoting separate sessions to each targeted outcome (e.g., the first two sessions devoted to improving parent symptoms and the rest devoted to improving parenting behavior); (2) addressing multiple targeted outcomes within a single session, but as distinct topics (e.g., a single session is divided into time devoted to improving parenting behavior and time devoted to improving parent symptoms); (3) a blended integration across multiple targeted outcomes (e.g., providing parenting skills which are specifically relevant in the context of parental psychopathology). This last category of interventions, which blended the integration of components, was further coded to specify which targeted outcomes were integrated in this manner (e.g., parenting behavior and parent symptoms).

Coding Sample Demographics

Coding for sample demographics was part of Aim 2c. Child age groups were assigned according to the mean age of the children's sample, with age groups divided into four categories: toddlers (ages 1–2), preschool (ages 3–5), school-age (ages 6–11) and adolescents (ages 12–18). A sample was coded as being predominantly low SES if most parents had high school diplomas or less, if lower income families were explicitly oversampled, if a majority of families were on public assistance, or if the average family income was below the national median for the year in which the study was conducted. Relationship status was coded to determine the percentage of participating parents who were not married (single, divorced and/or widowed). Minority status was coded as the percentage of participating parents who were from an ethnic or racial minority group in the country in which the study was conducted. When parents' ethnic and racial background was not reported, the percentage of children from an ethnic or racial minority was extracted.

Coding the Diagnostic Profile of the Samples

The final sample characteristic that was coded was diagnostic profile. The first step was to determine whether or not a given sample met a clinically meaningful threshold. This threshold was met in one of three ways: (1) if a study set its inclusion criteria to require that parents or children meet a clinical threshold for a DSM defined mental disorder; (2) if over 50% of the parent or child sample was reported as meeting a clinical threshold; (3) if the samples' mean baseline scores on measures of psychopathology met an established clinical cutoff. In this latter case, measures for which no clear clinical cutoffs could be determined (e.g., baseline means were not provided or raw scores were provided, but no T-scores) were coded as missing data. For papers in which we were unable to determine if either the parents or children met a clinically meaningful threshold ($n = 8$), we emailed the authors to attempt to obtain this information, before coding it as missing data. Parent or child samples which did meet a clinically meaningful threshold were further coded to categorize which forms of psychopathology met that threshold. This was determined on the basis of the relevant inclusion criteria (e.g., DSM diagnosis of Major Depressive Disorder) or the measure on which baseline levels had met a clinical cutoff (e.g., Beck Depression Inventory score met cutoff for clinical depression).

Reliability

Reliability coding was conducted on 20% of the studies. Percent agreement between coders was 90%, ranging from 80–100%, for all codes except one. Percent agreement for the significant improvement in parenting behavior code was 75%. Due to the importance of this code for our review, the first and second authors both coded all studies for this specific variable and when disagreements arose ($n = 6$), each item was discussed to reach consensus. Study quality was coded by the first author. Consensus coding was utilized for Aim 2b, to characterize the method used to integrate interventions with multiple treatment targets. All three authors coded items for this aim. Again, disagreements were resolved through consensus amongst all three authors. In cases of ambiguity or missing data, authors of papers were contacted to provide additional information to allow for more accurate coding.

Results

Aim 1) Categorize Interventions by Targeted Outcomes and Outcomes Improved

Targeting Parent Symptoms, Child Symptoms and/or Parenting Behavior

In the first aim, we sought to better understand which and how many of the three outcomes of interest—parent symptoms, child symptoms, and parenting behaviors—were directly targeted in the interventions, and which and how many of the three outcomes improved following the intervention (Table 1). A full list of studies included in the qualitative review is provided in Appendix B. Of the 56 interventions, nearly all of them targeted parenting behaviors ($n=53$), approximately half targeted parent symptoms ($n=27$), and approximately one-third targeted child symptoms ($n=17$). For the interventions that targeted a single outcome, only a few interventions exclusively targeted parent ($n=1$) or child ($n=2$) mental health symptoms. In contrast, 23 interventions exclusively targeted parenting behavior. The majority of interventions ($n=30$) directly targeted multiple intervention outcomes. These integrated interventions targeting multiple outcomes were further reduced into three categories: those that targeted parent symptoms and parenting behavior ($n=15$), those targeting child symptoms and parenting behavior ($n=4$), and finally, those targeting all three outcomes ($n=11$). None of the integrated interventions targeted parent and child symptoms without targeting parenting behavior. Having categorized interventions by their targeted outcomes, the next step of Aim 1 is to categorize interventions by the outcomes they improved.

Improvements in Parent Symptoms, Child Symptoms and Parenting Behavior: Single Target Interventions

Of the 26 interventions that targeted a single outcome, 46% ($n=12$) reported improvements in all three outcomes (all 12 interventions targeted parenting behavior only). Thirty five percent of the interventions ($n=9$) that targeted a single outcome reported improvements in both the targeted outcome and one secondary outcome (two of these targeted child symptoms, and seven targeted parenting behavior). Amongst the 26 interventions targeting a single outcome, two (which targeted parent symptoms or parenting behavior) reported improvement in the targeted outcome only. Additionally, three interventions (all targeting parenting behavior) reported improvements in secondary targets (child symptoms and parent symptoms), but not in the primary targeted outcome. Thus, 87% ($n=20$) of interventions that exclusively targeted parenting behavior reported improvement in a secondary outcome or in all three outcomes. Because a total of three studies targeted only parent symptoms or only child symptoms, it is inconclusive whether this single target approach is effective at improving the additional, untargeted outcomes.

Improvements in Parent Symptoms, Child Symptoms and Parenting Behavior: Integrated Interventions

Of the 30 integrated interventions which targeted multiple outcomes, 19 (63%) targeted two outcomes and the remaining 11 (37%) targeted all three. Amongst the 19 dual-target interventions, six interventions (32%) improved all three outcomes (five targeted parent symptoms and parenting behavior and another targeted child symptoms and parenting behavior). Another four interventions (21%), which all targeted parent symptoms and parenting behavior, showed improvements in at least one targeted outcome and the secondary untargeted outcome (child symptoms). Additionally, three (16%) dual-target interventions (all targeting parent symptoms and parenting behavior) showed improvement in the secondary outcome (child symptoms), but not in the primary targeted outcomes. Four interventions (21%) reported improvements in one or both of the targeted outcomes, but no improvement in the secondary outcome. Finally, one dual-target intervention (targeting parent symptoms and parenting behavior) showed no improvement across any of the three outcomes. Thus, 68% ($n=13$) of the dual-target interventions, all of which targeted parenting behavior, reported improvement in a secondary outcome that was not directly intervened upon. Nearly two-thirds of the 11 interventions targeting all three outcomes also reported improvements in all three ($n=7$). The remaining four interventions reported improvements in at least two outcomes. Overall, over three-quarters ($n=23$) of the integrated interventions reported improvements in two or all three outcomes assessed. Sixty-six percent of the 15 interventions targeting parent symptoms and parenting behavior improved two or more outcomes, 50% of the four interventions targeting child symptoms and parenting behavior improved two or more outcomes, and 100% of the 11 interventions targeting all three outcomes, reported improvement in two or three outcomes.

Aim 2) Describe Interventions Reporting Improvements across all 3 Treatment Outcomes

A major focus of this review was to describe and characterize interventions that yielded improvements in parent mental health symptoms, child mental health symptoms and parenting behavior. While Aim 1 centered around which outcomes were targeted and improved by the intervention, Aim 2 is narrowed to fully characterize only those interventions that

Table 1 Number of interventions targeting and reporting improvements in parent symptoms, child symptoms and parenting behavior

		Reported improvements							
Targeted outcome	# of Interventions	No improvements	Parent symptoms only	Child symptoms only	Parenting behavior only	Parent symptoms + Parenting behavior	Parent symptoms + Child symptoms	Child symptoms + Parenting behavior	All three outcomes
Parent symptoms only	1	–	1	–	–	–	–	–	–
Child symptoms only	2	–	–	–	–	–	–	2	–
Parenting behavior only	23	–	–	2	1	–	1	7	12
Multiple treatment targets	30								
Parent symptoms + Parenting behavior	15	1	–	3	1	1	2	2	5
Parent symptoms + Child symptoms	0	–	–	–	–	–	–	–	–
Child symptoms + Parenting behavior	4	–	–	1	1	–	–	1	1
All three targets	11	–	–	–	–	1	2	1	7
Total	56	1	1	6	3	2	5	13	25
% Total	~100%	2%	2%	10%	5%	3%	9%	25%	44%

yielded improvements in all three outcomes. Twenty-five of the 56 reviewed interventions reported improvements for all three outcomes. To characterize these 25 interventions, they were categorized by: (a) the quality of the study in which they were tested and the type of outcomes directly targeted within the intervention, as distilled from Aim 1 results, (b) the manner in which interventions targeting multiple outcomes integrated the treatment components, and (c) the populations with which they were implemented. Table 2 provides a detailed characterization of each of the 25 interventions and we review results of each of the sub-aims below.

Aim 2a) Assess Study Quality and Characterize Intervention Targets

The quality of the studies testing an intervention that reported improvements in all three outcomes varied, but was generally high (Table 2). In the case of 24 interventions (96%), studies reported an adequate sequence generation. One study randomized a portion of participating families, but due to time constraints towards the end of the study period, allocated some participants in a nonrandomized fashion (Hutchings et al., 2002). In the case of 14 interventions (56%), studies reported adequate concealment of allocation to conditions, while studies on 10 interventions (40%) did not provide information on whether allocation was concealed, and one study did not adequately conceal allocation. Studies on 14 interventions (56%) reported the use of blind assessors, while eight did not (32%), and three (12%) did not provide information on blinding of assessors. Intention-to-treat analyses were conducted for 24 out of 25 interventions (96%). Study quality for 10 of the interventions (40%) met all four criteria by which study quality was assessed, studies on seven of the interventions (28%) met three out of four criteria, studies on seven interventions (28%) met two criteria, and one study did not provide information on three of the four criteria for assessing study quality.

Information regarding the outcomes targeted by these 25 interventions was distilled from the results of Aim 1 (Table 2). All 25 interventions included parenting behavior as a target. Approximately half of the interventions ($n = 12$) targeted only parenting behavior. The remaining 13 interventions that reported improvements in all three outcomes targeted multiple outcomes. A majority targeted all three outcomes ($n = 7$) whereas the other six interventions each targeted two outcomes, namely parent symptoms and parenting behavior (Cohen et al., 1999, 2002; Chronis-Tuscano et al., 2013; DeGarmo et al., 2004; Jouriles et al., 2009; Slesnick & Zhang, 2016), or child symptoms and parenting behavior (Cohen et al., 2004).

Aim 2b) Characterize the Method for Integrating Components Targeting Multiple Outcomes

The 13 interventions that targeted multiple outcomes and reported improvements in all three were further examined to determine how treatment components were integrated to target multiple outcomes (Table 2). Only one intervention conducted separate sessions for each target. DeGarmo et al.'s (2004) study on a Parent Management Training Oregon Model (PMTO) intervention for recently separated single mothers and their sons devoted separate sessions to targeting parent symptoms and parenting behavior. Two of the 14 sessions focused on maternal emotion regulation, while the remainder of the sessions provided mothers with parent management techniques.

Four interventions targeted multiple outcomes within a single session, but kept these treatment components distinct from one another. For example, in Compas et al.'s (2009, 2010) Family Cognitive-Behavioral Preventive Intervention, sessions were divided such that parents and children met with therapists separately but concurrently, with parents learning parenting skills and children learning cognitive-behavioral skills in a single session. Parental depression was addressed separately in the first three sessions. In Jouriles et al.'s (2009) Project Support, a single session with mothers included both the provision of child management skills, as well as instrumental and emotional support, with time being devoted to each topic, separately. In Tuning into Kids (TIK), all three outcomes were targeted, with time being devoted to each of them separately (Havighurst et al., 2010). A single two-hour session included both the provision of parental emotion coaching skills (targeting parenting behavior), as well as discrete exercises to enhance parents' emotion awareness, regulation and wellbeing (targeting parent symptoms). Parents were then encouraged to use similar concepts to help their children develop greater emotional awareness (targeting child symptoms).

A majority of interventions ($n = 8$) blended the treatment components aimed at their targeted outcomes within sessions, such that multiple components were addressed in a single session, with a focus on how one target may influence another. Chronis-Tuscano et al.'s (2013) Integrated Parent Intervention for depressed mothers of children with ADHD provides an illustrative example of how parenting behavior is targeted within the context of a parent's symptoms of depression. The intervention included sessions devoted to how a mother's mood affects her parenting, recognizing and changing a mother's cognitive distortions with regard to her child, and increasing focus on positive aspects of the child during a session devoted to parental praise.

Table 2 Characteristics of interventions that reported improvements in all three outcomes: parent symptoms, child symptoms and parenting behavior (categorized by intervention targets and children's age group)

Study	Study quality[a]	Intervention	Children's age group	Clinical sample (Y = Yes, N = No)		Targeted			Treatment component integration
				Parent	Child	Parent symptoms	Child symptoms	Parenting behavior	
Interventions targeting parenting behavior									
Niccols, 2009	++++	COPEing with toddler behavior (CWTB)	Toddlers	N	N			x	N/A
Stanger et al., 2011	+U−+	Parent training + Contingency management	Preschool	Y (Externalizing/ Substance Abuse)	N			x	N/A
Luby et al., 2018 and Whalen et al., 2021	++++	Parent–child interaction therapy emotion development (PCIT-ED)	Preschool	N	Y (Depression)			x	N/A
Cummings & Wittenberg, 2008	++++	Supportive expressive therapy – Parent child (SET-PC)	Preschool	N	Y (Externalizing Behaviors)			x	N/A
Ward et al., 2020	++++	Parenting for lifelong health (PLH)	Preschool	Y (Depression)	Y (Externalizing Behaviors)			x	N/A
Hutchings et al., 2002	−−++	Intensive behavioral management treatment	School age	Y (Depression)	Y (Externalizing Behaviors)			x	N/A
Porzig-Drummond et al., 2014	++−+	1-2-3 Magic emotion coaching	School age	N	Y (Externalizing Behaviors)			x	N/A
Stattin et al., 2015	+U−+	Comet	School age	N	–			x	N/A
Stattin et al., 2015	+U−+	Connect	School age	N	–			x	N/A
Stattin et al., 2015	+U−+	Cope	School age	N	–			x	N/A
Irvine et al., 1999	+UU+	Adolescent transition program (ATP)	Adolescents	N	N			x	N/A
Schwenck et al., 2016	+U++	Plan E parent training	Adolescents		Y (Externalizing Behaviors)			x	N/A
Integrated interventions: targeting parent symptoms and parenting									
Cohen et al., 1999, 2002	++++	Watch, Wait, Wonder	Infants/Toddlers	N	Y (Insecure/ Disorganized Attachment)	x		x	Distinct components

Table 2 (continued)

Study	Study quality[a]	Intervention	Children's age group	Clinical sample (Y = Yes, N = No)		Targeted			Treatment component integration
				Parent	Child	Parent symptoms	Child symptoms	Parenting behavior	
DeGarmo et al., 2004	+UUU	Parent management training oregon model (PMTO)	School age	–	N	x		x	Separate sessions
Jouriles et al., 2009	++++	Project support	School age	–	Y (Externalizing Behaviors)	x		x	Distinct components
Chronis-Tuscano et al., 2013	++++	Integrated parenting intervention for ADHD (IPI-A)	School age	Y (Depression)	Y (ADHD)	x[b]		x[b]	Blended components
Slesnick & Zhang, 2016 and Wu & Slesnick, 2020	+UU+	Ecologically-based family therapy (EBFT)	School age	Y (Substance Abuse)	N	x[b]		x[b]	Blended components
Integrated interventions: targeting child symptoms and parenting									
Cohen et al., 2004	+U++	Trauma-focused CBT (TF-CBT)	School age	Y (Depression)	Y (PTSD)		x[b]	x[b]	Blended components
Integrated interventions: targeting parent symptoms, child symptoms and parenting									
Havighurst et al., 2010	++++	Tuning into kids (TIK)	Preschool	N	N	x	x	x	Distinct components
Silverman et al., 2009	+U++	Cognitive behavioral therapy with parent involvement (CBT/P)	School age	N	Y (Anxiety)	x[b]	x[b]	x[b]	Blended components
Compas et al., 2009, 2010	++−+	Family group cognitive behavioral intervention	School age	Y (Depression)	N	x	x	x	Distinct components
Creswell et al., 2015	++++	Child CBT + Mother–Child interaction (CCBT+MCI)	School age	Y (Anxiety)	Y (Anxiety)	x[b]	x	x[b]	Blended components
Ginsburg et al., 2015	++++	Coping and promoting strength (CAPS)	School age	Y (Anxiety)	N	x[b]	x[b]	x[b]	Blended components
Swenson et al., 2010	++−+	Multisystemic therapy for child abuse and neglect (MST-CAN)	Adolescents	–	–	x[b]	x[b]	x[b]	Blended components
Saldana, 2015	++−+	Families actively improving relationships program (FAIR)	–	Y (Substance Use and Depression)	Y (Externalizing Behaviors)	x[b]	x	x[b]	Blended components

[a]In this column, a positive (+), negative (−), or unknown (U) sign is given for four quality criteria, respectively: allocation sequence, concealment of allocation to conditions, blinding of assessors, and intention-to-treat analyses

[b]indicates which targeted outcomes were blended

Silverman et al.'s (2009) Cognitive Behavioral Therapy with active parent involvement (CBT/P) for child anxiety serves as an example of the blended integration of all three targeted outcomes. In CBT/P, parents and children met with the therapist in a dyadic format in which exposure tasks were planned for both the parent and child. The parent and child would then share their exposure experiences with one another. Parenting behavior was targeted within the context of the child's symptoms, as parent management training was specifically aimed at teaching parents to manage child anxiety symptoms and facilitate exposures through the use of contingencies.

Overall, three of the eight interventions that employed a blended integration of targeted outcomes integrated all three targeted outcomes (Ginsburg et al., 2015; Silverman et al., 2009; Swenson et al., 2010), four integrated the targeting of parent symptoms and parenting behavior (Chronis-Tuscano et al., 2013; Creswell et al., 2015; Saldana, 2015; Slesnick & Zhang, 2016), and Cohen et al.'s (2004) TF-CBT integrated the targeting of child symptoms and parenting behavior.

Aim 2c) Characterize the Participants

Sample Demographics Intervention studies which reported parent sex ($n=21$) focused primarily on mothers, who represented 94% of the parent samples. Children's sex was more equally represented in the studies, but slightly skewed towards males (on average, 57% were male, with $n=24$ reporting). Twenty-three studies reported children's age. The majority of interventions focused on school-age ($n=13$) and preschool-age children ($n=5$). Three interventions focused on adolescents and two on toddlers. Twenty-one studies reported on the parents' SES and over half of them ($n=11$) focused on families from low-SES backgrounds. Across 19 studies, an average of 36% (ranging from 11–100%) of participating parents were either not married, single, divorced or widowed. Participants' racial and ethnic background was reported in only 72% ($n=18$) of the studies and an average of 34% (ranging from 12–100%) of participating parents in these interventions were from an underrepresented minority group.

Parent Diagnostic Profiles Overall, clinical thresholds for parent samples were able to be determined for 80% ($n=20$) of the interventions showing improvements in all three outcomes. Parent samples met a clinical threshold in half of these cases ($n=10$): six parent samples met a clinical threshold for depression, two for anxiety, one for mothers with a substance use disorder, and another for externalizing symptoms of mothers who also had a history of substance abuse.

Child Diagnostic Profiles Clinical thresholds for children's symptoms could be determined for 84% of the interventions ($n=21$). Over 60% of these ($n=13$) included child samples that met a clinical threshold. Seven of these had clinical levels of externalizing behaviors. Two studies involved children with clinical levels of anxiety. The other four studies each had child samples characterized by a different diagnostic profile: one included children with clinical levels of ADHD, another with children exhibiting clinical level symptoms of PTSD, one with children with clinical levels of depression, and one intervention included a sample of infants and toddlers with insecure and disorganized attachment.

Parent–Child Dyads Clinical thresholds for *both* parents' and children's symptoms were able to be determined in approximately two-thirds ($n=17$) of the studies. Approximately one third of these ($n=6$) included samples in which both parents and children met a clinical threshold. Three interventions involved depressed mothers and children with externalizing behaviors (Hutchings et al., 2002; Saldana, 2015; Ward et al., 2020). The Saldana (2015) intervention study focused on a high-risk sample of substance using mothers who also met a clinical threshold for depression at baseline. Two additional interventions were tested with a sample of depressed parents. In one case, the intervention was designed for depressed mothers and children with ADHD (Chronis-Tuscano et al., 2013) and in the other instance, children had clinical levels of PTSD (Cohen et al., 2004). Finally, Creswell et al.'s, 2015 study on an integrated intervention combining CBT for children and a mother–child interaction intervention (CCBT + MCI) was tested with a sample of clinically anxious mother–child dyads.

Intervention Targets across Diagnostic Profiles Seventy percent ($n=7$) of the interventions tested with samples of parents who met a clinical threshold were integrated interventions that targeted multiple outcomes. Six of these utilized a blended approach to integrating treatment components aimed at different targeted outcomes. In contrast, 80% ($n=8$) of interventions tested with parent samples who did not meet a clinical threshold targeted the single outcome of parenting behavior. A majority of the interventions focused on clinical child samples integrated multiple targets ($n=7$), most commonly combining all three targets ($n=3$) or targeting parent symptoms and parenting behavior ($n=3$). One intervention for children with clinical levels of PTSD targeted child symptoms and parenting behavior. A majority of these integrated interventions ($n=5$) blended the delivery of components aimed at different targeted outcomes.

Of the six interventions tested with parent–child dyads where both members of the dyad met a clinical threshold, only two interventions targeted a single outcome (parenting behavior, Hutchings et al., 2002; Ward et al., 2020). The

other four interventions included treatment components targeting multiple outcomes (parent symptoms, child symptoms, and parenting behavior, Creswell et al., 2015; Saldana, 2015; or parent symptoms and parenting behavior, Chronis-Tuscano et al., 2013; or child symptoms and parenting behavior, Cohen et al., 2004). Within these four integrated interventions, three blended treatment components targeting parent symptoms and parenting behavior, while Cohen et al.'s (2004) TF-CBT blended treatment components targeting child symptoms and parenting behavior.

Discussion

The high rates of co-occurrence of psychopathology in parents and their children, the mechanistic role that parenting plays in the transmission, and the bidirectional effects of children's mental health symptoms on parenting, necessitates identifying evidence-based interventions that can improve symptoms in parents and their children, as well as improve parenting quality. This systematic review identified intervention trials that reported on pre- and post-assessments of psychopathology symptoms in parents and children and parenting behavior. It then further characterized interventions that reported improvements in all three outcomes. The generally high study quality of the RCTs included provides an additional measure of confidence in the interventions reviewed.

Overall, 56 unique interventions assessed treatment-related changes in parent symptoms, child symptoms and parenting behavior, and 25 of them reported improvements in all three outcomes. There were several main findings. First, results of Aim 1 suggest that a majority of the interventions reviewed were able to effect change in a secondary outcome which was not directly intervened upon. Moreover, intervening on parenting behavior, either alone or in an integrated intervention, may be particularly key to improving the secondary outcomes of parent and child symptoms. Indeed, all of the interventions that reported improvements in all three outcomes targeted parenting behavior. Second, among the interventions that demonstrated improvement in all three outcomes, a majority were integrated and targeted more than a single outcome. These often employed a blended approach to integrating treatment targets. Finally, the last major finding was that while some interventions effectively treated parents and children who met clinical thresholds, the majority of these focused on maternal depression and child externalizing problems. These key findings are discussed below in greater detail. Together, they provide insights into whom might be best served by future treatment development efforts and how best to approach intervention design when seeking to simultaneously improve parent symptoms, child symptoms and parenting behavior.

Targeting Parenting Behavior to Improve Secondary Outcomes

Nearly all of the interventions in this review targeted parenting behavior (95%, $n = 53$). The sizable proportion of interventions that targeted parenting behavior, whether solely or in an integrated manner, likely stems from the parent training literature's long-standing awareness of the downstream effects on child symptoms, the secondary effects on parent symptoms (Hoagwood et al., 2012; Kazdin & Wassell, 2000), and the bidirectional influences among these variables (Pardini, 2008). Indeed, the results from this review corroborate that changing parenting behavior yields downstream effects on children and secondary effects on parents' symptoms, in that 22 of the 23 interventions that exclusively targeted parenting behavior resulted in improvements in at least one of the other two outcomes. This finding highlights the key mechanistic role parenting behavior plays in the transmission and reduction of psychopathology in children (Goodman & Gotlib, 1999), and also prompts greater consideration of how targeting parenting behavior may positively impact parent mental health. Gonzales and Jones (2016) have proposed two mechanisms through which targeting parenting behavior may result in improvements in parental depression. First, parent training increases parental self-efficacy which may alleviate depressive symptoms. Second, parent training which requires parents to engage more with their child, may reflect a form of behavioral activation requiring and reinforcing positive parent–child interactions through improvements in child behavior. Additionally, as child externalizing behavior is linked to parenting stress (Barroso et al., 2018), it is likely that the subsequent improvements in child behavior as a result of targeting parenting have the additional effect of improving parental symptoms by reducing parenting stress. The finding that all 25 interventions that reported improvements in all three outcomes (i.e., parent symptoms, child symptoms and parenting behavior) included treatment components that targeted parenting behavior further underscores the importance of targeting parenting.

In highlighting parenting behavior as a key treatment target, it is worth noting an unavoidable confound in our review's distinction between interventions targeting parenting behavior and those targeting child externalizing disorder symptoms specifically. In the case of child externalizing disorders, parent training programs are typically the recommended treatment, for which parents are the primary recipients of the intervention. Parent training is the gold standard intervention in these cases precisely because of the robust evidence showing that *parenting behavior* is a mechanism of change through which to influence child externalizing symptoms. Thus, this confound likely bolstered the overall finding that targeting parenting led to improvements in additional

outcomes, as parenting interventions for children's externalizing problems are designed explicitly to do that.

Parenting Interventions and Integrated Interventions Reported Improvements in All Three Outcomes

Nearly half of the interventions reviewed (45%; $n = 25$) showed improvements in all three outcomes of interest. However, it would be highly inaccurate to conclude that the remaining interventions (55%; $n = 31$) were ineffective. This is because improving all three outcomes was very seldom the explicit goal of the interventions. There are additional important observations regarding interventions that did not improve all three outcomes compared to those that did. One observation is that none of the interventions that solely targeted parent or child symptoms ($n = 3$) resulted in improvements in all three outcomes. Instead, improvement in all three outcomes was more likely for interventions targeting parenting behavior alone (47% of which improved all three outcomes), or for integrated interventions that targeted some combination of parent symptoms, child symptoms and/or parenting behavior.

Even within integrated interventions, it is important to consider factors that may have differentiated between interventions that led to improvement in all three outcomes, versus those that did not. First and perhaps not surprisingly, interventions that targeted all three outcomes were more successful at improving all three outcomes (~67%) compared to integrated interventions that targeted two outcomes (~30%). Second, targeting parenting behavior alone was more likely to lead to improvement in all three outcomes (47%) compared to targeting parenting behavior plus an additional target (30%). These unexpected findings likely emerge because dual-target interventions were more often conducted with clinical samples compared to interventions that only targeted parenting behavior. This finding aligns with other work showing how difficult it can be to improve all of these outcomes when parents have psychopathology and are seeking treatment for their children (Maliken & Katz, 2013; Reyno & McGrath, 2006; Wergeland et al., 2016).

Types of Integrated Interventions Reporting Improvements in All Three Outcomes

In addition to the prevalence of interventions targeting parenting behavior only, two predominant types of integrated interventions showed improvement in all three outcomes: those targeting parent symptoms and parenting behavior ($n = 5$) and those targeting all three outcomes ($n = 7$). Together, they represented 92% of the integrated interventions that improved all three outcomes. This tentatively suggests that reduction in child symptoms may be insufficient to produce secondary improvements in parent symptoms, such that integrated interventions should strongly consider targeting parent symptoms. Lastly, a brief note is warranted on the one intervention exception, TF-CBT, that produced improvements in all three outcomes, despite not targeting parent symptoms. As discussed in detail in another paper (Martin et al., 2019), within TF-CBT, parents are provided with all coping skills taught to their child, which may contribute to reductions in their own symptoms.

Blended Approaches to Integration of Treatment Targets

A major goal of this review was to characterize interventions that improved all three outcomes for the purpose of informing future treatment development efforts. One aspect of this characterization focused on the design of integrated interventions (Aim 2b). For example, based on the growing recognition that parenting programs are less effective for families in which a parent has psychopathology, a number of interventions (e.g., Incredible Years ADVANCE, Enhanced Triple-P) have added modules or sessions aimed at improving parental psychopathology by increasing parents' self-regulation, social support, and providing stress coping strategies (Sanders et al., 2000; Webster-Stratton, 2014). However, it remains unclear if adding these sessions is sufficient for eliciting change when parents and/or children have clinical level symptoms. In contrast, a blended approach, which contextualizes parenting difficulties relevant to symptomatic parents and children, appears more suitable for a clinical population. Indeed, the present review points to the utilization of a blended approach amongst integrated interventions that reported improvements in all three outcomes of interest, especially for clinical samples. Integration was largely achieved by addressing multiple targets within a single session (either as distinct topics or with a more blended approach), rather than by devoting separate sessions to each targeted outcome.

Clinical Populations Served by Interventions Reporting Improvements in All Three Outcomes

As reviewed above, it is critical to describe *for whom* these interventions are effective, focusing specifically on the diagnostic profiles of parents and children. Overall, we identified 17 interventions that improved all three outcomes in a clinical sample of either parents, children or both parents and children. These were generally skewed towards samples of children with clinical-level externalizing behaviors ($n = 7$) and/or mothers with clinical-level depression ($n = 6$). While the prevalence of these disorders and the associations between them (Gopalan et al., 2011) may justify prioritizing the study of treatments for depressed mothers and children

with externalizing disorders, it is nonetheless concerning that treatments focused on other diagnostic categories were generally lacking. This is concerning because there is ample research to suggest that different disorders are uniquely associated with specific parenting behaviors (Eyden et al., 2016; Lovejoy et al., 2000; van der Bruggen et al., 2008). Recruiting parents with different forms of psychopathology and assessing for intervention-driven improvements in transdiagnostic features of psychopathology is key to making interventions suitable to more families.

Demographic variables were also examined and reveal populations for whom these interventions have largely been untested. Consistent with previous reviews (Magill-Evans et al., 2006; Panter-Brick et al., 2014), our results indicate that mothers are the primary participants in these interventions. Better recruitment and engagement of fathers is especially important given that mothers and fathers differ in their rates of psychopathology and that there are gender differences in prevalence rates of various diagnostic categories. Additionally, as evidence for nonrandom mating in psychiatric populations exists (Nordsletten et al., 2016), it may be more likely that mothers with psychopathology are co-parenting with fathers who also have psychopathology. With regard to children's age groups, nearly three-quarters of the interventions that reported improvements in all three outcomes focused on preschool and school-aged children, which is consistent with the overall focus of many of the interventions on targeting parenting behavior to reduce children's behavior problems. Further, two-thirds of the interventions with clinical samples were with school-aged children, suggesting a particular gap in the field concerning preschoolers and adolescents with clinical level symptoms.

Strengths and Limitations

There are a number of limitations to this systematic review. First, in determining whether interventions showed improvement in each of the three outcomes of interest, this review made use of "vote counting" or categorizing the results of interventions into positive or negative based on statistical significance. Siddaway and colleagues (2019) have pointed out the flaws of this approach which does not take into account sample size and does not provide an estimate of effect size. A meta-analysis would address these issues. However, we opted not to conduct a meta-analysis because the search was designed to identify a broad range of interventions that reported all three outcomes, resulting in a heterogeneous pool of studies. Moderation analyses would likely have resulted in very small cell sizes, rendering analyses significantly underpowered. Instead, this systematic review represents the first step towards identifying and characterizing which interventions are able to improve all three outcomes. A related concern about dichotomizing interventions as having improved all three outcomes or not, was that all factors used to make this determination were given equal weight, despite the fact certain indicators of significance may have been more meaningful than others. In other words, even amongst the interventions reporting improvements in all three outcomes, there may be considerable differences in regard to their overall effectiveness. For all of these reasons, when this topic is ready for a meta-analysis, the quantifiable methods around weighing effect sizes will offer more clarity around which interventions are most impactful.

Finally, in searching for interventions that have reported outcomes for parental psychopathology, child psychopathology and parenting behavior, and excluding studies that did not report on all three outcomes, this systematic review may be vulnerable to publication bias. Statistically significant outcomes are more likely to be reported than non-significant RCT results and this outcome reporting bias has been found to impact the results of systematic reviews (Kirkham et al., 2010). We attempted to mitigate the effect of publication bias on this review by using liberal criteria to identify interventions that improved all three of our outcomes of interest. Nonetheless, it is possible that additional interventions exist which have assessed all three outcomes, but did not report all three, and were thus, excluded from this review.

Despite these limitations, this review is the first to identify and characterize psychotherapeutic interventions that assessed and reported on parental psychopathology, child psychopathology and parenting behavior, permitting a broader perspective on how the field has addressed a potential transactional relationship between these outcomes (Sameroff, 2009). Prior reviews have taken a narrower approach by focusing on how interventions have addressed two of these outcomes, by concentrating only on a single diagnostic category, or by honing in on a specific intervention. This review cast a much wider net in an effort to synthesize data on all psychotherapies, for any diagnoses, which reported on all three outcomes. It allows for a preliminary understanding of what has worked and for whom, and clinicians may find Table 2 useful in identifying potentially effective interventions for a specific population of parents and/or children. The review also sheds light on what needs are still left understudied and unserved by the current state of evidence-based interventions. To advance efforts to fill this gap, we propose four high priority future directions regarding research design and intervention development in this area.

Future Directions

(1) Researchers studying adult psychotherapy and child-focused interventions should consider including secondary outcome measures pertaining to parent and

child symptoms, as well as parenting behavior. This would drastically increase the field's understanding of the secondary effects of psychotherapeutic interventions on family members who may not be direct recipients of the intervention. Expanding this research may eventually yield data on cost-effectiveness, by being able to compare dosage needed to yield improvement across different treatment approaches.

(2) More longitudinal research incorporating repeated measures of parent and child psychopathology and parenting quality are needed in order to examine the transactional relationships among these variables. Testing the degree of influence that each member has on each other would refine the field's understanding of potential transactional effects and provide insight into how best to intervene.

(3) There is a significant need to develop interventions for families in which both the parent and child have clinically elevated symptoms of psychopathology. These interventions should carefully consider whether they target two (e.g., parent psychopathology and parenting behavior) or three outcomes (also intervening directly on child psychopathology), depending on the population for which it is being developed. Further, intervention developers should consider the manner in which multiple targets are integrated within treatment. This may be achieved by providing supplemental sessions devoted to additional targets of intervention (an approach which may lend itself to combining two existing evidence-based interventions), but may also require taking a more blended approach that contextualizes the specific difficulties faced by dually-disordered parent–child dyads (suggesting novel intervention development). Once an intervention is developed and ready for testing, a final recommendation is to include repeated measures of all three outcomes in order to provide critical information regarding the processes of change within these families.

(4) The last recommendation also stems from the finding that with few exceptions, there are very limited options for families in which both parent and child have psychopathology. Progress in this area will be inefficient if future intervention development targets one combination of parent and child psychopathology at a time. Rather, intervention developers should strongly consider targeting transdiagnostic features of psychopathology. A transdiagnostic approach, such as focusing on emotion dysregulation, might increase the likelihood that the intervention would be effective at treating parents and children with various forms of psychopathology.

Conclusion

This systematic review fills a gap in the literature by identifying and characterizing interventions that improve psychopathology in parents and children, as well as improve parenting behavior. It also highlights the need for interventions which are able to improve multiple interlinked outcomes in families experiencing mental health symptoms. As psychopathology in parents and children is highly co-occurring, and parenting behavior is a key mechanism in the transmission of risk, with bidirectional associations, there is a need to know what types of interventions may be able to assist these families. While many interventions show evidence of improving all three outcomes, few report such improvements in dually symptomatic parent–child samples. Families in which both parents and children are experiencing clinical-level mental health symptoms may benefit most from integrated interventions, and our review suggests this is an area in need of future intervention development efforts.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10567-021-00355-3.

Acknowledgements The authors wish to thank Barbara Jenkins for providing peer-review of the search strategy. We would also like to thank Rebecca Daigle and Thomas Hodgson for their contributions to data extraction.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Research Involving Human and Animal Participants This article does not contain any studies with human participants or animals performed by any of the authors.

References

Acri, M. C., & Hoagwood, K. E. (2015). Addressing parental mental health within interventions for children: A review. *Research on Social Work Practice, 25*(5), 578–586. https://doi.org/10.1177/1049731514546027

Barlow, J., Smailagic, N., Huband, N., Rollof, V., & Bennett, C. (2012). Group-based parent training programmes for improving parental psychosocial health. *Cochrane Database of Systematic Reviews (online)*. https://doi.org/10.1002/14651858.CD002020.pub3

Barroso, N. E., Mendez, L., Graziano, P. A., & Bagner, D. M. (2018). Parenting stress through the lens of different clinical groups: A systematic review and meta-analysis. *Journal of Abnormal Child Psychology, 46*(3), 449–461. https://doi.org/10.1007/s10802-017-0313-6

Beardslee, W. R., Brent, D. A., Weersing, V. R., Clarke, G. N., Porta, G., Hollon, S. D., & Garber, J. (2013). Prevention of depression in at-risk adolescents: Longer-term effects. *JAMA Psychiatry, 70*(11), 1161–1170. https://doi.org/10.1001/jamapsychiatry.2013.295

Breaux, R. P., Brown, H. R., & Harvey, E. A. (2017). Mediators and moderators of the relation between parental ADHD symptomatology and the early development of child ADHD and ODD symptoms. *Journal of Abnormal Child Psychology, 45*(3), 443–456. https://doi.org/10.1007/s10802-016-0213-1

Bywater, T., Hutchings, J., Daley, D., Whitaker, C., Yeo, S. T., Jones, K., & Edwards, R. T. (2009). Long-term effectiveness of a parenting intervention for children at risk of developing conduct disorder. *The British Journal of Psychiatry, 195*(4), 318–324. https://doi.org/10.1192/bjp.bp.108.056531

Carrion, V. G., Kletter, H., Weems, C. F., Berry, R. R., & Rettger, J. P. (2013). Cue-centered treatment for youth exposed to interpersonal violence: A randomized controlled trial. *Journal of Traumatic Stress, 26*, 654–662. https://doi.org/10.1002/jts

Clarke, G. N., Hornbrook, M., Lynch, F., Polen, M., Gale, J., O'Connor, E., & Debar, L. (2002). Group cognitive-behavioral treatment for depressed adolescent offspring of depressed parents in a health maintenance organization. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(3), 305–313. https://doi.org/10.1097/00004583-200203000-00010

Cobham, V. E., Dadds, M. R., & Spence, S. H. (1998). The role of parental anxiety in the treatment of childhood anxiety. *Journal of Consulting and Clinical Psychology, 66*(6), 893–905. https://doi.org/10.1037/0022-006X.66.6.893

Colalillo, S., & Johnston, C. (2016). Parenting cognition and affective outcomes following parent management training: A systematic review. *Clinical Child and Family Psychology Review, 19*(3), 216–235. https://doi.org/10.1007/s10567-016-0208-z

Connell, S., Sanders, M. R., & Markie-Dadds, C. (1997). Self-directed behavioral family intervention for parents of oppositional children in rural and remote areas. *Behavior Modification, 21*(4), 379–408.

Cuijpers, P. (2016). *Meta-analyses in mental health research. A practical guide*. Pim Cuijpers Uitgeverij.

Cuijpers, P., Weitz, E., Karyotaki, E., Garber, J., & Andersson, G. (2015). The effects of psychological treatment of maternal depression on children and parental functioning: A meta-analysis. *European Child and Adolescent Psychiatry, 24*(2), 237–245. https://doi.org/10.1007/s00787-014-0660-6

Dittman, C. K., Farruggia, S. P., Keown, L. J., & Sanders, M. R. (2016). Dealing with disobedience: An evaluation of a brief parenting intervention for young children showing noncompliant behavior problems. *Child Psychiatry and Human Development, 47*(1), 102–112.

Elgar, F. J., Mills, R. S. L., McGrath, P. J., Waschbusch, D. A., & Brownridge, D. A. (2007). Maternal and paternal depressive symptoms and child maladjustment: The mediating role of parental behavior. *Journal of Abnormal Child Psychology, 35*(6), 943–955. https://doi.org/10.1007/s10802-007-9145-0

Escovar, E. L., Drahota, A., Hitchcock, C., Chorpita, B. F., & Chavira, D. A. (2019). Vicarious improvement among parents participating in child-focused cognitive-behavioral therapy for anxiety. *Child and Family Behavior Therapy, 41*(1), 16–31. https://doi.org/10.1080/07317107.2019.1571770

Eyberg, S. M., Nelson, M. M., & Boggs, S. R. (2008). Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. *Journal of Clinical Child and Adolescent Psychology, 37*(1), 215–237. https://doi.org/10.1080/15374410701820117

Eyden, J., Winsper, C., Wolke, D., Broome, M. R., & MacCallum, F. (2016). A systematic review of the parenting and outcomes experienced by offspring of mothers with borderline personality pathology: Potential mechanisms and clinical implications. *Clinical Psychology Review, 47*, 85–105. https://doi.org/10.1016/j.cpr.2016.04.002

Fearon, R. P., Bakersman-Kranenburg, M., van Ijzendoorn, M., Lapsley, A., & Roisman, G. (2010). The significance of insecure attachment and disorganization in the development of children's externalizing behavior: A meta-analytic study. *Child Development, 81*(2), 435–456. https://doi.org/10.1086/117168

Fite, P. J., Colder, C. R., Lochman, J. E., & Wells, K. C. (2006). The mutual influence of parenting and boys' externalizing behavior problems. *Journal of Applied Developmental Psychology, 27*(2), 151–164. https://doi.org/10.1016/j.appdev.2005.12.011

Franke, N., Keown, L. J., & Sanders, M. R. (2016). An RCT of an online parenting program for parents of preschool-aged children with ADHD symptoms. *Journal of Attention Disorders, 24*(12), 1716–1726.

Garber, J., Clarke, G. N., Weersing, V. R., Beardslee, W. R., Brent, D. A., Gladstone, T. R. G., & Iyengar, S. (2009). Prevention of depression in at-risk adolescents. *JAMA*. https://doi.org/10.1001/jama.2009.788

Ginsburg, G. S. (2009). The child anxiety prevention study: Intervention model and primary outcomes. *Journal of Consulting and Clinical Psychology, 77*(3), 580–587. https://doi.org/10.1038/jid.2014.371

Gonzales, M. A., & Jones, D. (2016). Cascading effects of BPT for child internalizing problems and caregiver depression. *Clinical Psychology Review, 50*, 11–21. https://doi.org/10.1002/jmri.25711.PET/MRI

Goodman, S. H., & Gotlib, I. H. (1999). Risk for psychopathology in the children of depressed mothers: A developmental model for understanding mechanisms of transmission. *Psychological Review, 106*(3), 458–490. https://doi.org/10.1037/0033-295X.106.3.458

Gopalan, G., Dean-Assael, K., Klingenstein, K., Chacko, A., & Mckay, M. M. (2011). Caregiver depression and youth disruptive behavior difficulties. *Social Work in Mental Health, 9*(1), 56–70. https://doi.org/10.1080/15332985.2010.494528

Groh, A. M., Roisman, G. I., van Ijzendoorn, M. H., Bakermans-Kranenburg, M. J., & Fearon, R. P. (2012). The significance of insecure and disorganized attachment for children's internalizing symptoms: A meta-analytic study. *Child Development, 83*(2), 591–610. https://doi.org/10.1111/j.1467-8624.2011.01711.x

Gross, H. E., Shaw, D. S., & Moilanen, K. L. (2008). Reciprocal associations between boys' externalizing problems and mothers' depressive symptoms. *Journal of Abnormal Child Psychology, 36*(5), 693–709. https://doi.org/10.1007/s10802-008-9224-x

Gunlicks, M. L., & Weissman, M. M. (2008). Change in child psychopathology with improvement in parental depression: A systematic review. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*(4), 379–389. https://doi.org/10.1097/CHI.0b013e3181640805

Haggerty, K. P., Skinner, M., Fleming, C. B., Gainey, R., & Catalano, R. F. (2008). Long-term effects of focus on families on substance use disorders among children of parents in methadone treatment. *Addiction, 103*(12), 2008–2016. https://doi.org/10.1038/jid.2014.371

Hoagwood, K. E., Jensen, P. S., Acri, M. C., Serene Olin, S., Eric Lewandowski, R., & Herman, R. J. (2012). Outcome domains in child mental health research since 1996: Have they changed and why does it matter? *Journal of the American Academy of Child and Adolescent Psychiatry*. https://doi.org/10.1016/j.jaac.2012.09.004

Hutchings, J., Gardner, F., Bywater, T., Daley, D., Whitaker, C., Jones, K., & Edwards, R. T. (2007). Parenting intervention in sure start services for children at risk of developing conduct disorder: Pragmatic randomised controlled trial. *BMJ, 334*(7595), 678–682. https://doi.org/10.1136/bmj.39126.620799.55

Hutchings, J., Griffith, N., Bywater, T., & Williams, M. E. (2017). Evaluating the incredible years toddler parenting programme with parents of toddlers in disadvantaged (flying start) areas of wales. *Child: Care, Health and Development, 43*(1), 104–113. https://doi.org/10.1111/cch.12415

Jans, T., Jacob, C., Warnke, A., Zwanzger, U., Groß-Lesch, S., Matthies, S., & Philipsen, A. (2015). Does intensive multimodal treatment for maternal ADHD improve the efficacy of parent training for children with ADHD? A randomized controlled multicenter trial. *Journal of Child Psychology and Psychiatry, 56*(12), 1298–1313. https://doi.org/10.1111/jcpp.12443

Kazdin, A. E., & Wassell, G. (2000). Therapeutic changes in children, parents, and families resulting from treatment of children with conduct problems. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*(4), 414–420. https://doi.org/10.1097/00004583-200004000-00009

Kierfeld, F., Ise, E., Hanisch, C., Görtz-Dorten, A., & Döpfner, M. (2013). Effectiveness of telephone-assisted parent-administered behavioural family intervention for preschool children with externalizing problem behaviour: A randomized controlled trial. *European Child and Adolescent Psychiatry, 22*(9), 553–565.

Kirkham, J. J., Dwan, K. M., Altman, D. G., Gamble, C., Dodd, S., Smyth, R., & Williamson, P. R. (2010). The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ (online), 340*(7747), 637–640. https://doi.org/10.1136/bmj.c365

Kluczniok, D., Boedeker, K., Hindi Attar, C., Jaite, C., Bierbaum, A. L., Fuehrer, D., & Bermpohl, F. (2018). Emotional availability in mothers with borderline personality disorder and mothers with remitted major depression is differently associated with psychopathology among school-aged children. *Journal of Affective Disorders, 231*, 63–73. https://doi.org/10.1016/j.jad.2018.02.001

Lieberman, A. F., Ghosh Ippen, C., & Van Horn, P. (2006). Child-parent psychotherapy: 6-month follow-up of a randomized controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 45*(8), 913–918. https://doi.org/10.1097/01.chi.0000222784.03735.92

Lieberman, A. F., Van Horn, P., & Ippen, C. G. (2005). Toward evidence-based treatment: Child-parent psychotherapy with preschoolers exposed to marital violence. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*(12), 1241–1248. https://doi.org/10.1097/01.chi.0000181047.59702.58

Loechner, J., Starman, K., Galuschka, K., Tamm, J., Schulte-Körne, G., Rubel, J., & Platt, B. (2018). Preventing depression in the offspring of parents with depression: A systematic review and meta-analysis of randomized controlled trials. *Clinical Psychology Review, 60*, 1–14. https://doi.org/10.1016/j.cpr.2017.11.009

Lovejoy, M. C., Graczyk, P. A., O'Hare, E., & Neuman, G. (2000). Maternal depression and parenting behavior: A meta-analytic review. *Clinical Psychology Review, 20*(5), 561–592. https://doi.org/10.1016/S0272-7358(98)00100-7

Luthar, S. S., & Suchman, N. E. (2000). Relational Psychotherapy mothers' group: A developmentally informed intervention for at-risk mothers. *Development and Psychopathology, 12*(2), 235.

Magill-Evans, J., Harrison, M. J., Rempel, G., & Slater, L. (2006). Interventions with fathers of young children: Systematic literature review. *Journal of Advanced Nursing, 55*(2), 248–264. https://doi.org/10.1111/j.1365-2648.2006.03896.x

Maliken, A. C., & Katz, L. F. (2013). Exploring the impact of parental psychopathology and emotion regulation on evidence-based parenting interventions: A transdiagnostic approach to improving treatment effectiveness. *Clinical Child and Family Psychology Review, 16*(2), 173–186. https://doi.org/10.1007/s10567-013-0132-4

Marston, N., Stavnes, K., Van Loon, L. M. A. A., Drost, L. M., Maybery, D., Mosek, A., & Reupert, A. (2016). A content analysis of intervention key elements and assessments (IKEA): What's in the black box in the interventions directed to families where a parent has a mental illness? *Child and Youth Services, 37*(2), 112–128. https://doi.org/10.1080/0145935X.2016.1104041

Martin, C. G., Everett, Y., Skowron, E. A., & Zalewski, M. (2019). The role of caregiver psychopathology in the treatment of childhood trauma with trauma-focused cognitive behavioral therapy: A systematic review. *Clinical Child and Family Psychology Review, 22*(3), 273–289. https://doi.org/10.1007/s10567-019-00290-4

McCart, M. R., & Sheidow, A. J. (2016). Evidence-based psychosocial treatments for adolescents with disruptive behavior. *Journal of Clinical Child and Adolescent Psychology, 45*(5), 529–563. https://doi.org/10.1080/15374416.2016.1146990

McGilloway, S., Mhaille, G. N., Bywater, T., Furlong, M., Leckey, Y., Kelly, P., & Donnelly, M. (2012). A parenting intervention for childhood behavioral problems: A randomized controlled trial in disadvantaged community-based settings. *Journal of Consulting and Clinical Psychology, 80*(1), 116–127. https://doi.org/10.1037/a0026304

McLaughlin, K. A., Gadermann, A. M., Hwang, I., Sampson, N. A., Al-Hamzawi, A., Andrade, L. H., & Kessler, R. C. (2012). Parent psychopathology and offspring mental disorders: Results from the WHO world mental health surveys. *British Journal of Psychiatry, 200*(4), 290–299. https://doi.org/10.1192/bjp.bp.111.101253

Mejia, A., Calam, R., & Sanders, M. R. (2015). A pilot randomized controlled trial of a brief parenting intervention in low-resource settings in Panama. *Prevention Science, 16*(5), 707–717.

Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., & Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: A comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Services Research, 14*(1), 1–10.

Middeldorp, C. M., Wesseldijk, L. W., Hudziak, J. J., Verhulst, F. C., Lindauer, R. J. L., & Dieleman, G. C. (2016). Parents of children with psychopathology: Psychiatric problems and the association with their child's problems. *European Child and Adolescent Psychiatry, 25*(8), 919–927. https://doi.org/10.1007/s00787-015-0813-2

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ, 339*, b2535–b2535. https://doi.org/10.1136/bmj.b2535

Niccols, A., Milligan, K., Smith, A., Sword, W., Thabane, L., & Henderson, J. (2012). Integrated programs for mothers with substance abuse issues and their children: A systematic review of studies reporting on child outcomes. *Child Abuse and Neglect, 36*(4), 308–322. https://doi.org/10.1016/j.chiabu.2011.10.007

Nordsletten, A. E., Larsson, H., Crowley, J. J., Almqvist, C., Lichtenstein, P., & Mataix-Cols, D. (2016). Patterns of nonrandom mating within and across 11 major psychiatric disorders. *JAMA Psychiatry, 73*(4), 354–361. https://doi.org/10.1001/jamapsychiatry.2015.3192

Palmer, M. L., Keown, L. J., Sanders, M. R., & Henderson, M. (2019). Enhancing outcomes of low-intensity parenting groups through sufficient exemplar training: A randomized control trial. *Child Psychiatry and Human Development, 50*(3), 384–399.

Panter-Brick, C., Burgess, A., Eggerman, M., McAllister, F., Pruett, K., & Leckman, J. F. (2014). Practitioner review: Engaging fathers—Recommendations for a game change in parenting interventions based on a systematic review of the global evidence. *Journal of Child Psychology and Psychiatry, 55*(11), 1187–1212. https://doi.org/10.1111/jcpp.12280

Pardini, D. A. (2008). Novel insights into longstanding theories of bidirectional parent-child influences: Introduction to the special section. *Journal of Abnormal Child Psychology, 36*(5), 627–631. https://doi.org/10.1007/s10802-008-9231-y

Patterson, G. R. (2002). The early development of coercive family process. In J. B. Reid, G. R. Patterson, & J. J. Snyder (Eds.), *Antisocial behavior in children and adolescents: A developmental*

analysis and the Oregon model for intervention (pp. 25–44). American Psychological Association.

Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problem–A meta-analytic review. *Journal of Child Psychology and Psychiatry, 47*(1), 99–111. https://doi.org/10.1111/j.1469-7610.2005.01544.x

Sameroff, A. (Ed.). (2009). *The transactional model of development: How children and contexts shape each other*. American Psychological Association.

Sampson, M., McGowan, J., Cogo, E., Grimshaw, J., Moher, D., & Lefebvre, C. (2009). An evidence-based practice guideline for the peer review of electronic search strategies. *Journal of Clinical Epidemiology, 62*(9), 944–952. https://doi.org/10.1016/j.jclinepi.2008.10.012

Sanders, M. R., Bor, W., & Morawska, A. (2007). Maintenance of treatment gains: A comparison of enhanced, standard, and self-directed triple P-positive parenting program. *Journal of Abnormal Child Psychology, 35*(6), 983–998. https://doi.org/10.1007/s10802-007-9148-x

Sanders, M. R., Dittman, C. K., Farruggia, S. P., & Keown, L. J. (2014). A comparison of online versus workbook delivery of a self-help positive parenting program. *The Journal of Primary Prevention, 35*(3), 125–133.

Sanders, M. R., & McFarland, M. (2000). Treatment of depressed mothers with disruptive children: A controlled evaluation of cognitive behavioral family intervention. *Behavior Therapy, 31*(1), 89–112. https://doi.org/10.1016/S0005-7894(00)80006-4

Schrank, B., Moran, K., Borghi, C., & Priebe, S. (2015). How to support patients with severe mental illness in their parenting role with children aged over 1 year? A systematic review of interventions. *Social Psychiatry and Psychiatric Epidemiology, 50*(12), 1765–1783. https://doi.org/10.1007/s00127-015-1069-3

Schneider, S., In-Albon, T., Nuendel, B., & Margraf, J. (2013). Parental panic treatment reduces children's long-term psychopathology: A prospective longitudinal study. *Psychotherapy and Psychosomatics, 82*(5), 346. https://doi.org/10.1159/000350448

Sellers, R., Harold, G. T., Mahedy, L., Potter, R., Langley, K., Thapar, A., & Collishaw, S. (2016). Examining whether offspring psychopathology influences illness course in mothers with recurrent depression using a high-risk longitudinal sample. *Journal of Abnormal Psychology, 125*(2), 256–266. https://doi.org/10.1037/abn0000080

Shonkoff, J. P., & Fisher, P. (2013). Rethinking evidence-based practice and two-generation programs to create the future of early childhood policy. *Developmental Psychopathology, 25*, 1635–1653. https://doi.org/10.1016/j.physbeh.2017.03.040

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology, 70*(1), 747–770. https://doi.org/10.1146/annurev-psych-010418-102803

Siegenthaler, E., Munder, T., & Egger, M. (2012). Effect of preventive interventions in mentally ill parents on the mental health of the offspring: Systematic review and meta-analysis. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*(1), 8-17.e8. https://doi.org/10.1016/j.jaac.2011.10.018

Sterne, J. A., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., & Higgins, J. P. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ*. https://doi.org/10.1136/bmj.14898

Swartz, H. A., Cyranowski, J. M., Cheng, Y., Zuckoff, A., Brent, D. A., Markowitz, J. C., & Frank, E. (2016). Brief psychotherapy for maternal depression: Impact on mothers and children. *Journal of the American Academy of Child and Adolescent Psychiatry, 55*(6), 495-503.e2. https://doi.org/10.1016/j.jaac.2016.04.003

Thanhäuser, M., Lemmer, G., De Girolamo, G., & Christiansen, H. (2017). Do preventive interventions for children of mentally ill parents work? Results of a systematic review and meta-analysis. *Current Opinion in Psychiatry, 30*(4), 283–299. https://doi.org/10.1097/YCO.0000000000000342

Turner, K. M., & Sanders, M. R. (2006). Help when it's needed first: A controlled evaluation of brief, preventive behavioral family intervention in a primary care setting. *Behavior Therapy, 37*(2), 131–142.

van der Bruggen, C. O., Stams, G. J. J. M., & Bögels, S. M. (2008). Research review: The relation between child and parent anxiety and parental control: A meta-analytic review. *Journal of Child Psychology and Psychiatry 49*(12), 1257–1269. https://doi.org/10.1111/j.1469-7610.2008.01898.x

Verduyn, C., Barrowclough, C., Roberts, J., Tarrier, T., & Harrington, R. (2003). Maternal depression and child behaviour problems. Randomised placebo-controlled trial of a cognitive-behavioural group intervention. *The British Journal of Psychiatry, 183*, 342–348. https://doi.org/10.1192/bjp.183.4.342

Vidair, H. B., Reyes, J. A., Shen, S., Parrilla-Escobar, M. A., Heleniak, C. M., Hollin, I. L., & Rynn, M. A. (2011). Screening parents during child evaluations: Exploring parent and child psychopathology in the same clinic. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*(5), 441–450. https://doi.org/10.1016/j.jaac.2011.02.002

Webster-Stratton, C. (2014). The incredible years: Parents, teachers, and children training series. *Innovative Mental Health Interventions for Children: Programs That Work*. https://doi.org/10.4324/9781315786179

Wergeland, G. J. H., Fjermestad, K. W., Marin, C. E., Bjelland, I., Haugland, B. S. M., Silverman, W. K., & Heiervang, E. R. (2016). Predictors of treatment outcome in an effectiveness trial of cognitive behavioral therapy for children with anxiety disorders. *Behaviour Research and Therapy, 76*, 1–12. https://doi.org/10.1016/j.brat.2015.11.001

Yap, M. B. H., Morgan, A. J., Cairns, K., Jorm, A. F., Hetrick, S. E., & Merry, S. (2016). Parents in prevention: A meta-analysis of randomized controlled trials of parenting interventions to prevent internalizing problems in children from birth to age 18. *Clinical Psychology Review, 50*, 138–158. https://doi.org/10.1016/j.cpr.2016.10.003

The Clinical Process in Psychiatry: A Clinimetric Approach

Giovanni A. Fava, MD; Chiara Rafanelli, MD, PhD; and Elena Tomba, PhD

Objective: The aim of this review was to examine the clinical process in psychiatry, with special reference to clinimetrics, a domain concerned with the measurement of clinical phenomena that do not find room in customary taxonomy.

Data Sources: A MEDLINE search from inception to August 2010 was performed for English-language articles using the keywords *clinical judgment*, *clinimetric*, *staging*, *comorbidity*, *sequential treatment*, and *subclinical symptoms* in relation to psychiatric illness. It was supplemented by a manual search of the literature.

Study Selection: Choice of assessment strategies was based on their established or potential incremental increase in clinical information compared to use of diagnostic criteria.

Data Extraction: Contributions were evaluated according to the principles of clinimetrics.

Results: Several innovative assessment strategies were identified: the use of diagnostic transfer stations with repeated assessments instead of diagnostic endpoints, subtyping versus integration of different diagnostic categories, staging methods, and broadening of clinical information through macroanalysis and microanalysis. The most representative examples were selected.

Conclusions: Current assessment strategies in psychiatric research do not reflect the sophisticated thinking that underlies clinical decisions in practice. The clinimetric perspective provides an intellectual home for the reproduction and standardization of these clinical intuitions.

J Clin Psychiatry 2011;72(00):000–000

Submitted: *July 27, 2010; accepted October 20, 2010.*
Online ahead of print: *Month 00, 2011 (doi:10.4088/JCP.10r06444).*
Corresponding author: *Giovanni A. Fava, MD, Department of Psychology, Viale Berti Pichat 5, 40127 Bologna, Italy (giovanniandrea.fava@unibo.it).*

Psychiatric diagnosis and classification have attracted considerable attention in the past decades.[1] The introduction of diagnostic criteria for the identification of psychiatric syndromes, such as the *DSM*,[2] has considerably decreased the variance of diagnoses due to different assessors and the use of inferential criteria rather than direct observation.

However, clinicians have become increasingly aware of the limitations of the current diagnostic systems[3] and concerned about future *DSM* or *ICD* developments.[4] The customary clinical taxonomy in psychiatry does not include patterns of symptoms, severity of illness, effects of comorbid conditions, timing of phenomena, rate of progression of illness, responses to previous treatments, and other clinical distinctions that demarcate major prognostic and therapeutic differences among patients who otherwise seem to be deceptively similar since they share the same psychiatric diagnosis.

Little consideration has been given to the clinical process in psychiatry, that is, how clinical judgment leading to medical decisions is formulated. The main emphasis has been given to the standardization of the assessment process by use of rating scales leading to diagnostic configuration.[5]

In 1967, Alvan Feinstein dedicated a monograph to an analysis of clinical reasoning that underlies medical evaluations, such as the appraisal of symptoms, signs, and the timing of individual manifestations.[6] In 1982, he introduced the term *clinimetrics*[7] to indicate a domain concerned with the measurement of clinical issues that do not find room in customary clinical taxonomy. Such issues include the types, severity, and sequence of symptoms; rate of progression in illness (staging); severity of comorbidity; problems of functional capacity; reasons for medical decisions (eg, treatment choices); and many other aspects of daily life, such as well-being and distress.[8] Feinstein, in his book on clinimetrics,[8] quotes Molière's bourgeois gentleman who was astonished to discover that he spoke in prose as an example of clinicians who may discover that they constantly communicate with clinimetric indices. Indeed, in clinical practice, psychiatrists weigh factors such as the progression of disease, the overall severity of the disorder, the patient's social support and adaptation, resilience and reaction to stressful life circumstances, and response to previous treatment.[9] However, current formal strategies of assessment fail to capture most of this information.

We will examine some emerging trends and perspectives in the clinical process in psychiatry, with special reference to the diagnostic process, the staging method, and the organization of information.

DATA SOURCES AND STUDY SELECTION

A review of the literature, based on a MEDLINE search from inception to August 2010 using the keywords *clinical judgment*, *clinimetric*, *staging*, *comorbidity*, *sequential treatment*, and *subclinical symptoms* in relation to psychiatric illness was performed. It was supplemented by a manual search of the literature. Choice of assessment strategies was based on clinimetric principles[8] and on the concept of incremental validity,[10] which refers to the unique contribution or incremental increase in predictive power associated with the inclusion of a particular assessment procedure in the

Clinical Points

- Exclusive reliance on diagnostic criteria has impoverished the clinical process and does not reflect the complex thinking that underlies decisions in psychiatric practice.
- The accuracy of clinical judgment can be greatly increased with specific strategies: global formulations, staging methods, and a better organization of clinical information (encompassing macroanalysis and microanalysis).
- The concept of disease is no longer adequate to guide psychiatric care; therefore, clinical decision making should be addressed to attainment of individual goals.

clinical decision process.[11,12] We will then discuss the implications that a renewed interest in these assessment strategies may entail.

DIAGNOSTIC ENDPOINTS VERSUS TRANSFER STATIONS

In most instances of diagnostic reasoning in psychiatry, the process ends with the identification of a disorder,[13] often subsumed under a rubric of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*). A single assessment generates the prognostic and therapeutic judgments of the clinician. A *DSM* diagnosis (eg, major depressive disorder), however, encompasses a wide range of manifestations, comorbidity, seriousness, prognosis, and responses to treatment.

The majority of patients with mood and anxiety disorders do not qualify for 1, but for several Axis I and Axis II disorders.[14] As Cloninger[15] remarks, mental disorders can be characterized as manifestations of complex adaptive systems that are multidimensional in their description and multifactorial in their origins, and they involve nonlinear interactions in their development. As a result, efforts to describe psychopathology in terms of discrete categorical diagnoses result in extensive comorbidity and do not lend themselves to adequate treatment strategies.[15]

Very seldom do comorbid diagnoses undergo hierarchical organization (eg, generalized anxiety disorder and major depression) or is attention paid to the longitudinal development of mental illnesses. There is comorbidity that wanes upon successful treatment of 1 mental disease, eg, recovery from panic disorder with agoraphobia may result in remission from co-occurring hypochondriasis, without any specific treatment for the latter.[10] Other times, treatment of 1 disorder does not result in the disappearance of comorbidity. For instance, successful treatment of depression may not affect preexisting anxiety disturbances.[16]

The diagnostic criteria are particularly helpful in setting a threshold for conditions worthy of clinical attention. Accordingly, the diagnostic criteria for a major depressive disorder identify a syndrome that may be responsive to antidepressant drugs. At least 5 of a set of 9 symptoms should be present (and 1 should be either depressed mood or loss of interest). However, according to the psychometric model, all items are weighed the same, unlike in clinical medicine, where major and minor symptoms are often differentiated (eg, Jones criteria for rheumatic fever).[7] As a result, a patient with severe and pervasive anhedonia, incapacitating fatigue, and difficulties concentrating, which make him unable to work, would not be diagnosed with a major depressive disorder, despite the clinical intuition of potential benefit from pharmacotherapy. This diagnosis could be performed in a patient who barely meets the criteria for 5 symptoms. The hidden conceptual model is psychometric: severity is determined by the number of symptoms, not by their intensity or quality, to the same extent that a score in a depression self-rating scale depends on the number of symptoms that are scored as positive.[10] This is not surprising in view of the fact that the development of psychometrics took place outside of the clinical field, mainly in educational and social areas.[17] Since the phenomena under observation in the development of psychometric principles were not clinical, they could not be automatically adapted to clinical psychology and psychiatry.

Similar considerations apply to the longitudinal development of the disorder (prodromal phase, the fully developed disorder, and residual states).[9] Detre and Jarecki[18] provided a model for relating prodromal and residual symptomatology in psychiatric illness, defined as the rollback phenomenon, ie, as the illness remits, it progressively recapitulates, even though in a reverse order, many of the stages and symptoms that were seen during the time it developed. The rollback phenomenon has been substantiated in mood and anxiety disorders.[19,20] There is limited awareness of the fact that the current patient's symptomatology may have developed over the years and have reflected previous treatments.

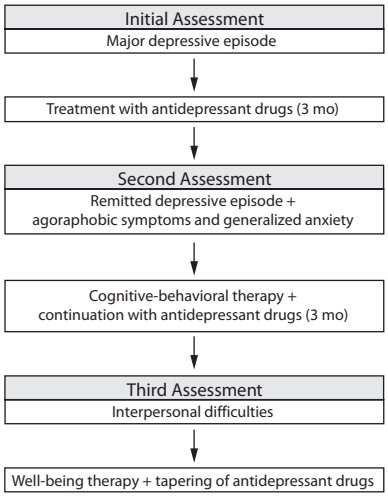
Feinstein[13] remarks that, when making a diagnosis, thoughtful clinicians seldom leap from a clinical manifestation to a diagnostic endpoint. The clinical reasoning goes through a series of "transfer stations," where potential connections between presenting symptoms and pathophysiological process are drawn. These stations are a pause for verification or change to another direction.[13] This strategy particularly applies to psychiatric disorders. An initial state of generalized anxiety may assume phobic connotations at some later point in time. If major depression then ensues, mood symptomatology may overshadow the previous anxiety disturbances, but the diagnosis of depression is only a transfer from prodromal to residual anxiety.

Some assessment strategies have been developed to overcome the flat, cross-sectional view of *DSM*.

Repeated Assessments

The use of diagnostic transfer stations has been suggested by the sequential treatment model,[21] an intensive, 2-stage approach, that includes the use of 1 treatment (eg, pharmacotherapy) after remission has been achieved. One type of treatment is thus employed to address the residual

Figure 1. Effects of Repeated Assessments on the Development of a Diagnostic Work



symptomatology that the other treatment was unable to affect. The sequential model relies on repeated assessments (after each line of treatment has been completed) that may modify an initial diagnosis (eg, preexisting anxiety disturbances may emerge after pharmacotherapy of a major depressive episode). Robins and Guze[22] developed the primary/secondary dichotomy in depression, which was based on chronology and course of follow-up. An episode of depression was defined as secondary when it was superimposed on a preexisting psychiatric or medical disease. The *DSM-IV*,[2] however, does not differentiate primary and secondary manifestations of depressive illness, as is performed in general medicine (eg, hypertension). As outlined in Figure 1, in view of the rollback phenomenon, Robins and Guze's primary/secondary distinction[22] becomes feasible: the major depressive episode appears to be superimposed on longstanding agoraphobic fears and avoidance and generalized anxiety. Symptoms are qualitatively differentiated (eg, the fact they persisted upon treatment against a background of improved symptomatology). They may be elicited by a diary or daily rating scales, which yield information that is not readily apparent in interview.

Subtyping

The need for subtyping major depressive disorder, since this category is too broad to yield meaningful treatment implications, has been recently underscored.[23,24] Lichtenberg and Belmaker,[23] for instance, differentiate between depression with anxiety (maintains functioning, positive response to favorable news or pleasurable activities) and late-life depression (no prior depressive history, reduced energy and interest, impaired cognitive function). Bech[24] has revived Robins and Guze's hierarchical primary/secondary distinction (eg, postnatal depression, poststroke late-life depression).[22] The basic assumption is that clinical manifestations that share the diagnosis of major depressive disorder may display substantial differences in prognostic and therapeutic terms.[23,24]

The underlying assumption is to increase the amount of clinical information that is conveyed by diagnosis. This requires use of instruments that yield a broad spectrum of information, such as hostility, irritable mood, and phobic avoidance, and are not ordinarily available.[24]

Building Unitary Concepts

Tyrer and associates[25] remarked that what is shared by syndromes such as anxiety, panic, phobic disturbances, and irritability may be as important as the differences between them, and conditions that are apparently comorbid could be part of the same clinical syndrome. They argued that the combination of mixed anxiety and depressive disorders together with a certain type of abnormal personality (excessive timidity, poor-self-esteem, avoidance of anxiety-provoking situations, and dependence on others) constitutes a single syndrome, the general neurotic syndrome.[25] The syndrome was shown to be associated with a poor response to treatment, frequent symptoms throughout the neurotic diagnostic spectrum, and tendency to relapse. The concept of neurosis, in its phenomenological[26] and psychodynamic[27] traditions, still has a lot to teach in terms of clinical thinking.[28]

Another example of search for unitary mechanisms of symptom formation is van Praag's Scale for Personality Disturbances.[29] On the basis of a structured interview, the rater is asked to score the following experiential qualities: (1) basic feelings of discontent with one's life situation and psychological make up, (2) unhappiness with one's personal relationships, and (3) emotional instability. The scale aims to overcome the difficulties in incorporating the I and II Axes of *DSM* and was found to allow important differentiations from residual symptomatology.[30]

The concept of allostatic load (the cumulative effects of stressful experiences in daily life) originated from basic science.[31] However, it offers another clinical opportunity of assessing the presence of a source of distress in the form of recent life events and/or chronic stress that exceed the individual's coping skills together with symptomatic manifestations encompassing psychological symptoms.[32] These approaches may be subsumed under the clinimetric rubric of global assessment indices. While the sensitivity of these methods is acknowledged in drug trials, where they often yield the most sensitive discrimination between drug and placebo effects,[33] the clinical value of these global evaluations in assessment and treatment planning is currently underestimated.

STAGING

In 1993, Fava and Kellner[9] introduced the clinimetric concept of staging in psychiatric classification. Unlike in clinical medicine, where this method had achieved wide currency (eg, the New York Heart Association Functional Classification, the Ann Arbor staging classification of Hodgkin's disease), staging was largely neglected in psychiatry. Staging

Table 1. Stages of a Psychiatric Disorder

Stage 1: Prodromal phase
Stage 2: Acute manifestations
Stage 3: Residual phase
Stage 4: Chronic (in attenuated or persistent form)

Table 2. Staging of Levels of Treatment Resistance

Stage 0: No history of failure to respond to therapeutic trial
Stage 1: Failure of at least 1 adequate therapeutic trial
Stage 2: Failure of at least 2 adequate therapeutic trials
Stage 3: Failure of 3 or more adequate therapeutic trials
Stage 4: Failure of 3 or more adequate trials including at least 1 concerned with augmentation/combination

Table 3. Staging of Loss of Therapeutic Effects During Continuation or Maintenance Treatment

Stage 0: No loss of therapeutic effect
Stage 1: Loss of therapeutic effects after adequate response in a therapeutic trial
Stage 2: Loss of therapeutic effects after adequate response in 2 therapeutic trials
Stage 3: Loss of therapeutic effects after adequate responses in 3 or more therapeutic trials

differs from the conventional diagnostic practice in that it not only defines the extent of progression of a disorder at a particular point in time but also reveals a person's current location on the continuum of the course of illness. Thus, once an index defines the existence of a particular disease state, its seriousness, extent, and longitudinal characteristics need to be evaluated.[8]

Fava and Kellner[9] developed staging methods for unipolar depression, bipolar disorder, panic disorder, and schizophrenia. Table 1 outlines the basic steps of development of a psychiatric disorder, ranging from the prodromal to the residual and chronic forms, in a longitudinal view of development of disturbances. Staging models have subsequently been refined in schizophrenia,[34] mood disorders,[35–38] and agoraphobia,[28] and they have been introduced in anorexia.[39] Staging instruments have also been developed.[40,41] In 2 randomized controlled trials,[42,43] psychotherapeutic intervention was applied according to a staging method and was found to yield long-term benefits.[44,45]

Further, the staging method has been applied to treatment response in depression.[46–48] It appears that the more information included in the method, the stronger its predictive value.[49] This information may encompass the number of trials completed,[49] the intensity/optimization of each trial,[49] issues of pseudoresistance (nonresponse to inadequate treatment in terms of duration, doses, or indications),[50] or occurrence of loss of therapeutic effects after clinical response.[51] Table 2 provides an illustration of the various levels of treatment resistance. By a clinical viewpoint, it is quite different to treat a patient with a major depressive episode who displayed positive responses to previous therapeutic trials (stage 0) and a patient who failed to respond to various adequate trials, including one concerned with augmentation/combination (stage 4). Similarly, if we encounter a depressed patient who repeatedly displayed loss of therapeutic response using various antidepressant drugs (Table 3), we should be aware that use of a new antidepressant is likely to yield the same phenomenon, probably because of a mechanism of oppositional tolerance.[51] For instance, many patients who did not respond to initial treatment in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial and went through various types of treatments, including augmentation/combination, were characterized by a refractory state with low remission, high relapse, and high intolerance rates.[35] Accordingly, their likelihood of lasting remission would be very low, as indicated by the staging methods of Tables 2 and 3.

Motivation to treatment and changing behavior has also been submitted to a staging system and may yield valuable insights into psychological resistances of the patient.[52] Di Clemente and Prochaska[52] developed a helpful staging method: "precontemplation" (people do not recognize that a problem exists and have no intention to change), "contemplation" (individuals accept that a problem exists but are ambivalent about it), "preparation/determination" (a perceived discrepancy between current and desired study), "action," and "maintenance" of the new patterns. It is difficult to suggest a psychotherapeutic treatment, despite pertinent indications, to a patient who is in the "precontemplation" stage. However, this is seldom considered, particularly in randomized controlled trials of psychotherapy.

ORGANIZATION OF CLINICAL INFORMATION

The information we previously mentioned adds to other customary domains of the clinical evaluation, such as psychiatric history, background of alcohol and other substance abuse, general medical history, physical examination, laboratory tests, and diagnostic interviews, whether they follow specific instruments or a more personal format.[5] There are other areas, however, that need to be addressed and are currently neglected.

Subclinical Distress and Illness Behavior

A diagnostic interview and a set of criteria have been used extensively in psychosomatic research.[53–55] The Diagnostic Criteria for Psychosomatic Research allow one to translate in clinical terms the spectrum of manifestations of illness behavior, ie, the ways in which individuals experience, perceive, evaluate, and respond to their health status.[53–55] The 2 main forms of abnormal illness behavior (illness affirming and illness denying) have several common expressions in psychiatric practice. However, the psychopathology of insight—as defined by Lewis[56]—is seldom examined. When this happens, the results can be quite interesting. For instance, in a recent investigation on the spectrum of anxiety disorders in the medically ill, agoraphobia without history of panic attacks was found to be closely related to the Diagnostic Criteria for Psychosomatic Research illness denial.[57] Persistent denial of having a medical disorder and

needing treatment frequently occurs in the medical setting.[53] If panic attacks have not taken place (illness denial was not associated with panic disorder and agoraphobia), agoraphobic fears tend to be highly rationalized and do not lead individuals to seek medical attention.[57] The identification of these fears requires careful expert interviewing, well beyond the checklist use of diagnostic instruments, to overcome the denial that underlies agoraphobia and other distress manifestations. The linking between agoraphobia without history of panic attacks and Diagnostic Criteria for Psychosomatic Research illness denial provides an explanation for some discrepancies that have occurred in the literature as to the prevalence of agoraphobia in clinical samples compared to epidemiologic studies.[28] Other important constructs covered by the Diagnostic Criteria for Psychosomatic Research are demoralization,[58] irritable mood,[53] and alexithymia.[27,59]

Psychological Well-Being

An area that is currently neglected in assessment is psychological well-being, despite the availability of validated instruments and its growing importance in establishing resilience.[3,60] Dimensions such as environmental mastery, personal growth, purpose in life, autonomy, self-acceptance, and positive relations with others were found to affect vulnerability to life adversities and complex balance between positive and negative affects in mood and anxiety disorders.[60]

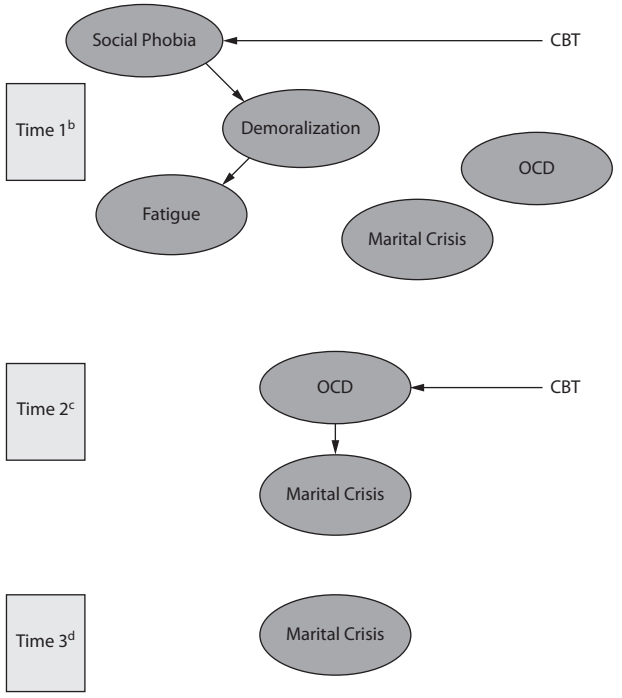
Mezzich and Salloum[3] developed the Person-centered Integrative Diagnosis, which encompasses both the positive and negative aspects of health, in an interactive way, within the person's life context. The Person-centered Integrative Diagnosis includes both the symptomatology of mental disorders and the positive aspects of health (adaptive functioning, protective factors, quality of life, etc) according to a holistic view of the person (including his/her dignity, values, and aspirations).[3] Rehabilitation of mental disorders is targeted as much on the patient's strengths and wishes as it is on alleviating symptoms and psychopathology.[61]

Macroanalysis and Microanalysis

Feinstein, when he introduced the concept of comorbidity, referred to any "additional coexisting ailment" separate from the primary disease, even if this secondary phenomenon does not qualify as a disease per se.[62] Indeed, in clinical medicine, the many methods that are available for measuring comorbidity are not limited to disease entities.[63]

A method has been developed in psychiatry for organizing clinical data as variables in clinical reasoning. Emmelkamp et al[64,65] have introduced the concept of macroanalysis (a relationship between co-occurring syndromes and problems is established on the basis of when treatment should commence). Fava and Sonino[54] have applied macroanalysis to assessing the relationship between medical and psychological variables. Macroanalysis starts from the assumption that, in most cases, there are functional relationships with other more or less clearly defined problem areas[64] and that the targets of treatment may vary during the course of disturbances.[54]

Figure 2. Example of Macroanalysis[a]



[a]A patient presents with work situational social phobia, demoralization, fatigue, obsessive-compulsive disorder (OCD) symptoms, and marital crisis.
[b]At time 1, the therapist could give priority to cognitive-behavioral therapy (CBT) of social phobia, expecting a consequent improvement in demoralization and sense of fatigue.
[c]At time 2, the therapist could decide to intervene on OCD symptoms by using CBT techniques to emphasize the negative effects of the patient's excessive preoccupation for order and precision, leading to a chronic malaise and communicative difficulties with the partner.
[d]At posttherapy assessment (time 3), the therapist could determine the relationship of OCD symptoms to marital crisis.

For instance, a patient may present with work situational social phobia (which leads him or her to avoid important opportunities for improving his or her job), demoralization (which increases his or her sense of fatigue), marital crisis (as a result of obsessional traits of mental order incompatible with that of his or her spouse), and obsessive ruminations (which lead to a chronic state of indecision). In terms of macroanalysis, the clinician, after a thorough interview with the patient, could place into a hierarchy the syndromes and symptoms of comorbidity by considering also the patient's needs. The clinician could thus give priority to the cognitive-behavioral treatment of social phobia, leaving to posttherapy assessment the determination of the relationship of social phobia to demoralization, marital crisis, and obsessional ruminations. Will they wane as anxious epiphenomena or will they persist, despite some degree of improvement? Should, in this latter case, further treatment be necessary? What type of relationship do demoralization and obsessive-compulsive symptoms entertain? If the clinical decision of tackling one syndrome may be taken during the initial assessment, the subsequent steps of macroanalysis require a reassessment after the first line of treatment has terminated (Figure 2).

The hierarchical organization that is chosen may depend on a variety of factors (urgency, availability of treatment tools, etc) that include also the patient's preferences and priorities. Macroanalysis is a tool that allows the therapist to not only increase accuracy in clinical decision making but also inform the patient about the relationship between different problem areas and motivate the patient for changing.[64,65] The concept of shared decision making is getting increasing attention in clinical medicine,[66] but it is still seldom practiced in psychiatry.[67] Macroanalysis also requires reference to the staging method, whereby a disorder is characterized according to seriousness, extension, and longitudinal development.[9] For instance, certain psychotherapeutic strategies can be deferred to a residual stage of depression when state-dependent learning has been improved by use of antidepressant drugs.[68] The planning of treatment thus requires determination of the symptomatic target of the first-line approach (eg, pharmacotherapy) and tentative identification of other areas of concern to be addressed by subsequent treatment (eg, psychotherapy).

Macroanalysis should be supplemented by microanalysis, a detailed analysis of specific symptoms (onset and course of the complaints, circumstances that worsen symptoms and consequences).[64,65] For instance, when anxiety characterizes the clinical picture, it is necessary to know under which circumstances the anxiety become manifest and how the patient responds when he/she becomes anxious, and also to know whether an avoidant behavior occurs and, if so, what are the long-term consequences of the avoidant behavior.

Targum and associates[69] have developed specific criteria (SAFER) to be used in drug trials for improving the assessment accuracy of symptoms: State versus trait (the identified symptoms must reflect the current state of illness and not long-standing traits), Accessibility, Face validity, Ecological validity, and Rule of the 3 *p*'s (symptoms must be present, persistent, and pathological). The SAFER criteria inventory constitutes a valid method of microanalysis. Microanalysis also consists of dimensional measurements, such as observer or self-rating scales for assessing anxiety and fears. Choice of these instruments is dictated by the clinimetric concept of incremental validity.[10–12] Each distinct aspect of psychological measurement should deliver a unique increase in information in order to qualify for inclusion. The concept can also be applied to the selection of instruments in a psychometric battery. In clinical research, several highly redundant scales are often used under the misguided assumption that nothing will be missed. On the contrary, violation of the concept of incremental validity leads to only conflicting results. Microanalysis is consequential and secondary to macroanalysis and leads to overcoming the assumption that there is a common assessment strategy for all clinical encounters.

CONCLUSION

Part of the challenge and, at the same time, fascination of being a clinician lies in applying scientific methods in the care of patients and in understanding disease.[70] Greater knowledge should result in significant benefits for the patients, and, in a sense, continued development on the part of the physician.[71] We are witnessing, however, a progressive detachment of clinicians from research, which is often accompanied by a sense of personal stagnation and tiredness.[71] This detachment is mainly the reflection of an intellectual crisis that has become more and more manifest in recent years.[71–73]

In 1967, Feinstein[6] urged clinicians to develop a "basic science" of their own—to study the clinical phenomena directly, to specify the importance of different types of clinical data, to create appropriate systems of taxonomy for classifying the information, and to develop intellectual models and pragmatic methods that would articulate the clinical process and use the results for quantified analyses.

More recently, Tinetti and Fried[74] have argued that time has come to abandon disease as the focus of medical care. Clinical decision making for all patients should be addressed to attainment of individual goals and identification and treatment of all modifiable and nonbiological factors, rather than solely to the diagnosis and treatment of individual diseases.[74]

Often, in their clinical practice, psychiatrists use sophisticated forms of clinical judgment that are suitable for clinical challenges but are not addressed by current research strategies. Exclusive reliance on diagnostic criteria has impoverished the clinical process and does not reflect the complex thinking that underlies decisions in psychiatric practice. The use of transfer stations with repeated assessments instead of diagnostic endpoints, the building of global formulations of clinical integration, staging methods, and a better organization of clinical information (encompassing subclinical distress, illness behavior, psychological well-being, macroanalysis, and microanalysis) may be an antidote to oversimplified models that derive from biological reductionism, neglect individual responses to treatment, and clash with clinical reality.[71,75]

The clinimetric perspective provides an intellectual home for the reproduction and standardization of the clinical intuitions. It allows the clinician to make full use of the clinical information that is available. It opens a new exciting area of research that is likely to yield improved targets for neurobiological studies and treatment trials.

Author affiliations: Affective Disorders Program, Department of Psychology, University of Bologna, Italy (all authors); and the Department of Psychiatry, State University of New York at Buffalo (Dr Fava).
Potential conflicts of interest: None reported.
Funding/support: None reported.
Previous presentation: A modified version of this article was presented at the workshop "Clinical Judgment in Psychiatry: a Tribute to George Molnar, MD"; Aug 13, 2010; Buffalo, New York.

REFERENCES

1. Maj M, Gaebel W, Lopez-Ibor JJ, et al, eds. *Psychiatric Diagnosis and Classification*. Chichester, UK: Wiley; 2002.
2. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition. Washington, DC: American Psychiatric Association; 1994.
3. Mezzich JE, Salloum IM. Towards innovative international classification and diagnostic systems: ICD-11 and person-centered integrative

diagnosis. *Acta Psychiatr Scand*. 2007;116(1):1–5.
4. Frances A. Whither *DSM-V*? *Br J Psychiatry*. 2009;195(5):391–392.
5. American Psychiatric Association. Psychiatric Evaluation of Adults. Second Edition. *Am J Psychiatry*. 2006; 163 (June suppl): 1–36
6. Feinstein AR. *Clinical Judgment*. Baltimore, MD: Williams & Wilkins; 1967.
7. Feinstein AR. T. Duckett Jones Memorial Lecture: The Jones criteria and the challenges of clinimetrics. *Circulation*. 1982;66(1):1–5.
8. Feinstein AR. *Clinimetrics*. New Haven, CT: Yale University Press; 1987.
9. Fava GA, Kellner R. Staging: a neglected dimension in psychiatric classification. *Acta Psychiatr Scand*. 1993;87(4):225–230.
10. Fava GA, Ruini C, Rafanelli C. Psychometric theory is an obstacle to the progress of clinical research. *Psychother Psychosom*. 2004;73(3):145–148.
11. Hunsley J, Meyer GJ. The incremental validity of psychological testing and assessment: conceptual, methodological, and statistical issues. *Psychol Assess*. 2003;15(4):446–455.
12. Garb HN. Incremental validity and the assessment of psychopathology in adults. *Psychol Assess*. 2003;15(4):508–520.
13. Feinstein AR. An analysis of diagnostic reasoning. I. The domains and disorders of clinical macrobiology. *Yale J Biol Med*. 1973;46(3):212–232.
14. Zimmerman M, Chelminski I, McDermut W. Major depressive disorder and Axis I diagnostic comorbidity. *J Clin Psychiatry*. 2002;63(3):187–193.
15. Cloninger CR. Implications of comorbidity for the classification of mental disorders: the need for a psychobiology of coherence. In: Maj M, Gaebel W, Lopez-Ibor JJ, et al, eds. *Psychiatric Diagnosis and Classification*. Chichester, UK: Wiley; 2002:79–106.
16. Sherbourne CD, Wells KB. Course of depression in patients with comorbid anxiety disorders. *J Affect Disord*. 1997;43(3):245–250.
17. Rust J, Golombok S. *Modern Psychometrics: The Science of Psychological Assessment*. London, UK: Routledge; 1989.
18. Detre TP, Jarecki H. *Modern Psychiatric Treatment*. Philadelphia, PA: Lippincott; 1971.
19. Fava GA. Subclinical symptoms in mood disorders: pathophysiological and therapeutic implications. *Psychol Med*. 1999;29(1):47–61.
20. Fava GA, Mangelli L. Subclinical symptoms of panic disorder: new insights into pathophysiology and treatment. *Psychother Psychosom*. 1999;68(6):281–289.
21. Fava GA, Ruini C, Rafanelli C. Sequential treatment of mood and anxiety disorders. *J Clin Psychiatry*. 2005;66(11):1392–1400.
22. Robins E, Guze SB. Classification of affective disorders: the primary-secondary, the endogenous-reactive, and the neurotic-psychotic concepts. In: Williams TA, Katz MM, Shield JA, eds. *Recent Advances in the Psychobiology of the Depressive Illness*. Washington, DC: Government Printing Office; 1972:283–293.
23. Lichtenberg P, Belmaker RH. Subtyping major depressive disorder. *Psychother Psychosom*. 2010;79(3):131–135.
24. Bech P. Struggle for subtypes in primary and secondary depression and their mode-specific treatment or healing. *Psychother Psychosom*. 2010; 79(6):331–338.
25. Tyrer P, Seivewright N, Ferguson B, et al. The general neurotic syndrome: a coaxial diagnosis of anxiety, depression and personality disorder. *Acta Psychiatr Scand*. 1992;85(3):201–206.
26. Slater E, Slater P. A heuristic theory of neurosis. *J Neurol Neurosurg Psychiatry*. 1944;7(1–2):49–55.
27. Taylor GJ. Affects, trauma, and mechanisms of symptom formation: a tribute to John C. Nemiah, MD (1918–2009). *Psychother Psychosom*. 2010;79(6):339–349.
28. Fava GA, Rafanelli C, Tossani E, et al. Agoraphobia is a disease: a tribute to Sir Martin Roth. *Psychother Psychosom*. 2008;77(3):133–138.
29. van Praag HM. Diagnosing depression—looking backward into the future. *Psychiatr Dev*. 1989;7(4):375–394.
30. Rafanelli C, Park SK, Ruini C, et al. Rating well-being and distress. *Stress Med*. 2000;16(1):55–61.
31. McEwen BS. Physiology and neurobiology of stress and adaptation: central role of the brain. *Physiol Rev*. 2007;87(3):873–904.
32. Fava GA, Guidi J, Semprini F, et al. Clinical assessment of allostatic load and clinimetric criteria. *Psychother Psychosom*. 2010;79(5):280–284.
33. Fava M, Evins AE, Dorer DJ, et al. The problem of the placebo response in clinical trials for psychiatric disorders: culprits, possible remedies, and a novel study design approach. *Psychother Psychosom*. 2003;72(3): 115–127.
34. McGorry P, Hickie IB, Yung AR, et al. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. *Aust N Z J Psychiatry*. 2006;40(8):612–622.
35. Fava GA, Tomba E, Grandi S. The road to recovery from depression—don't drive today with yesterday's map. *Psychother Psychosom*. 2007;76(5):260–265.
36. Berk M, Hallam KT, McGorry PD. The potential utility of a staging model as a course specifier: a bipolar disorder perspective. *J Affect Disord*. 2007;100(1–3):279–281.
37. Kapczinski F, Dias VV, Kauer-Sant'Anna M, et al. The potential use of biomarkers as an adjunctive tool for staging bipolar disorder. *Prog Neuropsychopharmacol Biol Psychiatry*. 2009;33(8):1366–1371.
38. Reinares M, Colom F, Rosa AR, et al. The impact of staging bipolar disorder on treatment outcome of family psychoeducation. *J Affect Disord*. 2010;123(1–3):81–86.
39. Maguire S, Le Grange D, Surgenor L, et al. Staging anorexia nervosa: conceptualizing illness severity. *Early Interv Psychiatry*. 2008;2(1):3–10.
40. Bilsbury CD, Richman A. A staging approach to measuring patient-centred subjective outcomes. *Acta Psychiatr Scand*. 2002; 106(suppl s414):5–40.
41. Andresen R, Caputi P, Oades L. Stages of recovery instrument: development of a measure of recovery from serious mental illness. *Aust N Z J Psychiatry*. 2006;40(11–12):972–980.
42. Fava GA, Grandi S, Zielezny M, et al. Cognitive behavioral treatment of residual symptoms in primary major depressive disorder. *Am J Psychiatry*. 1994;151(9):1295–1299.
43. Fava GA, Rafanelli C, Grandi S, et al. Prevention of recurrent depression with cognitive behavioral therapy: preliminary findings. *Arch Gen Psychiatry*. 1998;55(9):816–820.
44. Fava GA, Grandi S, Zielezny M, et al. Four-year outcome for cognitive behavioral treatment of residual symptoms in major depression. *Am J Psychiatry*. 1996;153(7):945–947.
45. Fava GA, Ruini C, Rafanelli C, et al. Six-year outcome of cognitive behavior therapy for prevention of recurrent depression. *Am J Psychiatry*. 2004;161(10):1872–1876.
46. Thase ME, Rush AJ. Treatment-resistant depression. In: Bloom FE, Kupfer DT, eds. *Psychopharmacology: the Fourth Generation of Progress*. New York, NY: Raven Press; 1995:1081–1097.
47. Fava M. Diagnosis and definition of treatment-resistant depression. *Biol Psychiatry*. 2003;53(8):649–659.
48. Fekadu A, Wooderson SC, Markopoulou K, et al. The Maudsley Staging Method for treatment-resistant depression: prediction of longer-term outcome and persistence of symptoms. *J Clin Psychiatry*. 2009;70(7): 952–957.
49. Petersen T, Papakostas GI, Posternak MA, et al. Empirical testing of two models for staging antidepressant treatment resistance. *J Clin Psychopharmacol*. 2005;25(4):336–341.
50. Nierenberg AA, Amsterdam JD. Treatment-resistant depression: definition and treatment approaches. *J Clin Psychiatry*. 1990;51(suppl):39–47, discussion 48–50.
51. Fava GA. Can long-term treatment with antidepressant drugs worsen the course of depression? *J Clin Psychiatry*. 2003;64(2):123–133.
52. Di Clemente CC, Prochaska JO. Toward a comprehensive, transtheoretical model of change: stages of change and addictive behaviours. In: Miller WR, Heather N, eds. *Treating Addictive Behaviors*. 2nd ed. New York, NY: Plenum; 1998:3–24.
53. Porcelli P, Sonino N, eds. *Psychological Factors Affecting Medical Conditions. A New Classification for* DSM-V. Basel, Switzerland: Karger; 2007.
54. Fava GA, Sonino N. Psychosomatic assessment. *Psychother Psychosom*. 2009;78(6):333–341.
55. Wise TN. Diagnostic criteria for psychosomatic research are necessary for *DSM V*. *Psychother Psychosom*. 2009;78(6):330–332.
56. Lewis A. The psychopathology of insight. *Br J Med Psychol*. 1934;14: 332–348.
57. Fava GA, Porcelli P, Rafanelli C, et al. The spectrum of anxiety disorders in the medically ill. *J Clin Psychiatry*. 2010;71(7):910–914.
58. Cockram CA, Doros G, de Figueiredo JM. Diagnosis and measurement of subjective incompetence: the clinical hallmark of demoralization. *Psychother Psychosom*. 2009;78(6):342–345.
59. Ogrodniczuk JS, Piper WE, Joyce AS, et al. Alexithymia and treatment preferences among psychiatric outpatients. *Psychother Psychosom*. 2009; 78(6):383–384.
60. Fava GA, Tomba E. Increasing psychological well-being and resilience by psychotherapeutic methods. *J Pers*. 2009;77(6):1903–1934.
61. Ralph RO, Corrigan PW, eds. *Recovery in Mental Illness. Broadening our Understanding of Wellness*. Washington, DC: American Psychological Association; 2005.
62. Feinstein AR. The pre-therapeutic classification of comorbidity in chronic disease. *J Chronic Dis*. 1970;23(7):455–468.
63. de Groot V, Beckerman H, Lankhorst GJ, et al. How to measure

comorbidity. A critical review of available methods. *J Clin Epidemiol*. 2003;56(3):221–229.
64. Emmelkamp PMG, Bouman TK, Scholing A. *Anxiety Disorders*. Chichester, UK: Wiley; 1993:55–67.
65. Emmelkamp PMG. The additional value of clinimetrics needs to be established rather than assumed. *Psychother Psychosom*. 2004;73(3): 142–144.
66. Joosten EAG, DeFuentes-Merillas L, de Weert GH, et al. Systematic review of the effects of shared decision-making on patient satisfaction, treatment adherence and health status. *Psychother Psychosom*. 2008;77(4): 219–226.
67. Joosten EAG, de Jong CAJ, de Weert-van Oene GH, et al. Shared decision-making reduces drug use and psychiatric severity in substance-dependent patients. *Psychother Psychosom*. 2009;78(4):245–253.
68. Reus VI, Weingartner H, Post RM. Clinical implications of state-dependent learning. *Am J Psychiatry*. 1979;136(7):927–931.
69. Targum SD, Pollack MH, Fava M. Redefining affective disorders: relevance for drug development. *CNS Neurosci Ther*. 2008;14(1):2–9.
70. Engel GL. Physician-scientists and scientific physicians: resolving the humanism-science dichotomy. *Am J Med*. 1987;82(1):107–111.
71. Fava GA. The intellectual crisis of psychiatric research. *Psychother Psychosom*. 2006;75(4):202–208.
72. Maj M. Are psychiatrists an endangered species? *World Psychiatry*. 2010; 9(1):1–2.
73. Katschnig H. Are psychiatrists an endangered species? observations on internal and external challenges to the profession. *World Psychiatry*. 2010; 9(1):21–28.
74. Tinetti ME, Fried T. The end of the disease era. *Am J Med*. 2004;116(3): 179–185.
75. Wise TN. Curiosity and crisis. *Psychother Psychosom*. 2006;75(4):199–201.

Clinical Child and Family Psychology Review
https://doi.org/10.1007/s10567-025-00535-5

Perinatal Mental Health Interventions with a Parent–Child Relational Component: A Systematic Review of the Effects on Mothers and Dyads

Sarah E. Freeman[1] · Joely Reznik[1] · Maitri Jain[1] · Emily E. Sokol[2] · Dana Manning[1] · Haley Rohde[1] · Tamara M. Loverich[1] · Jamie M. Lawler[1]

Accepted: 8 June 2025

Abstract

Perinatal mental health challenges are associated with deleterious next-generation outcomes due in part to their impact on early parent–child relationships. Many perinatal interventions for caregivers with mental health difficulties target *either* parental mental health *or* parent–child relationships, despite ample evidence that these are related. This systematic review aimed to examine mental health and relational outcomes associated with perinatal interventions targeting caregiver *and* relational wellbeing. Seven electronic databases were searched to identify studies meeting the following criteria: (a) intervention targets parental mental health and parent–infant relational wellbeing (b) study design includes a comparison group, (c) study reports mental health and observed relational outcomes for all groups, (d) intervention is delivered to perinatal mothers by a professional, and (e) article is written in English. Data were tabulated to facilitate narrative synthesis and risk of bias analysis. Sixteen articles evaluating 13 interventions delivered to 1070 participants were included. Four interventions exhibited improvements in both outcomes, four demonstrated improvements in only one outcome, and five demonstrated improvements in neither. Clinical implications include support for the use of brief group and individual interventions that target mental health outcomes via support for emotion expression and problem-solving, and relational outcomes via in vivo guidance for sensitive parent–child interactions, promotion of reflective functioning, and addressing caregiver perceptions of parent–child relationships. Limitations related to study quality, design heterogeneity, and participant demographic homogeneity demand caution in interpretation of results and there is a need for higher quality research in this area. The review protocol was pre-registered with Prospero (ID: CRD42022380278).

Keywords Perinatal mental health · Mother–infant relationships · Attachment · Systematic review · Intergenerational transmission

✉ Sarah E. Freeman
sfreem24@emich.edu

Joely Reznik
joelydrew@yahoo.com

Maitri Jain
mjain2@emich.edu

Emily E. Sokol
sokolem1@msu.edu

Dana Manning
manningdanac@gmail.com

Haley Rohde
haleyrohde9@gmail.com

Tamara M. Loverich
tpenix@emich.edu

Jamie M. Lawler
jlawler1@emich.edu

[1] Psychology Department, Eastern Michigan University, Ypsilanti, MI 48197, USA

[2] Psychology Department, Michigan State University, East Lansing, MI 48824, USA

Published online: 07 July 2025

Introduction

A vast body of literature has documented the broad array of intergenerational impacts of perinatal mental health challenges. These include both proximal outcomes, such as increased likelihood of birth complications, low birthweight, and poorer socioemotional and cognitive outcomes in infancy (Cook et al., 2018; Dowse et al., 2020; Feldman et al., 2009), as well as more distal outcomes, such as elevated rates of both internalizing and externalizing psychopathology among offspring across childhood, adolescence, and adulthood (Barker et al., 2012; Essau et al., 2018; Murray et al., 2011). Although there is evidence supporting genetic vulnerability as one mechanism of this transmission (e.g., Willems et al., 2019), genes shared by parents and offspring make up one piece of this complex, multi-faceted, and transactional puzzle (Avinun & Knafo, 2014; Branje et al., 2020). Importantly, other avenues (e.g., environmental avenues), are also more readily modifiable and therefore more reasonable targets for intervention. Accordingly, much of the literature to date has emphasized the impact of perinatal mental health difficulties on early parent–child interactions and relationship quality as a key mechanism for this transmission (McAndrew, 2019). Interventions delivered during the perinatal period vary in terms of the degree to which they target perinatal mental health challenges versus parent–child relationships; this review aims to fill a gap in the literature by providing an overview of interventions that target both and their effectiveness.

The Impact of Parent–Child Relationships

According to attachment theory (Bowlby, 1958, 1969), parent–child interactions during infancy play a key role in child development due to the impact early interactions have on the development of attachment security (Ainsworth, 1973). Attachment theory posits that these early interactional patterns have lifelong implications via their effect on a young child's mental representations, or their "internal working models," of relationships, which includes their representations or beliefs about the self, caregiver, and caregiving relationship. These internal working models, informed by early caregiving experiences, are carried forward to guide future behaviors and expectations of relationships (Main et al., 1985).

Although findings regarding the relative stability of attachment style over time have been somewhat mixed (Groh et al., 2014; Waters et al., 2000), the lifelong significance of attachment security during infancy has been well-documented. For instance, attachment *in*security in infancy has consistently been found to be associated with both internalizing and externalizing problems in childhood (Madigan et al., 2016). The impact of infant attachment security on emotion regulatory abilities may account for these problems (Brumariu & Kerns, 2013). In contrast, longitudinal research has shown that children with secure attachment histories in infancy demonstrate higher self-reliance, self-regulation, social competence, and resilience in early and middle childhood (Sroufe, 2005), in addition to better relationship functioning in adulthood (Simpson et al., 2011), compared to infants classified as insecure and disorganized.

Parental sensitivity (i.e., the caregiver's tendency to "detect and recognize signals or situations that might require her response, to respond promptly to these situations, and to respond appropriately;" Pederson et al., 1990, p. 1976) is widely considered one of the most influential predictors of infant attachment security, particularly when in response to infant distress (i.e., as opposed to non-distress; Fearon & Belsky, 2016; McElwain & Booth-LaForce, 2006). In line with Bowlby's concept of the caregiver as a "secure base" from which an infant can explore (1988), caregiver support for infant autonomy has also been found to meaningfully contribute to the development of secure attachment (Bernier et al., 2014). In contrast, caregiver behaviors during infancy that are intrusive, controlling, unresponsive, or even frightening (e.g., dissociation, as may occur in the context of parental trauma) are associated with insecure or, in the latter case, disorganized, attachment (Fearon & Belsky, 2016; Out et al., 2009). Disorganized attachment, which occurs when infants are unable to develop reliable strategies for getting their needs met, may be associated with the most dire developmental outcomes, including elevated stress reactivity and distress intolerance, as well as later aggression (Van Ijzendoorn et al., 1999).

Interventions Targeting Attachment Security

In line with these empirical findings, interventions targeting attachment security frequently emphasize sensitive and responsive caregiving behavior, maternal representations of the child and caregiving relationship, parental reflective functioning (i.e., ability to hold the child's mental states in mind; Slade, 2005), psychoeducation about attachment and child development, and skin-to-skin contact or infant massage to promote closeness and emotion regulation (Erickson et al., 2019; Wright et al., 2023). Early reviews and meta-analyses of interventions targeting attachment security have found evidence for medium to large effect sizes on parental sensitivity (Cohen's $d = 0.33$–0.58) but weaker, though still significant, effect sizes for attachment insecurity (Cohen's $d = 0.17$–0.20; Bakermans-Kranenburg et al., 2003; Van Ijzendoorn et al., 1995). In these reviews, interventions with

the largest effect sizes targeted overt parenting behaviors as opposed to parental attachment representations and were shorter-term (i.e., 16 sessions or fewer; Bakermans-Kranenburg et al., 2003; Van Ijzendoorn et al., 1995).

A more recently conducted systematic review and meta-analysis concluded that interventions emphasizing these targets were associated with both reduced odds of disorganized attachment (Odds Ratio [OR] = 0.54, 95% Confidence Interval [CI_{95}] = 0.39–0.77) and increased odds of secure attachment (OR = 1.85, CI_{95} = 1.36–2.52; Wright et al., 2023). Importantly, this review also highlighted the fact that attachment interventions with the strongest evidence base (i.e., those incorporating the use of video feedback to promote sensitive caregiving) are among those least frequently used in the United Kingdom (Wright et al., 2023), underscoring the importance of dissemination of evaluation research, as well as the need to evaluate more rigorously those interventions that are being more frequently used.

A recent meta-analysis of studies evaluating a more psychodynamically-oriented intervention targeting attachment security via changes in maternal representations (Parent–Infant Psychotherapy; PIP) showed favorable outcomes in terms of attachment security when compared to a control group, despite failing to find evidence of between-group differences in terms of parent–child interactions (Barlow et al., 2015), suggesting that other important mechanisms of change may also be at play in these treatments. Of note, however, no significant differences in attachment outcomes favoring PIP were detected when this intervention was compared to alternative treatments (e.g., video interaction guidance, psychoeducation, or counseling; Barlow et al., 2015), highlighting the important point that many different treatment modalities may achieve similar outcomes. In contrast, a review of more recent studies evaluating PIP found mixed results in terms of the impact of PIP on parent–child interactions and no evidence of changes in attachment security (Cucciniello & Melia, 2024).

The Impact of Perinatal Mental Health Challenges on Parent–Child Interactions

Given the well-established role that parent–child interactions play in the development of secure attachment, it is important to consider how parental mental health challenges may interfere with these interactions. Goodman and Gotlib's model (1999) of mechanisms underlying the intergenerational transmission of maternal depression highlights the impact of maternal depression on interaction qualities such as increased insensitivity, lower levels of stimulation, slower and noncontingent responding, and less frequent vocalization and physical contact. In general, mothers with depression may be at risk for impaired parent–child interactions due to the associations between depression and increased negative parenting behaviors (e.g., intrusiveness; Tronick & Reck, 2009) and decreased positive parenting behaviors (Erickson et al., 2019), including demonstrating less positive and more negative affect (Muzik et al., 2017). Maternal sensitivity has also been found to be compromised in anxious mothers (Feldman et al., 2009) and trauma-exposed mothers, whose interactions are characterized by less emotional availability and greater avoidance, control, intrusiveness, aggression, and hostility (van Ee et al., 2016).

Importantly, infant behavior also contributes to the quality of parent–child interactions and may interact bidirectionally with perinatal mental health problems. For instance, perinatal anxiety has been found to be associated with reduced infant attention, alertness, and social interaction (Brouwers et al., 2001; Feldman et al., 2009; Hernández-Martínez et al., 2008), while perinatally depressed mothers may be more likely to have infants with increased negative emotionality and poorer emotion regulation (Feldman et al., 2009). It has been postulated that mothers' own affective difficulties influence and exacerbate these infant outcomes through difficulties matching infant affect and repairing following mismatches in affect (Tronick & Reck, 2009). Although affective mismatches occur frequently in clinically-referred and non-clinically referred dyads alike, the ability to recover from mismatches may be impacted by the presence of maternal depression, leading to increased emotion regulation difficulties in infants (Tronick & Reck, 2009).

Interventions Targeting Perinatal Mental Health Challenges

Given cautions and contraindications regarding the use of certain psychotropic drugs during pregnancy and while breastfeeding (Sutter-Dallay & Riecher-Rössler, 2016), as well as a reported preference among pregnant and postpartum women favoring therapy over medication as treatment (Dennis & Chung-Lee, 2006; Goodman, 2009; Kim et al., 2011), it is important to study the efficacy of psychotherapeutic treatments. Although pregnant and postpartum individuals face a wide range of psychological concerns during the perinatal period, interventions targeting perinatal depression have been most commonly studied (Nillni et al., 2018).

Multiple systematic reviews and meta-analyses have established the efficacy of Interpersonal Therapy (IPT) and Cognitive Behavioral Therapy (CBT) in reducing symptoms of depression during the perinatal period (Miniati et al., 2014; Sockol, 2015; Sockol et al., 2011), with effect sizes favoring IPT over CBT (Sockol et al., 2011). IPT's emphasis on role transitions and interpersonal disputes, which may be particularly salient themes and sources of distress during the transition to parenthood, may help to account for its superiority over CBT. Individual, as opposed to group-delivered therapies, as well as those delivered postpartum, compared

to during pregnancy, have also been associated with larger effect sizes in terms of mental health outcomes (Sockol, 2015; Sockol et al., 2011), despite the fact that some women may prefer group-delivered therapies due to their ability to provide built-in social support (Dennis & Chung-Lee, 2006).

Effective treatments for perinatal anxiety disorders have largely been cognitive-behavioral in nature (Maguire et al., 2018; Marchesi et al., 2016; Nillni et al., 2018), including the use of specific strategies such as prolonged exposure to phobic stimuli (Lilliecreutz et al., 2010) and mindfulness-based cognitive therapy (Goodman et al., 2014). Similarly, perinatal traumatic stress has been effectively treated using CBT interventions that incorporate the development of a trauma narrative (Shaw et al., 2013). The treatment literature for more severe mental illness (i.e., perinatal bipolar disorder or psychosis) tends to emphasize pharmacotherapy or electroconvulsive therapy over psychotherapy (Batt et al., 2022; Bergink et al., 2016; Khan et al., 2016).

Studies evaluating the impact of treatments for perinatally depressed mothers on relational and child outcomes have demonstrated inconsistent findings and vary in the degree to which they specifically address maternal depressive symptoms (Letourneau et al., 2017; Tsivos et al., 2015b). Although home-based and dyadic interventions targeting parent–child relationships *among* women with depression have demonstrated efficacy at improving relational outcomes (Nylen et al., 2006), relationally-focused interventions (e.g., infant massage) do not show consistent evidence of improvements in maternal depression (Lindensmith, 2018). Likewise, interventions that target depressive symptoms (e.g., CBT and IPT) have demonstrated limited ability to influence infant or parent–child relational outcomes (Forman et al., 2007; Milgrom et al., 2006).

The Present Review

The perinatal period is associated with increased incidence and prevalence of mental health concerns, with some estimates suggesting that rates of perinatal depression and anxiety may reach as high as 15–25% during this period (Howard & Khalifeh, 2020). Not only are these experiences prevalent, but they are also costly; one recent study estimated the economic costs of untreated maternal mental health conditions to be $2.2 billion across the first 5 years postpartum in Texas alone (Margiotta et al., 2022). Estimates have suggested that roughly 60% of the cost of postnatal depression stems from its impact on offspring (Bauer et al., 2016), further underscoring the important point that perinatal mental health conditions have a multi-generational impact. At the same time, this dual-generation impact implies that interventions delivered during the perinatal period offer the unique opportunity to simultaneously treat *and* prevent mental illness.

Despite a well-documented association between maternal mental health challenges and parent–child attachment (e.g., Wan & Green, 2009), the majority of interventions delivered during the perinatal period target *either* maternal mental health difficulties *or* parent–child relationships, failing to integrate these two closely interrelated constructs. The former may be necessary yet insufficient to promote sensitive and attachment-oriented caregiving (Forman et al., 2007), while the latter may fail to address underlying difficulties that interfere with sensitive parenting behavior, while also overlooking the mental health needs and wellbeing of caregivers.

In order to promote an enhanced understanding of how to support secure parent–child attachment in the context of maternal mental health challenges, the aim of the present paper is to systematically review perinatal mental health interventions that directly address maternal mental health *and* parent–infant relational wellbeing. Recognizing the importance of the connection between these constructs, a recently published article (Newton et al., 2022) reviewed the impact of perinatal mental health interventions on relational and infant outcomes. However, though interventions included in their review were delivered to parents experiencing mental health difficulties, there were no explicit criteria requiring included interventions to target parental mental health, nor did authors report on caregiver mental health outcomes. Thus, the present review seeks to fill an important gap in the literature by highlighting effective treatment approaches associated with improvements in maternal mental health *and* relational outcomes.

Methods

All methods were conducted and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for systematic reviews (Page et al., 2021). The review protocol was pre-registered with Prospero (ID: CRD42022380278).

Eligibility Criteria

This review aimed to identify all articles that evaluated the effectiveness of interventions delivered during the perinatal period that seek to promote both maternal mental health and parent–child relational wellbeing, as opposed to one or the other of these targets. To meet this criterion, the mental health-promoting aspects of the intervention were required to be guided by an interventionist (i.e., not self-help) and directive (i.e., not merely supportive in nature). The relational component(s) were required to address the parent–child relationship and interactions directly (i.e., beyond

exploring the impact of the relationship and distorted cognitions related to parenting on maternal mental health).

Inclusion criteria were as follows: (a) study design includes some kind of comparison group, which could also include an active control group targeting mental health *or* relational processes, with randomization not required; (b) mental health and relational outcomes are reported for both the treatment and comparison group(s); (c) at least one relational outcome must be observer-rated or coded (i.e., as opposed to only parent–report); (d) intervention is delivered to mothers during pregnancy and/or the first 18 months postpartum (on average); and (e) article is published in English.

Although randomization is an important marker of study quality, the decision not to require randomization was made in recognition, based on the literature, that dual-generation interventions are relatively understudied. Therefore, authors aimed to "cast a wider net" to be as thorough as possible in identifying relevant studies while maintaining attention to the context of each study's quality. The decision to require at least one observer-rated or coded relational outcome was based on evidence that certain mental health concerns (e.g., depression) may impact a caregiver's representations of their child in a way that biases their perception of relationship quality (Trapolini et al., 2008). Thus, it was deemed important to have at least one more objectively measured outcome for comparison. The intervention time frame (i.e., under 18 months) was intended to capture interventions that are truly perinatal in focus, as opposed to more general or broad. The exclusion of articles published in a language other than English was based on literature suggesting that this practice is not associated with systematic bias (Morrison et al., 2012; Siddaway et al., 2019) and because resources were not available to review articles not in English.

Studies were excluded if any of the following criteria were met: (a) intervention targets *paternal* mental health and parent–child relationships; (b) the primary mental health referral concern is related to a substance use disorder; and (c) intervention is delivered by "peers" (i.e., non-professional interventionists, such as women who previously had postpartum depression).

Although valuable to understand the impact of the mental health of non-birthing parents (e.g., fathers) and other parent–child relationships, this research may be considered still in its infancy and thus was determined to be beyond the scope of the current review. Additional considerations related to the physical effects of perinatal substance use, addiction, detoxification, and withdrawal on both mothers and offspring informed the decision to exclude such studies from the present review (Burns et al., 2016; Forray & Foster, 2015). Finally, although possible that peer-delivered interventions may offer some benefits (e.g., relatability) for perinatal populations, the decision to limit this review to professionally-delivered interventions was driven by recognition of the complex interplay of biological, psychological, and social factors influencing perinatal wellbeing and a desire to increase the likelihood that interventionists are trained to consider and manage this complexity.

Search Strategy and Selection Process

Articles were identified using electronic database searches. Reference lists of included studies were also hand-searched to screen for additional studies possibly meeting criteria. Seven electronic databases were searched from November 21st through November 23rd of 2022: APA PsycArticles, Psychology Database, Sociological and Social Services Abstracts, and Sociology Database (each via ProQuest); PsycInfo; PubMed; and Cochrane Trials Database. Search terms used were: (*perinatal* OR *prenatal* OR *postpartum* OR *antenatal* OR *postnatal*) AND (*depress** OR *anxi** OR *trauma** OR "*mental* ill**" OR "*mental disorder**" OR "*mental health*" OR "*psychopatholog**") AND (*attachment* OR *parent–child* OR *parent–infant* OR *dyad** OR *relation**) AND (*interven** OR *therap** OR *psychotherap**). Where the use of such filters was possible, searches were also filtered based on participant gender (i.e., female; PubMed, PsycInfo) and publication type (i.e., scholarly journals OR dissertation & theses; PsycInfo, all ProQuest searches).

All results, including titles and abstracts, were imported into rayyan.ai (Ouzzani et al., 2016), which is a free, web-based tool designed to aid in the process of study screening for systematic reviews. Although rayyan.ai uses AI features to enhance screening procedures (e.g., identifying possible duplicate studies, suggesting a likelihood that an article meets inclusion criteria, etc.), all articles were thoroughly hand-reviewed by at least two reviewers to ensure accuracy. This tool was used primarily to identify possible duplicates (which were all individually reviewed), organize articles, manage blinded review, and track decisions. A single researcher (SF) was responsible for screening all titles and abstracts of identified articles to assess for possible inclusion. All articles were divided up and additionally screened and reviewed by a second researcher who was blinded to the first researcher's decisions (JR, MJ, ES, DM, or HR). A subset of potentially eligible articles identified through title and abstract screening was subsequently retrieved and uploaded into rayyan.ai to enable full-text review and further analysis of eligibility by at least two reviewers. Any disagreements among reviewers were discussed and resolved such that consensus was reached on a final list of included studies. These same procedures were repeated approximately 1 year later (October 18, 2023) and again prior to publication (March 23, 2025) to identify whether any additional eligible articles had been published during the intervening time.

Data Extraction and Analysis

The primary data of interest sought for synthesis included: descriptions of both mental health- and mother–infant relationship-promoting strategies used in interventions, the specific measures used to compare mental health and relational outcomes between groups and over time, and between- and within-group differences in relevant outcomes of interest post-intervention and across follow-up periods, as available. Other data of interest included information related to the study (e.g., the country where it was conducted, its design, and the type of comparison group used), participants (e.g., sample size; eligibility criteria; and demographic characteristics, including race/ethnicity and infant/fetal age at baseline), and details related to the intervention style, format, and delivery (e.g., individual or group therapy; whether the intervention had a *primary* focus on mental health, relational health, or was balanced; whether mental health and relational components were delivered simultaneously or sequentially; site of delivery; frequency and duration of the intervention; and whether treatment was delivered during pregnancy, postpartum, or both).

In addition to these data, information relevant to assessing risk of bias was extracted using the Critical Appraisal Skills Programme (CASP) Randomized Controlled Trial Checklist as a guide (Critical Appraisal Skills Programme, 2020). This tool was designed to aid in the critical evaluation of study quality based on features of its design, methodology, and results. A fourth category is included in this tool to assist users in analyzing the appropriateness of an intervention for a given population. This section was omitted from the present review to increase generalizability of findings since the population of interest of readers will vary. However, appropriateness of an intervention to a specific setting or context is nonetheless an important consideration that individuals must incorporate when considering implementation.

Following extraction, which was completed by a single researcher (SF), data related to study and participant characteristics, intervention features, and reported outcomes were tabulated by the same researcher to facilitate identification of common features of interventions associated with improvements in outcomes of interest (i.e., maternal mental health and maternal–infant relational wellbeing). Studies were grouped for synthesis based on whether they resulted in improvements in both outcomes, only one outcome, or neither outcome.

Results

Search and Selection Outcome

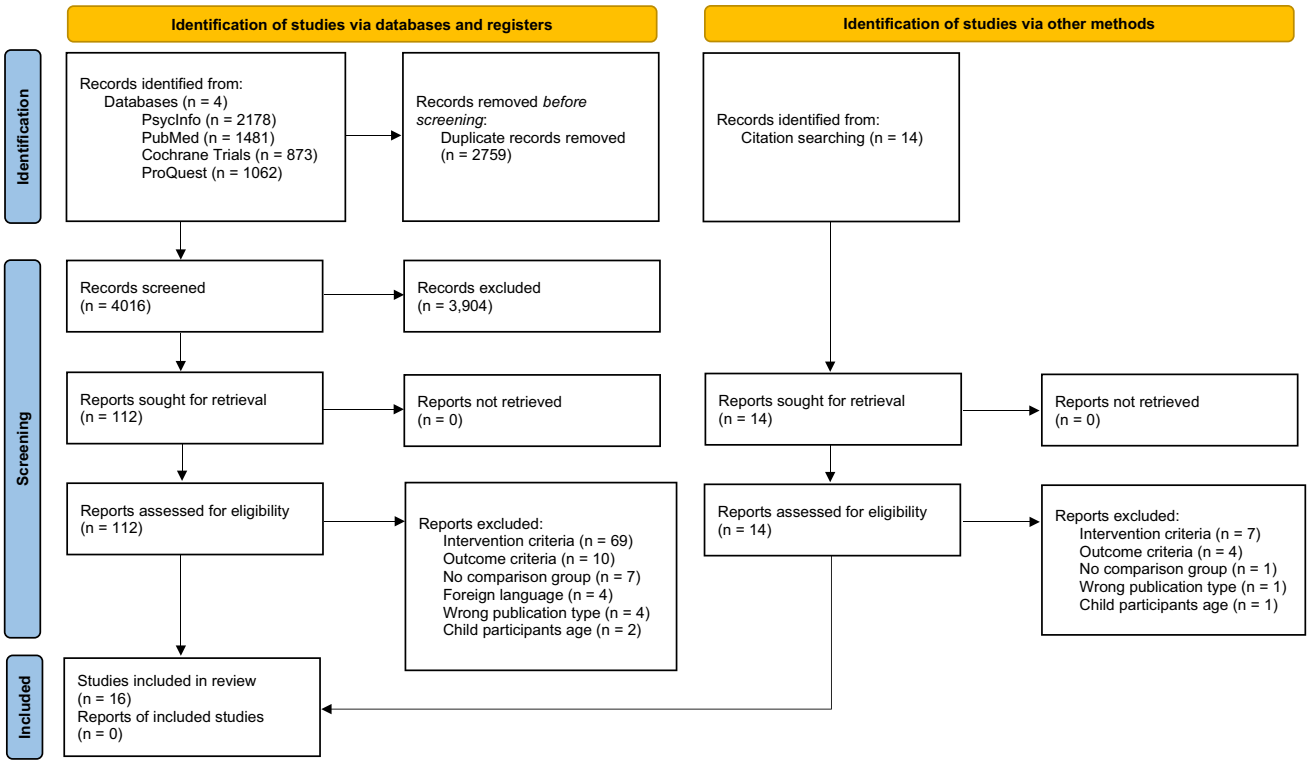
After duplicates were removed ($n = 2759$), titles and abstracts of 4016 unique results identified during the initial review period (November 2022) were screened. One hundred twelve articles were retrieved for full-text review. Disagreements between reviewers ($n = 6$) were discussed until final agreement on included studies ($n = 16$) was reached. Of the 112 articles subject to full-text review, the majority ($n = 69$) were excluded on the basis that intervention criteria were not met, such as inadequate emphasis on or targeting of either mental health or relational problems. The remainder were excluded due to failure to measure or report on outcomes of interest ($n = 10$), lacking a comparison group ($n = 7$), not being written in English ($n = 4$), being the wrong publication type (i.e., oral presentation at a conference; $n = 4$), and being delivered to mothers of older children (i.e., non-infants; $n = 2$).

Fourteen additional articles were retrieved based on a hand-search of reference lists of included studies. However, none of these articles were deemed eligible. See Fig. 1 for a flow diagram outlining the initial full search and selection process. During the second review period (October 2023), 539 additional articles were screened once duplicates ($n = 309$) were removed using identical procedures. Ten articles were retrieved for full-text review and all reviewers independently agreed that none were eligible. See Fig. 2 for a flow diagram outlining the second search and selection process. During the third review period (March 2025), 506 additional articles were screened once duplicates ($n = 327$) were removed using identical procedures. Twenty-three articles were retrieved for full-text review and all reviewers independently agreed that none were eligible. See Fig. 3 for a flow diagram outlining the third and final search and selection process.

Characteristics of Included Studies

A final total of 16 articles reporting outcomes from 14 unique studies and 13 unique interventions were included after screening against selection criteria as described (See Table 1 for an overview of included studies). Sample sizes ranged from 27 to 193, with 1070 total participants involved. Studies were conducted in the United Kingdom (Cooper et al., 2003; Fonagy et al., 2016; Murray et al., 2003; Stein et al., 2006, 2018; Tsivos et al., 2015a), United States (Alhusen et al., 2021; Clark et al., 2003, 2008; Goodman et al., 2015; Lenze et al., 2020), Australia (Ericksen et al., 2018; Holt et al., 2021; Smith et al., 2010), and Canada (Feeley et al., 2012; Zelkowitz et al., 2011) and published between 2003 (Clark et al., 2003; Cooper et al., 2003; Murray et al., 2003) and 2021 (Alhusen et al., 2021; Holt et al., 2021).

In terms of inclusion criteria, the majority of studies required a certain degree of depressive symptomatology for participant eligibility ($n = 10$). Eight of these studies included only participants who actually met criteria for



PRISMA flow diagram outlining review search and selection procedures for articles published through November 2022

some depressive disorder (Clark et al., 2003, 2008; Cooper et al., 2003; Holt et al., 2021; Lenze et al., 2020; Murray et al., 2003; Stein et al., 2018; Tsivos et al., 2015a), and two (Alhusen et al., 2021; Goodman et al., 2015) required elevated depressive symptoms but not a diagnosis. One study required participants to meet diagnostic criteria for an eating disorder for eligibility (bulimia nervosa or other similar eating disorder of clinical severity; Stein et al., 2006). Another study, which was discussed in two articles, recruited mothers of very low birth weight (VLBW) infants (i.e., < 1500 g) regardless of maternal mental health symptomatology, though anxiety symptoms were targeted in the intervention (Feeley et al., 2012; Zelkowitz et al., 2011). The remaining studies (Erickson et al., 2018; Fonagy et al., 2016; Smith et al., 2010) recruited participants based on broader risk categories (i.e., the presence of mental or physical health concerns or some form of socioeconomic risk).

Studies varied in terms of the type of comparison group used, with the majority utilizing a "treatment as usual" (TAU) comparison group (Alhusen et al., 2021; Clark et al., 2003, 2008; Cooper et al., 2003; Ericksen et al., 2018; Fonagy et al., 2016; Goodman et al., 2015; Lenze et al., 2020; Murray et al., 2003; Smith et al., 2010; Tsivos et al., 2015a). However, some TAU groups also included certain enhancements, such as the provision of mental health referrals as needed (Clark et al., 2008; Lenze et al., 2020) or "depression monitoring" by phone (Goodman et al., 2015, p. 493). A subset of these studies utilized a waitlist control (WLC) group so that all participants in the comparison group could receive treatment following evaluation if desired (Clark et al., 2003, 2008; Ericksen et al., 2018).

Other comparison groups included educational attention controls (Feeley et al., 2012; Zelkowitz et al., 2011) and other forms of active treatment, such as IPT (Clark et al., 2003), CBT (Holt et al., 2021; Stein et al., 2006, 2018), psychodynamic therapy (Cooper et al., 2003; Murray et al., 2003), and nondirective counselling (Cooper et al., 2003; Murray et al., 2003). A minority of studies included multiple comparison groups (Clark et al., 2003; Cooper et al., 2003; Murray et al., 2003). The comparison group in three studies featured the same mental health treatment as that which was received by the intervention group (Holt et al., 2021; Stein et al., 2006, 2018). For instance, all participants in Holt et al.' study (2021) received the same CBT for postnatal depression, but the comparison group participated in a nondirective play group instead of the relational play group attended by individuals in the treatment condition. In Stein et al.' study (2006), all women received guided cognitive behavioral self-help, but the comparison group received supportive counseling instead of video-feedback interactional treatment. Participants in Stein et al.' (2018) study all received CBT, and the comparison group received

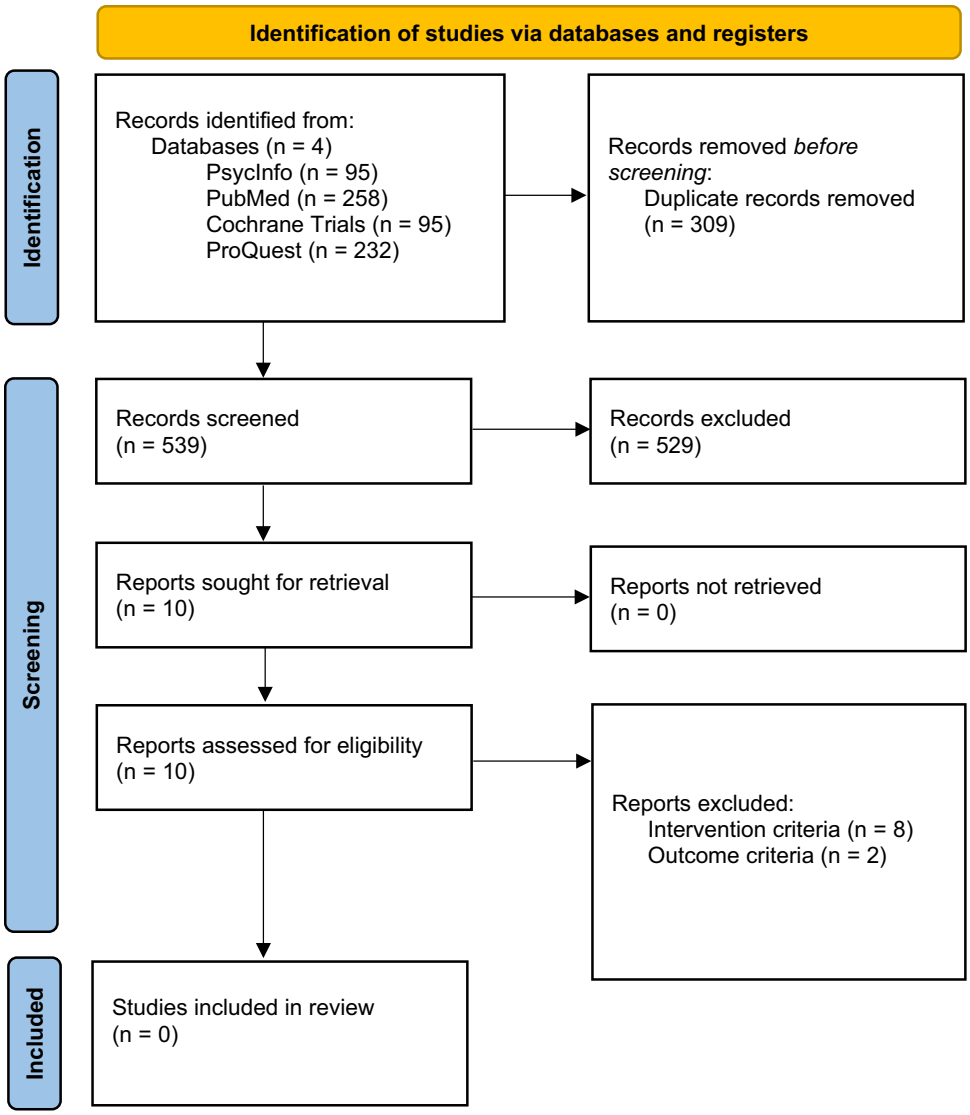


Fig. 2 PRISMA flow diagram outlining review search and selection procedures for articles published November 2022 through October 2023

progressive muscle relaxation therapy instead of video-feedback therapy.

It was difficult to compare studies based on sample demographic characteristics, as these were inconsistently reported. Information regarding race, ethnicity, and nationality of included participants, when reported, can be found in Table 1. It is worth highlighting that the majority of studies were relatively homogeneous and/or reported race and ethnicity in a binary way and relative to a majority group (e.g., White British vs. "Other"; Stein et al., 2018).

Characteristics of Interventions

Given that this review specifically sought interventions with a dual-focus on maternal mental health and mother–infant relational wellbeing, it was interesting to note the extent to which each emphasis was prioritized in the intervention. The majority of interventions primarily emphasized relational wellbeing (Clark et al., 2003, 2008; Ericksen et al., 2018; Fonagy et al., 2016; Goodman et al., 2015; Smith et al., 2010; Stein et al., 2006; Tsivos et al., 2015a) or appeared relatively balanced in emphasis (Cooper et al., 2003; Feeley et al., 2012; Holt et al., 2021; Lenze et al., 2020; Murray et al., 2003; Stein et al., 2018; Zelkowitz et al., 2011). Only one intervention had a primary focus on maternal mental health, with relational aspects, while present, appearing relatively less emphasized (Alhusen et al., 2021). This may have been related to the fact that this was the only intervention conducted solely during pregnancy (Alhusen et al., 2021), making it more difficult to attend to the relationship directly. Only one other intervention began during pregnancy, though the majority of sessions (i.e., 10 versus eight) were delivered postpartum (Lenze et al., 2020). The remainder were delivered solely postpartum (see Table 1 for average infant/fetal age at baseline).

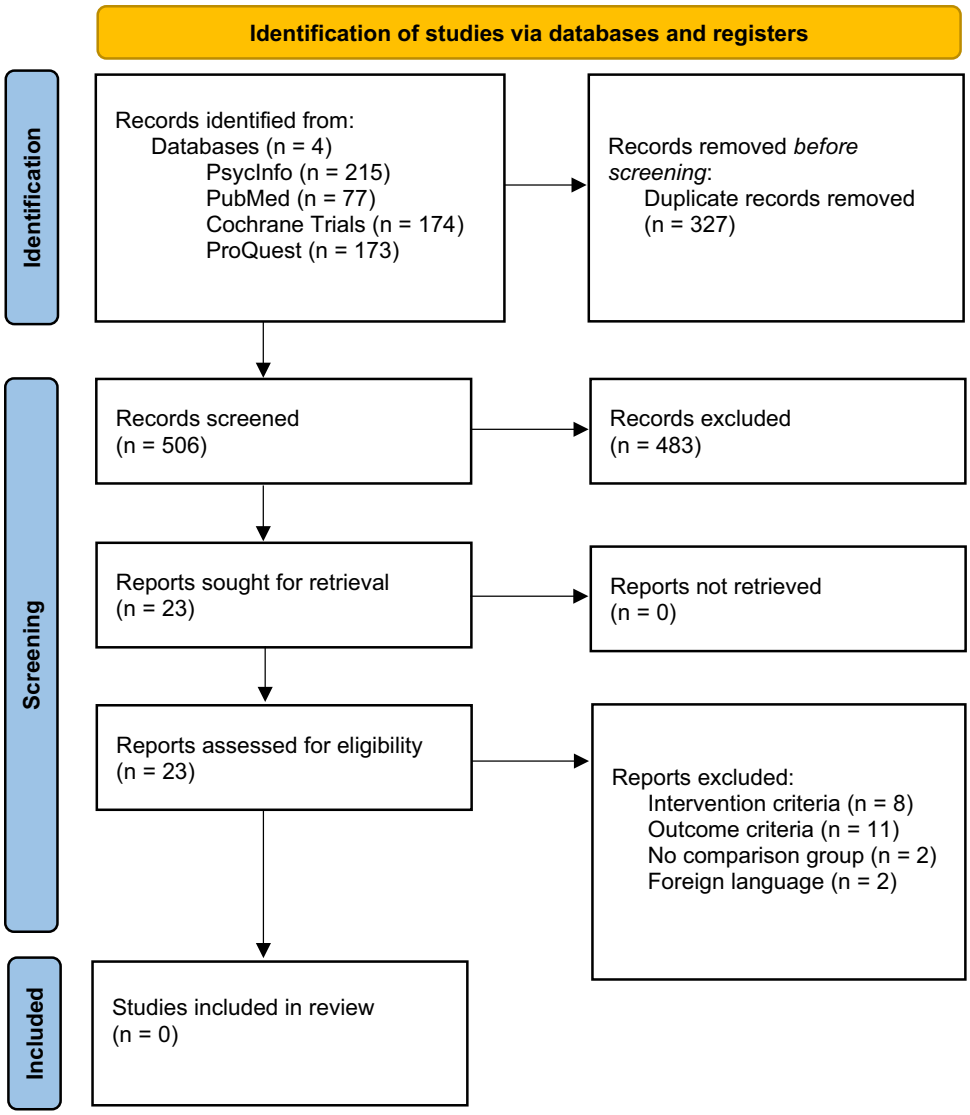


Fig. 3 PRISMA flow diagram outlining review search and selection procedures for articles published October 2023 through March 2025

The majority of interventions appeared to target mental health and relational wellbeing more or less simultaneously (i.e., focusing on both domains within each individual session), with only a subset (Feely et al., 2012; Holt et al., 2021; Zelkowitz et al., 2011) targeting mental health exclusively prior to addressing relational wellbeing. Intervention formats also appeared to favor individual therapy (Cooper et al., 2003; Feeley et al., 2012; Fonagy et al., 2016; Goodman et al., 2015; Lenze et al., 2020; Murray et al., 2003; Stein et al., 2006, 2018; Tsivos et al., 2015a; Zelkowitz et al., 2011) over group therapy (Alhusen et al., 2021; Clark et al., 2003, 2008; Ericksen et al.,; 2018; Holt et al., 2021; Smith et al., 2010), as well as shorter duration (i.e., 12 or fewer sessions; Alhusen et al., 2021; Clark et al., 2003, 2008; Cooper et al., 2003; Ericksen et al., 2018; Feeley et al., 2012; Goodman et al., 2015; Murray et al., 2003; Smith et al., 2010; Tsivos et al., 2015a; Zelkowitz et al., 2011) over longer duration (i.e., 13 or more; Fonagy et al., 2016; Holt et al., 2021; Lenze et al., 2020; Stein et al., 2006, 2018).

Quality Assessment

The quality of included studies varied. All but three studies (Clark et al., 2003, 2008; Smith et al., 2010) utilized a randomized controlled trial (RCT) design. Many studies were designated as pilot studies (Alhusen et al., 2021; Clark et al., 2008; Ericksen et al., 2018; Goodman et al., 2015; Lenze et al., 2020; Tsivos et al., 2015a) and were accordingly relatively small in scale [i.e., of the pilot studies, only Alhusen et al. (2021) had a sample size exceeding 45]. Many studies were self-described as underpowered (Alhusen et al., 2021; Ericksen et al., 2018; Holt et al., 2021; Lenze et al., 2020; Tsivos et al., 2015a) or struggled with high attrition rates (Ericksen et al., 2018; Fonagy et al., 2016; Holt et al., 2021;

Table 1 Characteristics and relevant results of included studies

Author and study information	Participant details	Intervention details	Outcome measures and time points	Results
Alhusen et al. (2021) USA Randomized controlled trial (RCT; pilot) Comparison group: Treatment as Usual (TAU)	$n = 30$ (MB); $n = 30$ (TAU) Inclusion criteria: Moderate to severe depressive symptoms (Edinburgh Postnatal Depression Scale [EPDS]) 90% (treatment and comparison) African American, 10% (treatment and comparison) White Fetal age: not reported, recruited 9–12 weeks gestation	*Mothers and Babies Course (MB)* Key mental health components Cognitive strategies (i.e., controlling negative thoughts) Behavioral activation Key relational components Psychoeducation on attachment, bonding, sensitive caregiving Sensitivity skill-building Style (emphasis/sequence, format, timing) Primarily mental health, simultaneous Group therapy Pregnancy only Delivery (site, frequency/duration) Prenatal clinics 6 sessions (120 min, weekly)	*Mental health* EPDS—pre (22–32 weeks gestation), 36 weeks gestation, 12 weeks postpartum *Relational* Maternal Fetal Attachment Scale (MFAS)—pre (22–32 weeks gestation), 36 weeks gestation Nursing Child Assessment Satellite Training Parent Child Interaction (NCAST)—12 weeks postpartum	EPDS: no evidence of differences MFAS: no evidence of differences NCAST: no evidence of differences Note: study deemed insufficiently powered for significance testing and estimating effect sizes. Authors compared the size of changes in average scores

Table 1 (continued)

Author and study information	Participant details	Intervention details	Outcome measures and time points	Results
Clark et al. (2003) USA Nonrandomized waitlist controlled trial Comparison groups: waitlist control (WLC) and Interpersonal Psychotherapy (IPT; IPT group added later)	$n = 13$ (M-ITG); $n = 15$ (IPT); $n = 11$ (WLC) Inclusion criteria: Depressive disorder diagnosis (questionnaire/interview not specified) 100% White (treatment), 93% White and 7% African American (active comparison), 100% White (waitlist comparison) Infant age (months): 6.3 (treatment), 8.2 (active comparison), 12.8 (waitlist comparison)	*Mother–Infant Therapy Group (M-ITG)* Key mental health components Cognitive behavioral: explore impact of cognitions and modify, increased expression and validation of emotion IPT to address interpersonal problems and increase social support and effective communication Key relational components Address internal working models (IWMs) and differentiate from family of origin Guided parent–child interactions promoting emotional availability and affective attunement Therapist "speaks for the infant" Promoting curiosity and reflective functioning by observing and wondering about infant cues Style (emphasis/sequence, format, timing) Primarily relational, simultaneous Group therapy Postpartum Delivery (site, frequency/duration) Site of delivery not specified 12 sessions (weekly)	*Mental health* Beck Depression Inventory (BDI)—pre/post Center for Epidemiologic Studies Depression Scale (CES-D)—pre/post *Relational* Parenting Stress Index (PSI)—pre/post Parent–Child Early Relational Assessment (PCERA)—pre/post	CES-D: M-ITG and IPT significantly lower than WLC, not different from each other BDI: not significant (n.s.) PSI: M-ITG and IPT report child is significantly more reinforcing and adaptable than WLC, not different from each other (otherwise n.s.) PCERA: M-ITG and IPT significantly higher maternal positive affective involvement and verbalization, not different from each other Note: based on ANCOVA with pretreatment scores as covariates
Clark et al. (2008) USA Nonrandomized WLC trial (pilot) Comparison group: WLC (referrals made as needed)	$n = 18$ (M-ITG); $n = 14$ (WLC) Inclusion criteria: Depressive disorder diagnosis (questionnaire/interview not specified) "ethnically homogeneous" Infant age (months): 7.9 (treatment) and 11.7 (comparison)	See Clark et al. (2003)	*Mental health* BDI—pre/post *Relational* PSI—pre/post PCERA—pre/post	BDI: M-ITG reported fewer depressive symptoms than WLC, Effect Size (ES) = 1.126 PSI: M-ITG reported child is more reinforcing than WLC, ES = 1.32 (otherwise n.s.) PCERA: M-ITG significantly higher maternal positive affective involvement and verbalization than WLC, ES = 1.084 Note: based on ANCOVA with pretreatment scores as covariates

Table 1 (continued)

Author and study information	Participant details	Intervention details	Outcome measures and time points	Results
Cooper et al. (2003) UK RCT Comparison groups: TAU, psychodynamic therapy, non-directive counselling	$n=43$ (r-CBT); $n=50$ (psychodynamic); $n=48$ (non-directive counselling); $n=52$ (TAU) Inclusion criteria: Major depressive disorder diagnosis (questionnaire/interview not specified) Race/ethnicity not reported Infant age: not reported, recruited between 8 and 18 weeks postpartum	*Relationally-Focused Cognitive Behavioral Therapy (r-CBT)* Key mental health components CBT-oriented, with an emphasis on patterns of thinking regarding parenting and infant and systematic problem-solving (not described in detail) Key relational components Reported under Murray et al. (2003) Style (emphasis/sequence, format, timing) Balanced, simultaneous Individual therapy Postpartum Delivery (site, frequency/duration) Home-based 10 sessions (weekly, 8–18 weeks postpartum)	*Mental health* EPDS—pre, post, 9 month follow-up (f/u), 18 month f/u, 5 year f/u Structured Clinical Interview for the DSM-III-R (SCID)—pre, post, 9 month follow-up (f/u), 18 month f/u, 5 year f/u *Relational* Reported under Murray et al. (2003)	EPDS: All treatment groups had significantly lower EPDS than control group at post. No significant differences at 9 month, 18 month, or 5 year f/u SCID: Psychodynamic therapy group significantly more effective in reducing rates of depression (71% remitted) than controls (40%). CBT (57%) and non-directive counselling (54%) no different than controls. No significant differences at 9 month, 18 month, or 5 year f/u Note: EPDS based on comparison of mean scores, SCID based on random effects logistic model
Ericksen et al. (2018) Australia RCT (pilot) Comparison group: WLC	$n=16$ (CHUGS); $n=15$ (WLC) Inclusion criteria: Mental health problems based on recent consultation with health professional Race/ethnicity not reported Infant age (months): 4.9 (treatment and comparison)	*Community HUGS (CHUGS)* Key mental health components CBT-oriented Anxiety/stress management Self-care Discussion of family of origin Key relational components Increasing sensitivity and reflective functioning through guided play, dance/movement Psychoeducation on attachment, temperament Style (emphasis/sequence, format, timing) Primarily relational, simultaneous Group therapy Postpartum Delivery (site, frequency/duration) Public hospital and community settings 10 sessions (60–90 min, weekly)	*Mental health* The Depression, Anxiety, and Stress Scale (DASS)—pre, post, 6 month f/u *Relational* PSI—pre, post Pediatric Infant Parent Exam (PIPE)—pre, post	DASS: CHUGS significantly lower depression, anxiety, and stress than WLC at post. 6 month f/u not reported for WLC, but noted that CHUGS improvements were sustained at f/u PSI: CHUGS significantly higher difficult child subscale (otherwise n.s.) PIPE: Neither CHUGS nor WLC showed changes in observer-rated parent–child interaction quality from pre to post Note: Generalized estimating equations accounting for baseline scores used for between-group comparisons. Dependent samples t-tests for PIPE, limited data available ($n=7$ CHUGS, $n=5$ WLC)

Table 1 (continued)

Author and study information	Participant details	Intervention details	Outcome measures and time points	Results
Feeley et al. (2012) Canada RCT Comparison group: attention control (education on infant care)	$n = 48$ (Cues); $n = 50$ (control; subset from Zelkowitz et al. (2011) available for longer follow-up) Inclusion criteria: Very low birth weight (< 1500 g) infants 70% (intervention) and 58% (comparison) born in North America Infant age (gestational age): 28 weeks (treatment) and 28.2 weeks (comparison); 33 days at start of intervention	*Cues* Key mental health components Anxiety management—psychoeducation, learn and practice relaxation strategies (deep breathing, imagery), challenging negative automatic thoughts Key relational components Psychoeducation on interacting with VLBW infants, reading cues Practice observing cues Guided sensitive interactions during feeding Style (emphasis/sequence, format, timing) Balanced, sequential Individual therapy Postpartum Delivery (site, frequency/duration) Neonatal intensive care unit (5 sessions), home (1 session) 6 sessions (60 min; 5 weekly, last session 2–4 weeks post-discharge)	*Mental health* State-Trait Anxiety Inventory (STAI; state portion)—6 month f/u Perinatal PTSD Questionnaire (PPQ)—6 month f/u EPDS—6 month f/u *Relational* Global Rating Scales of Mother–Infant Interaction (GRS)—6 month f/u	No significant between-group differences observed at 6 month f/u on any mental health or relational measures Note: unpaired t-tests used for between-group comparisons at post

Table 1 (continued)

Author and study information	Participant details	Intervention details	Outcome measures and time points	Results
Fonagy et al. (2016) UK RCT Comparison group: TAU	$n = 38$ (PIP); $n = 38$ (TAU) Inclusion criteria: Mental health problems reaching "probable psychiatric case" criteria based on GHQ and 1 + indicator of social exclusion 68% (treatment) and 58% (comparison) White Infant age (months): 3.9 (treatment) and 3.8 (comparison)	*Parent–Infant Psychotherapy (PIP)* Key mental health components Caregivers encouraged to express difficult feelings and thoughts Addressed emotional and behavioral difficulties (not described in detail) Key relational components Addressed IWMs based on past experiences Support for positive parent–child interactions Addressed "communication errors" associated with disorganized attachment—support for appropriate responding to infant signals Style (emphasis/sequence, format, timing,) Primarily relational, simultaneous Individual therapy Postpartum Delivery (site, frequency/duration) Hospitals, community center Weekly/every other week Variable duration ($M = 16$ sessions, range = 1–49)	*Mental health* CES-D—pre, 6 month f/u, 12 month f/u Brief Symptom Inventory (BSI)—pre, 6 month f/u, 12 month f/u *Relational* PSI—pre, 6 month f/u, 12 month f/u Mother's Object Relations Scales (MORS)—6 month f/u, 12 month f/u Parent Development Interview (PDI)—Parental Reflective Functioning (RF) and Assessment of Representational Risk (ARR)—pre, 12 month f/u Emotional Availability Scales (EAS)—pre, 6 month f/u, 12 month f/u Coding Interactive Behavior system (CIB)—pre, 6 month f/u, 12 month f/u SSP—12 month f/u	CES-D: PIP significantly greater improvements in depression BSI: n.s. PSI: PIP significantly less parenting stress over time, less reported perceived relational difficulties MORS: n.s. PDI: PIP significant reduction in coded representational risk, RF n.s. EAS/CIB: No significant treatment group effects for observed sensitivity or parent–child interactions SSP: n.s. Note: Multilevel mixed-effects linear regression used to examine treatment differences and changes over time. Chi-squared test used to analyze SSP differences

Table 1 (continued)

Author and study information	Participant details	Intervention details	Outcome measures and time points	Results
Goodman et al. (2015) USA RCT (pilot) Comparison group: TAU + depression monitoring	$n = 21$ (PDP); $n = 21$ (TAU) Inclusion criteria: 2 consecutive elevated postpartum depression scores 1 week apart (EPDS) 57.1% (treatment) and 61.9% (comparison) non-Hispanic White Infant age: not reported, recruited between 4 and 6 weeks postpartum	*Perinatal Dyadic Psychotherapy (PDP)* Key mental health components Mood monitoring, self-care, recognizing and modifying unhelpful thinking Increase social support Key relational components Promoting observation/curiosity and perspective-taking about infant behavior Encouraging appropriate interactions, including sensitivity to cues Exploration of IWMs Style (emphasis/sequence, format, timing) Primarily relational, simultaneous Individual therapy Postpartum Delivery (site, frequency/duration) Home-based 8 sessions (60 min; 4 weekly, 4 every other week)	*Mental health* SCID (depression and anxiety)—pre, post, 3 month f/u EPDS—pre, post, 3 month f/u STAI—pre, post, 3 month f/u *Relational* PSI—post, 3 month f/u CIB—post, 3 month f/u	No significant between-group differences observed on any mental health or relational measures Note: mixed effect regression analyses used to evaluate between-group differences at each time point

Table 1 (continued)

Author and study information	Participant details	Intervention details	Outcome measures and time points	Results
Holt et al. (2021) Australia RCT Comparison group: CBT + nondirective play group	$n = 38$ (HUGS); $n = 39$ (control) Inclusion criteria: Depressive disorder diagnosis (major or minor; SCID) 68% (treatment) 67% (comparison) born in Australia Infant age (months): 3.1 (treatment) and 4.0 (comparison)	*Cognitive Behavioral Therapy + Happiness, Understanding, Giving, and Sharing (CBT + HUGS Playgroup)* Key mental health components CBT for postnatal depression, including partner sessions Key relational components Psychoeducation on infant cues Interaction coaching to promote responsiveness Challenge distorted cognitions regarding IWMs Style (emphasis/sequence, format, timing) Balanced, sequential Group therapy (both) Postpartum Delivery (site, frequency/duration) Hospital or community center 12 CBT sessions (60–90 min, weekly; 9 group + 3 couple sessions), 4 HUGS sessions (60–90 min, weekly)	*Mental health* BDI—pre, post CBT, post HUGS, 6 month f/u Beck Anxiety Inventory (BAI)—pre, post CBT, post HUGS, 6 month f/u *Relational* PSI—pre, post CBT, post HUGS, 6 month f/u PCERA—pre, post HUGS, 6 month f/u Postpartum Bonding Questionnaire (PBQ)—pre, post CBT, post HUGS, 6 month f/u	BAI/BSI: n.s. PSI: n.s., PCERA: HUGS significantly higher maternal positive affective involvement and verbalization at 6 month f/u, n.s., at post PBQ: HUGS significantly lower in self-reported impaired bonding at 6 month f/u, n.s., at post Note: based on ANCOVA with pretreatment scores as covariates (PSI, PCERA, PBQ) and mixed between-within subjects ANOVA (BAI, BDI)

Table 1 (continued)

Author and study information	Participant details	Intervention details	Outcome measures and time points	Results
Lenze et al. (2020) USA RCT (pilot) Comparison group: Enhanced TAU (encouraged to engage in mental health services as needed, assistance obtaining community providers given)	$n = 21$ (IPT-Dyad); $n = 21$ (Enhanced TAU) Inclusion criteria: Depressive disorder diagnosis (MDD, Dysthymia, Depression NOS; SCID) 78.5% African American (not reported by group) Fetal age: not reported, recruited 12–30 weeks gestation	*Interpersonal Psychotherapy for the Mother–Infant Dyad (IPT-Dyad)* Key mental health components IPT Key relational components Interactive activities to explore relationship, promote sensitivity/responsiveness Modeling and imitating to promote communication, speaking for the infant Exploration and resolution of distorted IWMs Style (emphasis/sequence, format, timing) Balanced, simultaneous Individual therapy Pregnancy and postpartum Delivery (site, frequency/duration) Home or clinic or community 8 sessions (weekly) during pregnancy, up to 10 (4 weekly, then flexible) sessions postpartum	*Mental health* EPDS—pre, 37–39 weeks gestation, 3, 6, 9, 12 months f/u STAI (brief)—pre, 37–39 weeks gestation, 3, 6, 9, 12 months f/u *Relational* PSI—6, 12 month f/u CIB—3, 6, 9, 12 month f/u	EPDS/STAI: no differences between groups, significant decreases within both PSI: no differences between or within groups CIB: no differences between groups. Significant *decrease* in maternal sensitivity within IPT-Dyad group over time Note: study deemed insufficiently powered for between-group significance testing. Used linear mixed models to assess within-group differences over time
Murray et al. (2003) UK (see Cooper et al., 2003)	see Cooper et al. (2003)	*Cognitive Behavioral Therapy + Interaction Guidance Treatment* Key mental health components Reported under Cooper et al. (2003) Key relational components Modified interactive guidance (e.g., McDonough, 1993): review of videotaped parent–infant interactions (not described in detail) Soliciting reactions/reflections from caregiver, highlighting positive parenting/sensitivity, encouraging alternative behaviors when needed Style (emphasis/sequence, format, timing) Reported under Cooper et al. (2003) Delivery (site, frequency/duration) Reported under Cooper et al. (2003)	*Mental health* Reported under Cooper et al. (2003) *Relational* Problems in mother–infant relationship checklist (unnamed)—pre/post Global Rating Scales of Mother–Infant Interaction (GRS)—pre/post Strange Situation Procedure (SSP)—18 month f/u	Self-report checklist: CBT lowest relative risk of reporting relationship problems; all treatment groups significantly lower than control GRS: comparable increases across all conditions (n.s.) SSP: differences between groups all n.s. Note: Checklist based on relative risk of relationship problems, controlling for baseline; interactions based on mean differences pre/post; SSP based on relative risk of insecure attachment by group

Table 1 (continued)

Author and study information	Participant details	Intervention details	Outcome measures and time points	Results
Smith et al. (2010) Australia Pre- and post-group repeated measures design Comparison group: TAU (added later)	$n = 74$ (PAIRS); $n = 32$ (TAU) Inclusion criteria: General socioeconomic/medical/ mental health risk, though targeting depression Race/ethnicity not reported Infant age (months): 9 (treatment) and 9.9 (comparison)	*Parent and Infant Relationship Support (PAIRS)* Key mental health components Exploring traumas associated with motherhood, emotions, feelings of ambivalence/loss, joint problem-solving (not described in detail) Key relational components Psychoeducation on attachment, in vivo support for separation and reunion Modeling, interactive coaching Promoting reflective functioning (watch, wait, wonder) Style (emphasis/sequence, format, timing) Primarily relational, simultaneous Group therapy Postpartum Delivery (site, frequency/duration) Community clinics 10 sessions (120 min, weekly)	*Mental health* EPDS—pre, post *Relational* Dyadic Mutuality Code (DMC)—pre, post	EPDS: PAIRS significant decrease in depression. Control n.s. DM: PAIRS significant increase in observed mother–infant interaction quality. Control n.s. Note: statistical analyses not described. Significance of pre/post differences within groups described

Table 1 (continued)

Author and study information	Participant details	Intervention details	Outcome measures and time points	Results
Stein et al. (2006) UK RCT Comparison group: supportive counseling + guided cognitive-behavioral self-help	$n = 40$ (VF + CBT); $n = 40$ (control) Inclusion criteria: Eating disorder diagnosis (bulimia nervosa or similar; clinical interview) 70% (treatment and comparison) White Infant age: not reported, recruited between 4 and 6 months	*Video-Feedback Interactional Treatment + Guided Cognitive Behavioral Self-Help (VF + CBT)* Key mental health components Guided cognitive behavioral self-help Psychoeducation regarding eating problems Strategies to regain control over eating, reduce purging behaviors, address cognitions regarding shape/weight Key relational components Review of videotaped interactions during mealtime to promote identification of and responsiveness to infant cues Style (emphasis/sequence, format, timing) Primarily relational, simultaneous Individual therapy Postpartum Delivery (site, frequency/duration) Home-based 13 sessions (60 min, delivered over 6–8 months)	*Mental health* EPDS—pre, post Eating Disorder Examination Interview (EDE)—pre, post *Relational* Video-taped mother–infant interaction to rate conflict/harmony (unnamed)—pre/post	EPDS/EDE: n.s. Observed parent–child interactions: VF + CBT significantly more harmony, less severe conflict, greater maternal facilitation of interactions, more infant autonomy, and more appropriate nonverbal responses to infant cues. Appropriate verbal responses to infant cues n.s. Note: two-sample t-tests and Wilcoxon rank-sum tests used to compare treatment groups

Table 1 (continued)

Author and study information	Participant details	Intervention details	Outcome measures and time points	Results
Stein et al. (2018) UK RCT Comparison group: Cognitive Behavioral Therapy + Progressive Muscle Relaxation (CBT + PMR)	$n = 72$ (CBT + VFT); $n = 72$ (CBT + PMR) Inclusion criteria: Persistent postnatal MDD (SCID) 85% (treatment) and 81% (comparison) White British Infant age: 56.9% (treatment) and 59.7% (comparison) between 4.5 and 7 months; remainder between 7 and 9 months	*Cognitive Behavior Therapy + Video-Feedback Therapy (CBT + VFT)* Key mental health components Cognitive strategies Behavioral activation Self-care, activating social supports Key relational components Targeting parent behaviors using video feedback—contingent responding to cues, emotional scaffolding, and sensitivity/treating child as a psychological agent Style (emphasis/sequence, format, timing) Balanced, simultaneous Individual therapy Postpartum Delivery (site, frequency/duration) Primarily home-based (children's center if requested) 13 sessions (60–90 min; 6 weekly, 5 every other week, 2 post-treatment boosters)	*Mental health* EPDS—pre, post, 12 month f/u SCID (depression)—pre, post, 12 month f/u *Relational* Attachment Q-sort (AQS)—12 month f/u Videotaped mother–child interaction (unnamed)—pre, post, 12 month f/u	EPDS: n.s. at post. CBT + PMR group significantly lower depression at 12 month f/u SCID: n.s. Observed parent–child interactions: between group differences n.s. for following child's attention, contingent responsiveness, sensitivity, or warmth, but warmth improved significantly in VFT group but not PMR AQS: n.s. Note: linear mixed-effects models used to analyze group differences EPDS/SCID. Mann–Whitney *U* tests used for between group differences in parent–child interactions. ANCOVA with pretreatment depression and demographic variables as covariates used for AQS

Table 1 (continued)

Author and study information	Participant details	Intervention details	Outcome measures and time points	Results
Tsivos et al. (2015a) UK RCT (pilot) Comparison group: TAU	$n = 14$ (Baby Triple P); $n = 13$ (TAU) Inclusion criteria: Depression diagnosis (SCID) + risk of postpartum depression (EPDS) 100% (treatment) and 53% (comparison) "British" Infant age (months): 6.7 (treatment) and 5.7 (comparison)	*Baby Triple P (Positive Parenting Program)* Key mental health components Relaxation/stress management Identifying and challenging unhelpful thoughts/thinking traps, using coping statements and positive self-talk Activating social support, increasing interpersonal skills Emotion exploration Key relational components Teaching positive parenting skills In vivo practice of responsive/sensitive caregiving Style (emphasis/sequence, format, timing) Primarily relational, simultaneous Individual therapy Postpartum Delivery (site, frequency/duration) Home-based 8 sessions (60–90 min, weekly)	*Mental health* BDI—pre, post, 3 month f/u *Relational* PBQ—pre, post, 3 month f/u CARE index—pre, post, 3 month f/u	BDI: n.s. PBQ: n.s. CARE Index (maternal sensitivity, controlling behavior, unresponsiveness; infant cooperativeness, difficult behavior, compulsivity, passivity): n.s. Note: based on ANCOVA with pretreatment scores as covariates or Mann–Whitney *U* analysis
Zelkowitz et al. (2011) (see Feeley et al., 2012)	$n = 60$ (Cues); $n = 61$ (control) Inclusion criteria: Very low birth weight (< 1500 g) infants 68.3% (intervention) and 44.3% (comparison) born in North America	See Feeley et al. (2012)	*Mental health* State-Trait Anxiety Inventory (STAI; state portion)—pre/post Parental Stress Scale: Neonatal Intensive Care Unit (PSS:NICU)—pre/post Perinatal PTSD Questionnaire (PPQ)—pre/post EPDS—pre/post *Relational* Global Rating Scales of Mother–Infant Interaction (GRS)—pre/post	No significant between-group differences observed at post-evaluation on any mental health or relational measures Note: unpaired t-tests used to compare group scores at post

Table 2 Summary of risk of bias analyses

References	Study design			Methodology			Results		
	Focused question	Randomization	Participants accounted for	Blinding	Baseline similarities[a]	Equal treatment	Comprehensive reporting	Precise reporting	Benefits[b]
Alhusen et al. (2021)	+	+	+	/	/	+	−	−	/
Clark et al. (2003)	+	−	−	/	−	+	−	−	+
Clark et al. (2008)	+	−	−	/	−	+	/	−	+
Cooper et al. (2003) and Murray et al. (2003)	+	+	+	/	−	+	+	+	+
Ericksen et al. (2018)	+	+	/	−	/	+	−	−	+
Feeley et al. (2012) and Zelkowitz et al. (2011)	+	+	+	/	/	+	+	+	−
Fonagy et al. (2016)	+	+	+	/	−	+	/	+	+
Goodman et al. (2015)	+	+	+	−	−	+	+	+	−
Holt et al. (2021)	+	+	+	/	/	+	/	−	+
Lenze et al. (2020)	+	+	+	−	+	+	−	−	−
Smith et al. (2010)	/	−	−	/	/	+	−	−	+
Stein et al. (2006)	+	+	+	/	/	+	+	/	+
Stein et al. (2018)	+	+	+	/	/	+	+	+	+
Tsivos et al. (2015a)	+	+	+	−	+	+	+	/	−

Risk of bias analysis based on modified Critical Appraisal Skills Programme (CASP) Randomised Controlled Trial Checklist (Critical Appraisal Skills Programme, 2020). A plus sign (+) denotes low risk of bias, a slash (/) denotes questionable risk of bias, and a minus sign (−) denotes definite risk of bias

[a]See De Boer et al. (2015) for discussion of the relevance of baseline demographic differences in RCTs; this may not be a valid marker of low quality/bias

[b]Treatments were considered beneficial (+) if statistically significant benefits were reported on at least one outcome of interest compared to at least one comparison group. Questionable benefit (/) denotes descriptive differences favoring the intervention in the absence of statistically significant differences. Lack of benefit (−) denotes lack of evidence of intervention superiority on any outcome

Lenze et al., 2020). See Table 2 for an overview of study quality grouped by study design, methodology, and results.

Study Outcomes

Since interventions were included if they addressed both maternal mental health and maternal–infant relational well-being, the most successful interventions may be considered those that resulted in improvements in both of these outcomes statistically significantly beyond those that were observed in a comparison group, as, although meaningful, clinical significance was not addressed in the majority of articles. This criterion was considered to have been met on the basis of either statistically significant between-group differences following intervention, or statistically significant within-group improvements in the treatment condition in the absence of significant improvements in the comparison condition. Outcomes are presented grouped by effectiveness, starting with those providing evidence of improvements in both outcomes, followed by improvements in mental health outcomes only, and then relational outcomes only. Interventions lacking evidence of statistically significant improvements in either outcome will be discussed last.

Improvements in Both

Three interventions evaluated in four different studies were associated with improvements in both types of outcomes: Mother–Infant Therapy Group (M-ITG; Clark et al., 2003, 2008), Parent–Infant Psychotherapy (PIP; Fonagy et al., 2016), and Parent and Infant Relationship Support (PAIRS; Smith et al., 2010).

In 2003, Clark et al. published an evaluation of M-ITG compared to a waitlist control (WLC) group and an IPT comparison group (the latter having been added *after* initial data on M-ITG and WLC groups were collected). Recipients of M-ITG were found to have significantly lower depression scores on the Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977), which assesses symptoms of depression in the general population, than the WLC group following intervention. The severity of depression scores, as measured by the Beck Depression Inventory (BDI; Beck et al., 1961), was not

significantly different across groups post-intervention. Parent–child relational outcomes were measured using a self-report measure (the Parenting Stress Index [PSI]; Abidin et al., 2006) and an observer-rated measure (Parent–Child Early Relational Assessment [PCERA]; Clark, 1985). Individuals in the M-ITG group significantly outperformed the WLC group post-intervention in terms of observed maternal positive affective involvement and verbalization during interactions, as well as mother-rated reinforcement value of the child. However, M-ITG failed to outperform IPT in either of these domains (i.e., mental health or relational; Clark et al., 2003). Clark et al. published an additional evaluation of M-ITG in 2008, which utilized only a WLC group. Results were similar in terms of relational measures and outcomes. Depression was measured using the BDI only (Beck et al., 1961), and M-ITG significantly outperformed the WLC group in terms of improvements in depression symptoms.

PIP was compared to a Treatment As Usual (TAU) group by Fonagy et al. (2016). PIP was associated with significantly greater reductions in depression, as measured by the CES-D (Radloff, 1977), but no significant differences were observed in terms of overall psychological symptoms, measured using the Brief Symptom Inventory (BSI; Derogatis & Melisaratos, 1983). Relational outcomes were assessed using several measures. PIP participants self-reported lower parenting stress levels and perceived relational difficulties on the PSI (Abidin et al., 2006) compared to the control group. The Parent Development Interview (Slade et al., 2004) was used to code parental reflective functioning and representational risk. PIP treated individuals demonstrated lower representational risk following intervention compared to TAU (i.e., interview themes were coded as less hostile and helpless in terms of representations of the child and parenting; Fonagy et al., 2016). Self-reported maternal representations and observational measures of parent–child attachment and interactions were not significantly different between groups (see Table 1 for details).

Finally, PAIRS was evaluated less formally over the course of nearly 10 years of groups being delivered across Australia, with the same outcome measures being collected from a smaller group of individuals with similar referral concerns but who did not receive the intervention (Smith et al., 2010). Control group pre- and post-data were collected about 10–12 weeks apart to approximate the length of treatment. Although between-group significance testing was not conducted, PAIRS-treated individuals experienced significant reductions in depression (measured by the Edinburgh Postnatal Depression Scale [EPDS]; Cox et al., 1987) and significant increases in observed mother–infant interaction quality (measured using Dyadic Mutuality Codes [DMC]; Censullo, 1991), while no significant changes were observed among TAU participants.

In terms of study quality, it is important to note that three of these four studies (Clark et al., 2003, 2008; Smith et al., 2010) were nonrandomized. In fact, these were the only three studies among those that qualified for inclusion in the present review that did not utilize a randomized design. This consideration alone greatly compromises the confidence with which these interventions can be concluded to be effective. In addition to the lack of randomization, both studies conducted by Clark et al. (2003, 2008) were limited by small sample sizes, baseline demographic differences between groups (which is of concern given the lack of randomization), and missing outcome data that were not accounted for. The earlier study opted only to analyze data from recruited participants with more severe symptoms, as opposed to all recruited participants (Clark et al., 2003). Furthermore, an active treatment comparison group in Clark et al.' (2003) evaluation demonstrated that M-ITG failed to outperform an already well-established treatment (IPT). In fact, IPT may even be preferable due to the lack of observed increase in maternal negative affect and behavior that was observed among M-ITG participants. The paper by Smith et al. (2010) raised similar concerns in terms of quality due to the lack of randomization, the extended time period over which data were collected, and the general lack of rigorous protocols and procedures guiding recruitment, data collection, and data analysis. In contrast, Fonagy et al.' study (2016) featured randomization and intent-to-treat analyses to address moderate levels of attrition, in addition to conducting a power analysis to determine that the study was sufficiently powered for planned analyses, bolstering confidence in these findings relative to the others within this group. See Table 2 for a summary of risk of bias analyses.

Improvements in Mental Health Only

Three papers reporting outcomes from two studies demonstrated statistically significant improvements in mental health outcomes, but not in coded or observer-rated relational outcomes (Cooper et al., 2003; Ericksen et al., 2018; Murray et al., 2003).

A relationally-focused CBT intervention (r-CBT) was compared to two active treatment conditions and a TAU control condition, with maternal mental health outcomes reported by Cooper et al. (2003) and relational outcomes reported in Murray et al. (2003) for this longitudinal RCT. Mental health outcomes were measured at multiple time points up to 5 years post-intervention using the EPDS (Cox et al., 1987) and the Structured Clinical Interview for the DSM-III-R (SCID; Spitzer et al., 1992). All treatment conditions (i.e., r-CBT, psychodynamic, and non-directive counselling) were associated with significantly lower depressive symptoms (i.e., EPDS scores) post-intervention than the control condition, but were not significantly different from

one another (Cooper et al., 2003). No significant between-group differences were observed during follow-up assessments (9 months, 18 months, and 5 years post-intervention). In terms of depression remission, defined as no longer meeting DSM-III-R criteria for a depressive disorder, a significantly higher percentage (71%) of individuals receiving psychodynamic therapy experienced remission by the end of treatment compared to all other groups (57% r-CBT, 54% non-directive counselling, and 40% control). Similarly, however, no significant differences in terms of remission were detected during subsequent follow-up assessments. Relational outcomes assessed included infant attachment classification (Strange Situation Procedure [SSP]; Ainsworth et al., 1978), observer-rated parent–child interaction quality (Global Rating Scales of Mother–Infant Interaction [GRS]; Murray et al., 1996), and a parent-report relationship problems checklist that was developed by the authors (Murray et al., 2003). No significant between-group differences were found in terms of attachment or observed parent–child interactions. However, moms in all three treatment groups had a significantly lower relative risk of reporting relationship problems post-intervention than the control group, with the r-CBT group having the lowest relative risk (0.46; Murray et al., 2003). Thus, although the relationally-focused CBT did not result in improvements in observable relational outcomes, these self-report outcomes suggest there may still be value in terms of improving caregivers' perceptions of early relational wellbeing, albeit no more than that of the psychodynamic or non-directive counselling group. At the same time, demand characteristics related to the intervention may have influenced reporting, thus requiring caution in this interpretation.

Community HUGS (Happiness, Understanding, Giving, and Sharing; CHUGS) is a mother–infant playgroup intervention that was evaluated by Ericksen et al. (2018). Although primarily relational in emphasis, CHUGS participation was associated with significantly improved mental health symptoms but not relational outcomes. Specifically, CHUGS participants reported lower depression, anxiety, and stress than the WLC group, which was measured using the Depression, Anxiety, and Stress Scale (DASS; Lovibond & Lovibond, 1995). CHUGS participants rated their infants as significantly *more* difficult than WLC participants (based on the PSI; Abidin et al., 2006), and no differences were found in terms of parent–child interaction quality, rated by the Pediatric Infant Parent Exam (PIPE; Fiese et al., 2001).

The quality of these studies was variable, with both utilizing a randomized design and intent to treat principles to handle missing data (Cooper et al., 2003; Ericksen et al., 2018; Murray et al., 2003). However, although a power calculation was completed in both studies to determine an appropriate sample size, Ericksen et al. (2018) had to discontinue recruitment early due to funding constraints. As such, their study was considered underpowered. There was also a high degree of missing parent–child interactional data even when considering the already small sample size, and coders were not blinded to treatment condition. In contrast, the quality of the study evaluated by Cooper et al. (2003) and Murray et al. (2003) was markedly higher due to their use of adequate sample sizes and blinded coders. See Table 2 for a summary of risk of bias information.

Improvements in Relational Wellbeing Only

Three interventions were associated with improvements in relational but not mental health outcomes above and beyond those observed in the comparison conditions (Holt et al., 2021; Stein et al., 2006, 2018). However, it is worth noting that the comparison group in each of these studies received the same mental health treatment (i.e., cognitive-behavioral) as the treatment condition. This suggests that the relational add-on did not confer statistically meaningful *additional* improvements in mental health, but not that the intervention was not successful in treating maternal mental health. The absence of a no-treatment comparison group in these studies makes it difficult to infer the cause of the improvements in mental health observed across both groups.

Holt et al. (2021) evaluated the impact of adding a four-session HUGS playgroup (from which Ericksen et al.'s CHUGS intervention was adapted) following the receipt of 12 sessions of CBT for postnatal depression on maternal mental health and relational wellbeing. Mental health outcomes were measured using the BDI (Beck et al., 1961) and Beck Anxiety Inventory (BAI; Beck et al., 1988). No differences between groups were observed on these measures. The PCERA (Clark, 1985), Postpartum Bonding Questionnaire (PBQ; Brockington et al., 2006), and PSI (Abidin et al., 2006) were used to measure relational outcomes. Although no differences were detected in terms of individual items assessing observed maternal sensitivity or responsivity using the PCERA, one PCERA factor (i.e., positive affective involvement and verbalization), was found to be rated significantly higher among HUGS mothers at the 6-month follow-up (Holt et al., 2021). Similarly, mothers in the HUGS group self-reported lower impaired bonding at the 6-month follow-up period on the PBQ (Holt et al., 2021). These between-group differences were not observed immediately post-intervention, underscoring the importance of longer-term follow-up. No between-group differences were observed on the PSI.

Stein et al. (2006) evaluated guided cognitive behavioral self-help in conjunction with either video-feedback interactional treatment or supportive counseling in the treatment of bulimia. Although groups experienced similar rates of improvements in depression (EPDS; Cox et al., 1987) and eating disorder pathology (Eating Disorder Examination

[EDE]; Fairburn et al., 1993), mothers in the video feedback group had significantly better ratings in terms of levels of harmony, conflict, maternal facilitation, infant autonomy, and nonverbal responses to infant cues on an observational measure designed by the authors for this study (Stein et al., 2006). Similarly, Stein et al. (2018) compared CBT and video-feedback therapy (VFT) to CBT and progressive muscle relaxation (PMR) in the treatment of persistent postnatal depression. No significant between-group differences were found in terms of depression diagnosis (SCID; First & Gibbon, 2004), though EPDS (Cox et al., 1987) scores were significantly lower for the PMR group at 12-months follow-up. No significant *between*-group differences were observed in terms of following child's attention, contingent responsiveness, sensitivity, or warmth. However, significant improvements in maternal warmth were observed over time in the VFT group but not the PMR group (Stein et al., 2018). These outcomes were measured using a scale designed by the authors for this study to evaluate behaviors specifically targeted by the intervention (Stein et al., 2018). No between-group differences were observed in terms of attachment security, measured using the Attachment Q-Sort (AQS; Vaughn & Waters, 1990).

Despite all using comparison groups that limited conclusions regarding the effectiveness of the mental health components of treatment, studies within this group were among the higher quality studies reviewed. Both Stein et al. studies (2006, 2018) used power analyses to establish an adequate sample size and conducted analyses using intent-to-treat principles while also maintaining decent retention. Randomization was carefully carried out, with Stein et al. (2018) also describing concealment strategies used. Although Holt et al. (2021) conducted a power analysis to establish a target sample size, funding limitations impacted their ability to recruit an adequate sample size, and thus the study was considered underpowered. Rates of attrition were moderately high but groups remained comparable despite attrition. See Table 2 for additional details related to study quality.

No Evidence of Statistically Significant Improvement

Six remaining studies evaluating five interventions did not demonstrate statistically significant improvements in mental health or relational wellbeing, either because no statistically significant between-group differences were observed (Feeley et al., 2012; Goodman et al., 2015; Tsivos et al., 2015a; Zelkowitz et al., 2011), or because studies were described as underpowered to conduct between-group significance testing and significant within group changes were not differentially observed (Alhusen et al., 2021; Lenze et al., 2020).

Alhusen et al. (2021) evaluated a CBT-based Mothers and Babies Course (MB) delivered during pregnancy in a pilot randomized controlled trial. Although average improvements observed in the MB group were technically larger in magnitude than those observed in the comparison group in terms of depression (EPDS; Cox et al., 1987), observer-rated parent–child interactions (NCAST; Farel et al., 1991), and mother-reported fetal attachment (MFAS; Cranley, 1981), these differences were not tested for significance and effect sizes were not estimated, so it is not possible to determine whether these differences are meaningful. Lenze et al. (2020) evaluated Interpersonal Psychotherapy for the Mother–Infant Dyad (IPT-Dyad) and observed statistically significant decreases in depression (EPDS; Cox et al., 1987) and anxiety (Brief State-Trait Anxiety Inventory [STAI]; Marteau & Bekker, 1992) in both the IPT-Dyad and enhanced TAU condition. No differences were observed in terms of parenting stress (PSI; Abidin et al., 2006) in either group over time, and in terms of interactive behavior, there was a significant *decrease* in maternal sensitivity (based on the Coding Interactive Behaviors [CIB] manual; Feldman, 1998) associated with the treatment group only (Lenze et al., 2020).

Cues, an anxiety and parent–child interaction focused intervention delivered to parents of very low birthweight (VLBW) infants was compared to an attention control (Feeley et al., 2012; Zelkowitz et al., 2011). No significant between-group differences were observed in terms of anxiety (STAI; Spielberger et al., 1983), parental stress (Parental Stress Scale: Neonatal Intensive Care Unit [PSS:NICU]; Miles et al., 1993), trauma (Perinatal PTSD Questionnaire [PPQ]; DeMier et al., 1996), depression (EPDS; Cox et al., 1987), or parent–child interaction quality (GRS; Murray et al., 1996) either post-intervention (Zelkowitz et al., 2011) or at the 6-month follow-up (Feeley et al., 2012).

Perinatal Dyadic Psychotherapy (PDP) was compared to a TAU group that also received depression monitoring (Goodman et al., 2015). The SCID, EPDS, and STAI (Cox et al., 1987; First & Gibbon, 2004; Spielberger et al., 1983) were used to measure mental health outcomes, and no significant differences were observed between groups post-intervention or at the 3-month follow-up in terms of depression or anxiety. Relational outcomes were measured using the PSI (Abidin et al., 2006) and CIB (Feldman, 1998). Likewise, no differences were observed in relational outcomes.

Finally, Tsivos et al. (2015a) evaluated Baby Triple P (Positive Parenting Program) against TAU in a small ($n=27$) pilot RCT. No differences were observed between groups in terms of depression (BDI; Beck et al., 1961), self-reported bonding (PBQ; Brockington et al., 2006), or observer-rated parent–child interaction (CARE Index; Crittenden, 2004).

The quality of studies within this group was somewhat variable but overall decent to high (See Table 2). However, sample sizes were small and/or considered by authors to be underpowered in several of the studies (Alhusen et al., 2021; Goodman et al., 2015; Lenze et al., 2020; Tsivos et al.,

2015a), making it difficult to detect or measure between-group differences. The majority of studies either had low rates of attrition (Alhusen et al., 2021; Goodman et al., 2015) or used reasonable strategies to account for attrition (Feeley et al., 2012; Tsivos et al., 2015a; Zelkowitz et al., 2011). Inadequate blinding strategies were reported in two studies (Lenze et al., 2020; Tsivos et al., 2015a), where interaction coders were not blinded to treatment condition. The design for the Cues trial (Feeley et al., 2012; Zelkowitz et al., 2011) was overall rather high in quality. However, the use of an educational comparison group, in addition to the fact that high anxiety following the birth of a VLBW infant may remit naturally once the infant is healthy, may account for their null findings (Feeley et al., 2012; Zelkowitz et al., 2011). Likewise, depression monitoring that occurred in the comparison group of the PDP trial may have had more of a therapeutic impact than intended (Goodman et al., 2015).

Commonalities Among Successful Interventions

Because fewer studies in the present review utilized comparison groups enabling rigorous between-group comparisons in terms of effective mental health treatment components, in addition to the fact that commonalities of effective relational interventions have been reported on elsewhere (e.g., Newton et al., 2022), this review addresses commonalities among those interventions that were associated with improvements in both domains. Additionally, although parent-report and observer- or coder-rated relational outcomes were largely consistent across studies, there was one study where improvements were demonstrated in mother-reported relational outcomes but not in observer-rated outcomes (Cooper et al., 2003; Murray et al., 2003). Improvements in maternal perceptions can still be considered meaningful, as maternal perception may mediate links between depression and observable differences in interaction quality (Trapolini et al., 2008). Thus, the relationally-enhanced CBT treatment evaluated by Cooper et al. and Murray et al. (2003) will be considered moderately successful for the purposes of the present discussion, despite the lack of between-group differences reported in observable or coded outcomes.

Format and Style

With the exception of the relationally-enhanced CBT treatment (Cooper et al., 2003; Murray et al., 2003), which appeared more or less balanced in emphasis, interventions in this category (M-ITG, PIP, and PAIRS) were all primarily relational in focus (Clark et al., 2003, 2008; Fonagy et al., 2016; Smith et al., 2010). In all interventions, mental health and relational components of treatment were delivered simultaneously, with both aspects being attended to within each individual session. Most treatments were briefer in nature, with all except for PIP being delivered in 12 sessions or fewer, whereas the number of sessions used in PIP was flexible depending on the needs of the dyad ($M = 16$ sessions among participants who attended at least one session; Fonagy et al., 2016).

There did not appear to be a clear advantage conferred by a group versus individual format for treatment delivery. It is worth highlighting, however, that effective group interventions (M-ITG and PAIRS) utilized similar structures that included both conjoint dyadic time and parallel groups where mothers and infants met with interventionists separately (Clark et al., 2003, 2008; Smith et al., 2010). PIP and r-CBT were delivered individually, with the infant attending the entire session of PIP along with the mother (Fonagy et al., 2016), and no indication of whether infants in r-CBT were present throughout the entire sessions or just when relational functioning was being addressed (Cooper et al., 2003; Murray et al., 2003).

Mental Health Approaches

In general, effective interventions tended to make use of strategies that are common features across many treatment modalities, as opposed to strict adherence to a single modality or manual. For example, all interventions sought to promote expression and exploration of maternal thoughts and feelings, both in general, and as they pertain to parenting, the parent–child relationship, and their infant (Clark et al., 2003, 2008; Cooper et al., 2003; Fonagy et al., 2016; Murray et al., 2003; Smith et al., 2010). All utilized some form of collaborative problem-solving between the caregiver and therapist. Group interventions (Clark et al., 2003, 2008; Smith et al., 2010) also incorporated aspects of enhancing social support, with M-ITG explicitly utilizing IPT strategies (Clark et al., 2003, 2008) to enhance interpersonal communication and relationships, and both M-ITG and PAIRS inviting partners/fathers/social supports to a handful of sessions to provide psychoeducation about depression and promote communication and interpersonal functioning (Clark et al., 2003, 2008; Smith et al., 2010). Additionally, the use of a group format itself was considered a strategy to enhance social support and interpersonal communication. Unfortunately, mental health components of r-CBT and PIP were poorly described, limiting further discussion (Cooper et al., 2003; Fonagy et al., 2016; Murray et al., 2003).

Mother–Infant Relational Approaches

All interventions (M-ITG, r-CBT, PIP, and PAIRS) incorporated some form of interaction coaching or guidance, with therapists taking care to highlight and reinforce positive and sensitive caregiving moments and appropriate affect in vivo, while gently encouraging and supporting

practice of new ways of interacting as indicated (Clark et al., 2003, 2008; Cooper et al., 2003; Fonagy et al., 2016; Murray et al., 2003; Smith et al., 2010). Smith et al. (2010) specifically utilized within-session separations and reunions to orient caregivers to attachment-related behaviors. Only r-CBT (Cooper et al., 2003; Murray et al., 2003) made use of in-session video recording and review to achieve these interactive goals, though specific strategies were not described within the text. Authors succinctly described the intervention as an "appropriately modified form of the interaction guidance treatment described by McDonough (1993)" (Cooper et al., 2003, p. 412). Assumptions were made about specific strategies utilized in their relational-CBT based on descriptions of interaction guidance within the cited text (McDonough, 1993).

Facilitators created opportunities for positive interaction in sessions (e.g., using games, toys, infant massage, music, and movement) and occasionally used modeling to promote positive affect and interaction styles and strategies (Clark et al., 2003, 2008; Cooper et al., 2003; Fonagy et al., 2016; Murray et al., 2003; Smith et al., 2010). M-ITG and PAIRS also utilized infant developmental therapy procedures during the parallel portion of their groups to promote infant emotion regulation, thereby enhancing infants' ability to engage socially with and be reinforcing to their caregivers upon reunion (Clark et al., 2003, 2008; Smith et al., 2010).

Enhanced verbal and nonverbal communication were targeted through several avenues. Strategies to promote curiosity and reflection about the meaning behind infant cues and affect (e.g., watch, wait, and wonder; the therapist "speaking for the infant;" and therapists soliciting caregiver reflections regarding interactions prior to offering feedback) were used to increase reflective functioning and the capacity of caregivers to understand infant behavior (Clark et al., 2003; Clark et al., 2008; Cooper et al., 2003; Fonagy et al., 2016; Murray et al., 2003; Smith et al., 2010). Based on McDonough's description of the interactive guidance intervention (1993), mothers in the r-CBT were also prompted to reflect on the content and tone of their verbal interactions with the infant.

Finally, most interventions contained components that were more psychodynamic in nature and approach (Clark et al., 2003, 2008; Fonagy et al., 2016; Smith et al., 2010). These interventions sought to address caregivers' representations, or internal working models (IWMs), related to caregiving and their infants, using discussion about the mother's family of origin to make sense of these IWMs, while encouraging appropriate differentiation between the family of origin and current context (Clark et al., 2003, 2008; Fonagy et al., 2016; Smith et al., 2010).

Discussion

The aim of the present review was to summarize the state of the literature evaluating interventions delivered during the perinatal period that target both maternal mental health and relational outcomes. Thirteen interventions, evaluated in 16 articles, were reviewed. Eight interventions demonstrated improvements in at least one outcome of interest (Clark et al., 2003, 2008; Cooper et al., 2003; Ericksen et al., 2018; Fonagy et al., 2016; Holt et al., 2021; Murray et al., 2003; Smith et al., 2010; Stein et al., 2006, 2018). However, even when including parent-reported relational outcomes, only four interventions exhibited improvements in both outcomes (Clark et al., 2003, 2008; Cooper et al., 2003; Fonagy et al., 2016; Murray et al., 2003; Smith et al., 2010), and five demonstrated improvements in neither (Alhusen et al., 2021; Feeley et al., 2012; Goodman et al., 2015; Lenze et al., 2020; Tsivos et al., 2015a; Zelkowitz et al., 2011). The lack of randomization among many of the studies classified as showing improvements in both is noteworthy, however, and demands caution when considering clinical implications since other factors may account for these improvements. Furthermore, given the small number of interventions included, mixed findings, and diversity among interventions reviewed, no definitive consensus can be reached at this time regarding "what works" or, likewise, what does not.

Indeed, although it is certainly possible that many of these interventions were simply not effective, alternative explanations for null findings reported in several studies are worth considering. First, the timing of outcome measurement may have played a role. The perinatal period is a time of great transition that is also characterized by disturbances in sleep, physical health, and social relationships that can greatly influence mental health. Thus, it is common for mental health symptoms that emerge during this period to remit naturally over time as individuals adjust and circumstances normalize, which may obscure between-group differences if data are not collected at intervals that may be more likely to detect these differences. For instance, Holt et al. (2021) reported relational improvements at the 6-month follow-up assessment, but not immediately post-intervention, while other interventions reported short-term, but not long-term between-group differences (Cooper et al., 2003; Murray et al., 2003). Despite the fact that symptoms may naturally remit with time, the developmental salience of the postpartum period for infants and laying the groundwork for parent–child relationships easily justifies ongoing efforts to promote caregiver wellbeing during this period as proactively as possible and demands attention to the timing of evaluating these efforts to detect their impact.

Additionally, intervention effects may have been obscured due to variability in what comparison groups

received. Importantly, the meaning of "treatment as usual" can vary drastically depending on the accessibility and affordability of healthcare services where treatment is provided in addition to cultural norms regarding help-seeking. As such, TAU could imply a relatively high level of care in some cases but not others. Finally, null findings may also be explained by the use of active treatment and attention comparison groups in several studies. Although an active comparison group avoids the ethical dilemma of denying care to individuals during a critical window for receiving mental healthcare, this can make it difficult to make claims regarding the effectiveness of a treatment. At the same time, the potential effectiveness of many of these alternative treatment options may be a cause for hopefulness, as this could suggest that treatment recommendations can reasonably be tailored to mothers' preferences without sacrificing effectiveness.

Clinical Implications

With these limitations in mind, findings from the present review may tentatively inform and guide future clinical practice in terms of both the format and content of interventions sharing these goals.

In terms of format and approach, the majority of successful interventions were characterized as primarily relational in emphasis, though both mental health and relational targets were addressed in all intervention sessions (Clark et al., 2003, 2008; Cooper et al., 2003; Fonagy et al., 2016; Murray et al., 2003; Smith et al., 2010). Although it may seem intuitive that interventions should seek to improve maternal mental health *prior* to addressing relational functioning (as in Feeley et al., 2012; Holt et al., 2021; and Zelkowitz et al., 2011), our findings suggest that a simultaneous approach may be more appropriate. Effective interventions also tended to be briefer in nature (i.e., 12 sessions or fewer; Clark et al., 2003, 2008; Cooper et al., 2003; Murray et al., 2003; Smith et al., 2010), which is in line with past meta-analytic findings (Bakermans-Kranenburg et al., 2003; Van Ijzendoorn et al., 1995). It is possible that these shorter-term interventions are more targeted and goal-driven, and therefore more effective in enhancing outcomes of interest than interventions that may be longer-term due to being broader in scope or less focused. Furthermore, shorter-term interventions may simply be more effective if they are more accessible and feasible given the unique demands facing perinatal populations. Future research should explore these possibilities.

In the present review, there did not appear to be a clear treatment benefit in terms of using a group versus individual format, although previous studies suggest that individual therapy may be associated with larger improvements in mental health outcomes (Sockol, 2015; Sockol et al., 2011). This finding is important given that women's preferences in terms of format may vary based on a variety of considerations (Dennis & Chung-Lee, 2006; Goodman, 2009), and the absence of a clearly superior approach suggests that these preferences may reasonably guide treatment decisions. Effective group interventions in the present review used a "parallel conjoint" model, in which dyads spent part of each session in separate mother and infant groups, and part of each session together (Clark et al., 2003, 2008; Smith et al., 2010), which enabled devotion of attention to infants and mothers individually.

In terms of intervention content that promoted maternal mental health, more general strategies such as promoting emotion expression and problem-solving were emphasized across effective interventions (Clark et al., 2003, 2008; Cooper et al., 2003; Fonagy et al., 2016; Murray et al., 2003; Smith et al., 2010). Group interventions also targeted more interpersonal skills and social supports (Clark et al., 2003, 2008; Smith et al., 2010). Although many intervention descriptions specifically highlighted the use of cognitive-behavioral or interpersonal *strategies*, the overall more general nature of strategies described in interventions that qualified as effective in this review calls into question the need to utilize strictly manualized interventions.

Relationally, effective interventions incorporated guided parent–child interactions and used in vivo coaching and modeling to promote sensitive and responsive caregiving (Clark et al., 2003, 2008; Cooper et al., 2003; Fonagy et al., 2016; Murray et al., 2003; Smith et al., 2010). Interventions also targeted reflective functioning to enhance understanding of infant cues, behavior, and affect, and incorporated discussion of past experiences and how they inform parents' IWMs of their infants and relationships (Clark et al., 2003, 2008; Cooper et al., 2003; Fonagy et al., 2016; Murray et al., 2003; Smith et al., 2010). Although video feedback has previously been highlighted as an important feature of effective parent–child relationship-focused interventions (Newton et al., 2022; Wright et al., 2023), only one of the interventions in this grouping (Cooper et al., 2003; Murray et al., 2003) made use of video feedback specifically. These findings suggest that video feedback, although certainly valuable, may not be necessary per se, as long as interventions make use of some form of interaction guidance, and this possibility should be explored in future work.

Although these guidelines may be useful in clinical practice, it is important to hold in mind social determinants of mental and relational wellbeing, including but not limited to factors such as health care access and quality, economic stability, exposure to discrimination (at individual- and systems-levels), the built environment, safety, and other contextual factors. Therefore, in addition to considering these guidelines in clinical practice, it remains incumbent upon professionals within the perinatal mental health field to advocate for policy and systems-level changes that can

impact these outcomes beyond what can be addressed in the context of psychotherapeutic interventions.

Strengths, Limitations, and Future Directions

This review has many strengths. Search strategies were conducted thoroughly and systematically and reported in a manner consistent with best practices (Page et al., 2021; Siddaway et al., 2019). The fact that no additional articles were identified for inclusion when reference lists of included studies were hand-searched further supports this assertion. Furthermore, although posing a challenge for synthesis, broader inclusion criteria in terms of study design and comparison groups enabled thorough review of what is known regarding interventions targeting these outcomes, which is an important first step on which to base recommendations for future research.

Furthermore, this review fills an important gap in the literature by highlighting outcomes associated with interventions that are both mental health and relational in focus. Much of the literature on perinatal mental illness has emphasized its impact on subsequent generations (e.g., Letourneau et al., 2017; Poobalan et al., 2007; Tsivos et al., 2015b). While this impact is important and has significant public health implications, the wellbeing of caregivers matters beyond merely its impact on the next generation. The present review acknowledges the importance of maternal wellbeing in its own right by prioritizing dual-generation interventions.

A few meaningful limitations in terms of both review procedures and included evidence also merit discussion. First, although it may be considered beneficial to have been inclusive in terms of study design in the initial stages of evaluating this literature, this also presented significant challenges. Specifically, variability in terms of comparison groups used and a wide range in terms of study design (and thus, quality) makes it difficult to draw clear or confident conclusions regarding what is effective, which was a key aim of the review. For instance, all three interventions that were associated with improvements in relational outcomes only (Holt et al., 2021; Stein et al., 2006, 2018) may have been found to be associated with improvements in both, had they used a TAU or WLC comparison group. At the same time, when new or more specialized interventions fail to outperform more well-established (e.g., IPT or CBT) or less resource-intensive (e.g., educational or depression monitoring) interventions, one can fairly question whether the studied intervention is "worth it."

As discussed, the quality of included studies varied, and many studies were small in scale or experienced high rates of attrition, which is a common challenge in perinatal research due to the high demands placed on individuals' attention and energy during this life transition. Lower quality and non-randomized studies further interfered with the ability to draw confident conclusions from included studies, and the lack of randomization in three of the studies (Clark et al., 2003, 2008; Smith et al., 2010) significantly compromises their internal validity (Dunbar-Jacob, 2018). This is particularly noteworthy given that these studies were also among those identified as demonstrating the most favorable outcomes, and it is curious that larger-scale follow-up studies of these interventions have not yet been conducted.

Additionally, inclusion criteria intentionally restricted eligible studies to those evaluating interventions that are strictly perinatal (i.e., delivered during pregnancy and/or up to 18 months postpartum) due to a desire to focus on this period and population specifically. However, other interventions that share these dual-generation targets exist and have undergone evaluation (e.g., Mom Power; Lawler et al., 2018, or the Michigan Model of Infant Mental Health Home visiting; Weatherston et al., 2020). It is important to consider the impact of these interventions and how they may be similar or different to those included in this review, as well as potential "lessons learned" from interventions that are broader in terms of developmental range.

As funding allows, future studies should utilize an RCT design and oversample to ensure studies are sufficiently powered in the face of attrition, while also taking steps to minimize attrition (e.g., Coen et al.' integrated management approach [1996]). Additionally, given that the perinatal period involves rapid change and development, a greater frequency of data collection may improve a study's ability to detect meaningful yet nuanced changes associated with interventions. Future studies may also benefit from the use of multiple comparison groups to enable inferences regarding whether studied interventions are better than nothing, or TAU, as well as whether they outperform existing interventions. In the case of comparison to existing interventions, attention to differential attrition, ratings of participant satisfaction, and other qualitative data may be informative in scenarios where other outcomes of interest are comparable. As such, data related to study acceptance by participants should also routinely be collected and reported.

Relatedly, homogeneity in terms of the racial and ethnic identity of participants is an important limitation of studies included in the present review, as well as a long-standing problem in the mental health treatment evaluation literature overall (Polo et al., 2019). When race and ethnicity were reported, the majority of participants in the included studies were members of racial/ethnic majority groups and samples were largely homogenous. Only two studies (Alhusen et al., 2021; Lenze et al., 2020) reported demographic characteristics related to racial and ethnic identity that did *not* appear to be primarily from a majority group (i.e., African American), while several studies failed to report on these demographic characteristics completely or primarily reported race and ethnicity in a binary

manner that referenced the majority (e.g., X % majority vs. Y% non-majority). This is problematic, as research has shown that traditional mental health service models are not as widely accepted by or beneficial to minoritized populations whose needs and preferences are less frequently taken into account (Parker, 2021; Stevens et al., 2018).

Other aspects of diversity in the field of perinatal mental health, including diversity related to parent gender, sexual orientation, and mental health condition have been insufficiently explored. Fathers and non-birthing caregivers experience mental health challenges during the perinatal period that impact child development, yet far fewer studies have examined perinatal mental health and relationships outside of the context of mothers (Wong et al., 2016). Likewise, interventions targeting perinatal depression have been widely studied, whereas interventions designed to treat perinatal PTSD and anxiety disorders are far less prevalent (Nillniet al., 2018), even though these conditions are common (Howard & Khalifeh, 2020; Leach et al., 2017) and impact parent–child relationships (Erickson et al., 2019; Gaudet et al., 2010). Future studies should seek to understand and address the needs of different perinatal populations and should study interventions across diverse groups to improve generalizability.

Finally, it is worth noting that the exclusion of articles published in non-English languages may be considered a limitation of the present review, despite some literature suggesting that this practice is acceptable. There were 141 articles excluded automatically on the basis of this criterion, and future reviews may benefit from incorporating strategies for reviewing articles published in all languages.

Conclusions

This review offers an important step toward identifying what is known and what remains to be discovered regarding perinatal mental health interventions targeting both mental health and parent–child relational outcomes. Findings regarding clinical outcomes have important implications for primary and secondary prevention of mental illness and important future directions for research have been highlighted.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Review Registration The review protocol was pre-registered with Prospero (ID: CRD42022380278).

References

Abidin, R., Flens, J. R., & Austin, W. G. (2006). The parenting stress index. In R. P. Archer (Ed.), *Forensic uses of clinical assessment instruments* (pp. 297–328). Lawrence Erlbaum Associates Publishers.

Ainsworth, M. D. (1973). The development of infant mother attachment. In B. M. Caldwell & H. N. Ricciuti (Eds.), *Review of child development research.* (Vol. 3). University of Chicago Press.

Ainsworth, M. D. S. (1978). *Patterns of attachment: A psychological study of the strange situation*. Lawrence Erlbaum Associates.

Alhusen, J. L., Hayat, M. J., & Borg, L. (2021). A pilot study of a group-based perinatal depression intervention on reducing depressive symptoms and improving maternal-fetal attachment and maternal sensitivity. *Archives of Women's Mental Health, 24*(1), 145–154. https://doi.org/10.1007/s00737-020-01032-0

Avinun, R., & Knafo, A. (2014). Parenting as a reaction evoked by children's genotype: A meta-analysis of children-as-twins studies. *Personality and Social Psychology Review: an Official Journal of the Society for Personality and Social Psychology, Inc, 18*(1), 87–102. https://doi.org/10.1177/1088868313498308

Bakermans-Kranenburg, M. J., van Ijzendoorn, M. H., & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin, 129*(2), 195–215. https://doi.org/10.1037/0033-2909.129.2.195

Barker, E., Copeland, W., Maughan, B., Jaffee, S., & Uher, R. (2012). Relative impact of maternal depression and associated risk factors on offspring psychopathology. *The British Journal of Psychiatry, 200*(2), 124–129. https://doi.org/10.1192/bjp.bp.111.092346

Barlow, J., Bennett, C., Midgley, N., Larkin, S. K., & Wei, Y. (2015). Parent–infant psychotherapy for improving parental and infant mental health. *The Cochrane Database of Systematic Reviews, 1*(1), CD010534. https://doi.org/10.1002/14651858.CD010534.pub2

Batt, M. M., Olsavsky, A. K., Dardar, S., St. John-Larkin, C., Johnson, R. L., & Sammel, M. D. (2022). Course of illness and treatment updates for bipolar disorder in the perinatal period. *Current Psychiatry Reports, 24*(2), 111–120.

Bauer, A., Knapp, M., & Parsonage, M. (2016). Lifetime costs of perinatal anxiety and depression. *Journal of Affective Disorders, 192*, 83–90. https://doi.org/10.1016/j.jad.2015.12.005

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56*(6), 893.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*(6), 561–571.

Bergink, V., Rasgon, N., & Wisner, K. L. (2016). Postpartum psychosis: Madness, mania, and melancholia in motherhood. *American Journal of Psychiatry, 173*(12), 1179–1188.

Bernier, A., Matte-Gagné, C., Bélanger, M. È., & Whipple, N. (2014). Taking stock of two decades of attachment transmission gap: Broadening the assessment of maternal behavior. *Child Development, 85*(5), 1852–1865. https://doi.org/10.1111/cdev.12236

Bowlby, J. (1958). The nature of the child's tie to his mother. *The International Journal of Psychoanalysis, 39*, 350–373.

Bowlby, J. (1969). *Attachment and loss*. Basic Books.

Bowlby, J. (1988). *A secure base: Parent–child attachment and healthy human development*. Basic Books.

Branje, S., Geeraerts, S., de Zeeuw, E. L., Oerlemans, A. M., Koopman-Verhoeff, M. E., Schulz, S., Nelemans, S., Meeus,

W., Hartman, C. A., Hillegers, M. H. J., Oldehinkel, A. J., & Boomsma, D. I. (2020). Intergenerational transmission: Theoretical and methodological issues and an introduction to four Dutch cohorts. *Developmental Cognitive Neuroscience, 45*, Article 100835.

Brockington, I. F., Fraser, C., & Wilson, D. (2006). The Postpartum Bonding Questionnaire: A validation. *Archives of Women's Mental Health, 9*, 233–242.

Brouwers, E. P., van Baar, A. L., & Pop, V. J. (2001). Maternal anxiety during pregnancy and subsequent infant development. *Infant Behavior and Development, 24*(1), 95–106.

Brumariu, L. E., & Kerns, K. A. (2013). Pathways to anxiety: Contributions of attachment history, temperament, peer competence, and ability to manage intense emotions. *Child Psychiatry and Human Development, 44*(4), 504–515. https://doi.org/10.1007/s10578-012-0345-7

Burns, L., Coleman-Cowger, V. H., & Breen, C. (2016). Managing maternal substance use in the perinatal period: current concerns and treatment approaches in the United States and Australia. *Substance Abuse: Research and Treatment, 10*, SART-S34558.

Censullo, M. (1991). *Dyadic mutuality code manual*. Wellesley College Center for Research on Women.

Clark, R. (1985). Parent–child early relational assessment (ERA). *APA PsycTests*. https://doi.org/10.1037/t07207-000

Clark, R., Tluczek, A., & Brown, R. (2008). A mother–infant therapy group model for postpartum depression. *Infant Mental Health Journal, 29*(5), 514–536. https://doi.org/10.1002/imhj.20189

Clark, R., Tluczek, A., & Wenzel, A. (2003). Psychotherapy for postpartum depression: A preliminary report. *American Journal of Orthopsychiatry, 73*(4), 441–454. https://doi.org/10.1037/0002-9432.73.4.441

Coen, A. S., Patrick, D. C., & Shern, D. L. (1996). Minimizing attrition in longitudinal studies of special populations: An integrated management approach. *Evaluation and Program Planning, 19*(4), 309–319.

Cook, N., Ayers, S., & Horsch, A. (2018). Maternal posttraumatic stress disorder during the perinatal period and child outcomes: A systematic review. *Journal of Affective Disorders, 225*, 18–31. https://doi.org/10.1016/j.jad.2017.07.045

Cooper, P. J., Murray, L., Wilson, A., & Romaniuk, H. (2003). Controlled trial of the short- and long-term effect of psychological treatment of post-partum depression: 1. Impact on maternal mood. *British Journal of Psychiatry, 182*(5), 412–419. https://doi.org/10.1192/bjp.182.5.412

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale. *The British Journal of Psychiatry, 150*(6), 782–786.

Cranley, M. S. (1981). Development of a tool for the measurement of maternal attachment during pregnancy. *Nursing Research, 30*(5), 281–284.

Critical Appraisal Skills Programme. (2020). *CASP randomized controlled trial checklist.* https://casp-uk.net/news/updated-rct-checklist/

Crittenden, P. M. (2004*). CARE-Index: Coding manual.* Unpublished manuscript. Family Relations Institute.

Cucciniello, C., & Melia, Y. (2024). Effectiveness of parent–infant psychotherapy for infant development, parenting, parental mental health, and the attachment relationship: A review of the literature. *Journal of Public Child Welfare, 2024*, 1–35.

De Boer, M. R., Waterlander, W. E., Kuijper, L. D., Steenhuis, I. H., & Twisk, J. W. (2015). Testing for baseline differences in randomized controlled trials: An unhealthy research behavior that is hard to eradicate. *International Journal of Behavioral Nutrition and Physical Activity, 12*, 1–8.

DeMier, R. L., Hynan, M. T., Harris, H. B., & Manniello, R. L. (1996). Perinatal stressors as predictors of symptoms of posttraumatic stress in mothers of infants at high risk. *Journal of Perinatology, 16*, 276–280.

Dennis, C. L., & Chung-Lee, L. (2006). Postpartum depression help-seeking barriers and maternal treatment preferences: A qualitative systematic review. *Birth, 33*(4), 323–331.

Derogatis, L. R., & Melisaratos, N. (1983). The brief symptom inventory: An introductory report. *Psychological Medicine, 13*(3), 595–605.

Dowse, E., Chan, S., Ebert, L., Wynne, O., Thomas, S., Jones, D., Fealy, S., Evans, T., & Oldmeadow, C. (2020). Impact of perinatal depression and anxiety on birth outcomes: A retrospective data analysis. *Maternal and Child Health Journal, 24*, 718–726.

Dunbar-Jacob, J. (2018). Minimizing threats to internal validity. In D. Morrison-Beedy & B. M. Melnyk (Eds.), *Intervention research and evidence-based quality improvement: Designing, conducting, analyzing, and funding* (pp. 129–143). Springer.

Ericksen, J., Loughlin, E., Holt, C., Rose, N., Hartley, E., Buultjens, M., Gemmill, A. W., & Milgrom, J. (2018). A therapeutic playgroup for depressed mothers and their infants: Feasibility study and pilot randomized trial of Community HUGS. *Infant Mental Health Journal, 39*(4), 396–409. https://doi.org/10.1002/imhj.21723

Erickson, N., Julian, M., & Muzik, M. (2019). Perinatal depression, PTSD, and trauma: Impact on mother–infant attachment and interventions to mitigate the transmission of risk. *International Review of Psychiatry, 31*(3), 245–263.

Essau, C. A., Sasagawa, S., Lewinsohn, P. M., & Rohde, P. (2018). The impact of pre-and perinatal factors on psychopathology in adulthood. *Journal of Affective Disorders, 236*, 52–59.

Fairburn, C. G., Cooper, Z., & O'Connor, M. (1993). The eating disorder examination. *International Journal of Eating Disorders, 6*, 1–8.

Farel, A. M., Freeman, V. A., Keenan, N. L., & Huber, C. J. (1991). Interaction between high-risk infants and their mothers: The NCAST as an assessment tool. *Research in Nursing and Health, 14*(2), 109–118.

Fearon, R. M., & Belsky, J. (2016). Precursors of attachment security. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (3rd ed., pp. 291–313). Guilford Press.

Feeley, N., Zelkowitz, P., Shrier, I., Stremler, R., Westreich, R., Dunkley, D., Steele, R., Rosberger, Z., Lefebvre, F., & Papageorgiou, A. (2012). Follow-up of the cues and care trial: Mother and infant outcomes at 6 months. *Journal of Early Intervention, 34*(2), 65–81. https://doi.org/10.1177/1053815112453767

Feldman, R. (1998). *Coding interactive behavior manual*. Unpublished manual. Bar-Ilan University.

Feldman, R., Granat, A., Pariente, C., Kanety, H., Kuint, J., & Gilboa-Schechtman, E. (2009). Maternal depression and anxiety across the postpartum year and infant social engagement, fear regulation, and stress reactivity. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*(9), 919–927. https://doi.org/10.1097/CHI.0b013e3181b21651

Fiese, B. H., Poehlmann, J., Irwin, M., Gordon, M., & Curry-Bleggi, E. (2001). A pediatric screening instrument to detect problematic infant–parent interactions: Initial reliability and validity in a sample of high-and low-risk infants. *Infant Mental Health Journal, 22*(4), 463–478.

First, M. B., & Gibbon, M. (2004). The structured clinical interview for DSM-IV axis I disorders (SCID-I) and the structured clinical interview for DSM-IV axis II disorders (SCID-II). In M. J. Hilsenroth & D. L. Segal (Eds.), *Comprehensive handbook of psychological assessment* (pp. 134–143). Wiley.

Fonagy, P., Sleed, M., & Baradon, T. (2016). Randomized controlled trial of parent–infant psychotherapy for parents with mental health problems and young infants. *Infant Mental Health Journal, 37*(2), 97–114. https://doi.org/10.1002/imhj.21553

Forman, D. R., O'Hara, M. W., Stuart, S., Gorman, L. L., Larsen, K. E., & Coy, K. C. (2007). Effective treatment for postpartum depression is not sufficient to improve the developing mother–child relationship. *Development and Psychopathology, 19*(2), 585–602. https://doi.org/10.1017/S0954579407070289

Forray, A., & Foster, D. (2015). Substance use in the perinatal period. *Current Psychiatry Reports, 17*, 1–11.

Gaudet, C., Séjourné, N., Camborieux, L., Rogers, R., & Chabrol, H. (2010). Pregnancy after perinatal loss: Association of grief, anxiety and attachment. *Journal of Reproductive and Infant Psychology, 28*(3), 240–251.

Goodman, J. H. (2009). Women's attitudes, preferences, and perceived barriers to treatment for perinatal depression. *Birth, 36*(1), 60–69.

Goodman, J. H., Guarino, A., Chenausky, K., Klein, L., Prager, J., Petersen, R., Forget, A., & Freeman, M. (2014). CALM Pregnancy: Results of a pilot study of mindfulness-based cognitive therapy for perinatal anxiety. *Archives of Women's Mental Health, 17*(5), 373–387. https://doi.org/10.1007/s00737-013-0402-7

Goodman, J. H., Prager, J., Goldstein, R., & Freeman, M. (2015). Perinatal dyadic psychotherapy for postpartum depression: A randomized controlled pilot trial. *Archives of Women's Mental Health, 18*(3), 493–506. https://doi.org/10.1007/s00737-014-0483-y

Goodman, S. H., & Gotlib, I. H. (1999). Risk for psychopathology in the children of depressed mothers: A developmental model for understanding mechanisms of transmission. *Psychological Review, 106*(3), 458–490. https://doi.org/10.1037/0033-295x.106.3.458

Groh, A. M., Roisman, G. I., Booth-LaForce, C., Fraley, R. C., Owen, M. T., Cox, M. J., & Burchinal, M. R. (2014). Stability of attachment security from infancy to late adolescence. *Monographs of the Society for Research in Child Development, 79*(3), 51–66.

Hernández-Martínez, C., Arija, V., Balaguer, A., Cavallé, P., & Canals, J. (2008). Do the emotional states of pregnant women affect neonatal behaviour? *Early Human Development, 84*(11), 745–750.

Holt, C., Gentilleau, C., Gemmill, A. W., & Milgrom, J. (2021). Improving the mother–infant relationship following postnatal depression: A randomised controlled trial of a brief intervention (HUGS). *Archives of Women's Mental Health, 24*(6), 913–923. https://doi.org/10.1007/s00737-021-01116-5

Howard, L. M., & Khalifeh, H. (2020). Perinatal mental health: A review of progress and challenges. *World Psychiatry, 19*(3), 313–327.

Khan, S. J., Fersh, M. E., Ernst, C., Klipstein, K., Albertini, E. S., & Lusskin, S. I. (2016). Bipolar disorder in pregnancy and postpartum: Principles of management. *Current Psychiatry Reports, 18*, 1–11.

Kim, D. R., Sockol, L., Barber, J. P., Moseley, M., Lamprou, L., Rickels, K., O'Reardon, J., & Epperson, C. N. (2011). A survey of patient acceptability of repetitive transcranial magnetic stimulation (TMS) during pregnancy. *Journal of Affective Disorders, 129*(1–3), 385–390.

Lawler, J. M., Rosenblum, K. L., Schuster, M., & Muzik, M. (2018). Mom Power: A parenting group intervention for mothers with trauma histories. In M. Muzik & K. L. Rosenblum (Eds.), *Motherhood in the face of trauma: Pathways towards healing and growth* (1st ed., pp. 165–180). Springer. https://doi.org/10.1007/978-3-319-65724-0_7

Leach, L. S., Poyser, C., & Fairweather-Schmidt, K. (2017). Maternal perinatal anxiety: A review of prevalence and correlates. *Clinical Psychologist, 21*(1), 4–19.

Lenze, S. N., Potts, M. A., Rodgers, J., & Luby, J. (2020). Lessons learned from a pilot randomized controlled trial of dyadic interpersonal psychotherapy for perinatal depression in a low-income population. *Journal of Affective Disorders, 271*, 286–292. https://doi.org/10.1016/j.jad.2020.03.084

Letourneau, N. L., Dennis, C., Cosic, N., & Linder, J. (2017). The effect of perinatal depression treatment for mothers on parenting and child development: A systematic review. *Depression and Anxiety, 34*(10), 928–966. https://doi.org/10.1002/da.22687

Lilliecreutz, C., Josefsson, A., & Sydsjö, G. (2010). An open trial with cognitive behavioral therapy for blood-and injection phobia in pregnant women—A group intervention program. *Archives of Women's Mental Health, 13*, 259–265.

Lindensmith, R. (2018). Interventions to improve maternal-infant relationships in mothers with postpartum mood disorders. *MCN: the American Journal of Maternal/child Nursing, 43*(6), 334–340.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy, 33*(3), 335–343.

Madigan, S., Brumariu, L. E., Villani, V., Atkinson, L., & Lyons-Ruth, K. (2016). Representational and questionnaire measures of attachment: A meta-analysis of relations to child internalizing and externalizing problems. *Psychological Bulletin, 142*(4), 367–399. https://doi.org/10.1037/bul0000029

Maguire, P. N., Clark, G. I., & Wootton, B. M. (2018). The efficacy of cognitive behavior therapy for the treatment of perinatal anxiety symptoms: A preliminary meta-analysis. *Journal of Anxiety Disorders, 60*, 26–34.

Main, M., Kaplan, N., & Cassidy, J. (1985). Security in infancy, childhood, and adulthood: A move to the level of representation. *Monographs of the Society for Research in Child Development, 50*(1/2), 66–104. https://doi.org/10.2307/3333827

Marchesi, C., Ossola, P., Amerio, A., Daniel, B. D., Tonna, M., & De Panfilis, C. (2016). Clinical management of perinatal anxiety disorders: A systematic review. *Journal of Affective Disorders, 190*, 543–550.

Margiotta, C., Gao, J., O'Neil, S., Vohra, D., & Zivin, K. (2022). The economic impact of untreated maternal mental health conditions in Texas. *BMC Pregnancy and Childbirth, 22*(1), 1–10.

Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology, 31*(3), 301–306.

McAndrew, A. J. (2019). Maternal perinatal depression: Emotional development in offspring from infancy to adolescence. *Early Child Development and Care, 189*(1), 168–177.

McDonough, S. C. (1993). Interaction guidance: Understanding and treating early infant–caregiver relationship disturbances. In C. H. Zeanah (Ed.), *Handbook of Infant Mental Health* (pp. 414–426). The Guilford Press.

McElwain, N. L., & Booth-LaForce, C. (2006). Maternal sensitivity to infant distress and nondistress as predictors of infant–mother attachment security. *Journal of Family Psychology, 20*(2), 247.

Miles, M. S., Funk, S. G., & Carlson, J. (1993). Parental stressor scale: Neonatal intensive care unit. *Nursing Research, 42*, 148–152.

Milgrom, J., Ericksen, J., McCarthy, R., & Gemmill, A. W. (2006). Stressful impact of depression on early mother–infant relations. *Stress and Health: Journal of the International Society for the Investigation of Stress, 22*(4), 229–238.

Miniati, M., Callari, A., Calugi, S., Rucci, P., Savino, M., Mauri, M., & Dell'Osso, L. (2014). Interpersonal psychotherapy for postpartum depression: A systematic review. *Archives of Women's Mental Health, 17*, 257–268.

Morrison, A., Polisena, J., Husereau, D., Moulton, K., Clark, M., Fiander, M., Mierzwinski-Urba, M., Clifford, T., Hutton, B., &

Rabb, D. (2012). The effect of English-language restriction on systematic review-based meta-analyses: A systematic review of empirical studies. *International Journal of Technology Assessment in Health Care, 28*(2), 138–144.

Murray, L., Arteche, A., Fearon, P., Halligan, S., Goodyer, I., & Cooper, P. (2011). Maternal postnatal depression and the development of depression in offspring up to 16 years of age. *Journal of the American Academy of Child & Adolescent Psychiatry, 50*(5), 460–470. https://doi.org/10.1016/j.jaac.2011.02.001

Murray, L., Cooper, P. J., Wilson, A., & Romaniuk, H. (2003). Controlled trial of the short- and long-term effect of psychological treatment of post-partum depression. 2. Impact on the mother–child relationship and child outcome. *The British Journal of Psychiatry, 182*(5), 420–427. https://doi.org/10.1192/bjp.182.5.420

Murray, L., Fiori-Cowley, A., Hooper, R., & Cooper, P. (1996). The impact of postnatal depression and associated adversity on early mother–infant interactions and later infant outcome. *Child Development, 67*(5), 2512–2526.

Muzik, M., Morelen, D., Hruschak, J., Rosenblum, K. L., Bocknek, E., & Beeghly, M. (2017). Psychopathology and parenting: An examination of perceived and observed parenting in mothers with depression and PTSD. *Journal of Affective Disorders, 207*, 242–250.

Newton, K., Buck, E. T., Weich, S., & Uttley, L. (2022). A review and analysis of the components of potentially effective perinatal mental health interventions for infant development and mother–infant relationship outcomes. *Development and Psychopathology, 34*(1), 37–54.

Nillni, Y. I., Mehralizade, A., Mayer, L., & Milanovic, S. (2018). Treatment of depression, anxiety, and trauma-related disorders during the perinatal period: A systematic review. *Clinical Psychology Review, 66*, 136–148.

Nylen, K. J., Moran, T. E., Franklin, C. L., & O'Hara, M. W. (2006). Maternal depression: A review of relevant treatment approaches for mothers and infants. *Infant Mental Health Journal, 27*(4), 327–343.

Out, D., Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2009). The role of disconnected and extremely insensitive parenting in the development of disorganized attachment: Validation of a new measure. *Attachment and Human Development, 11*(5), 419–443.

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews, 5*, 1–10.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery, 88*, 105906.

Parker, A. (2021). Reframing the narrative: Black maternal mental health and culturally meaningful support for wellness. *Infant Mental Health Journal, 42*(4), 502–516.

Pederson, D. R., Moran, G., Sitko, C., Campbell, K., Ghesquire, K., & Acton, H. (1990). Maternal sensitivity and the security of infant–mother attachment: A Q-sort study. *Child Development, 61*(6), 1974–1983.

Polo, A. J., Makol, B. A., Castro, A. S., Colón-Quintana, N., Wagstaff, A. E., & Guo, S. (2019). Diversity in randomized clinical trials of depression: A 36-year review. *Clinical Psychology Review, 67*, 22–35.

Poobalan, A. S., Aucott, L. S., Ross, L., Smith, W. C. S., Helms, P. J., & Williams, J. H. (2007). Effects of treating postnatal depression on mother–infant interaction and child development: Systematic review. *The British Journal of Psychiatry, 191*(5), 378–386.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385–401.

Shaw, R. J., St John, N., Lilo, E. A., Jo, B., Benitz, W., Stevenson, D. K., & Horwitz, S. M. (2013). Prevention of traumatic stress in mothers with preterm infants: A randomized controlled trial. *Pediatrics, 132*(4), e886–e894.

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology, 70*, 747–770. https://doi.org/10.1146/annurev-psych-010418-102803

Simpson, J. A., Collins, W. A., & Salvatore, J. E. (2011). The impact of early interpersonal experience on adult romantic relationship functioning: Recent findings from the Minnesota longitudinal study of risk and adaptation. *Current Directions in Psychological Science, 20*(6), 355–359.

Slade, A., Aber, J. L., Berger, B., Bresgi, I., & Kaplan, M. (2004). *Parent development interview—Revised*. Unpublished manuscript. Yale Child Study Center.

Slade, A. (2005). Parental reflective functioning: An introduction. *Attachment and Human Development, 7*(3), 269–281.

Smith, J. C., Cumming, A., & Xeros-Constantinides, S. (2010). A decade of parent and infant relationship support group therapy programs. *International Journal of Group Psychotherapy, 60*(1), 59–89. https://doi.org/10.1521/ijgp.2010.60.1.59

Sockol, L. E. (2015). A systematic review of the efficacy of cognitive behavioral therapy for treating and preventing perinatal depression. *Journal of Affective Disorders, 177*, 7–21.

Sockol, L. E., Epperson, C. N., & Barber, J. P. (2011). A meta-analysis of treatments for perinatal depression. *Clinical Psychology Review, 31*(5), 839–849.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the state-trait anxiety inventory*. Consulting Psychologists Press.

Spitzer, R. L., Williams, J. B., Gibbon, M., & First, M. B. (1992). The structured clinical interview for DSM-III-R (SCID): I: History, rationale, and description. *Archives of General Psychiatry, 49*(8), 624–629.

Sroufe, L. A. (2005). Attachment and development: A prospective, longitudinal study from birth to adulthood. *Attachment and Human Development, 7*(4), 349–367.

Stein, A., Netsi, E., Lawrence, P. J., Granger, C., Kempton, C., Craske, M. G., Nickless, A., Mollison, J., Stewart, D. A., Rapa, E., West, V., Scerif, G., Cooper, P. J., & Murray, L. (2018). Mitigating the effect of persistent postnatal depression on child outcomes through an intervention to treat depression and improve parenting: A randomised controlled trial. *The Lancet Psychiatry, 5*(2), 134–144. https://doi.org/10.1016/S2215-0366(18)30006-3

Stein, A., Woolley, H., Senior, R., Hertzmann, L., Lovel, M., Lee, J., Cooper, S., Wheatcroft, R., Challacombe, F., Patel, P., Nicol-Harper, R., Menzes, P., Schmidt, A., Juszczak, E., & Fairburn, C. G. (2006). Treating disturbances in the relationship between mothers with bulimic eating disorders and their infants: A randomized, controlled trial of video feedback. *The American Journal of Psychiatry, 163*(5), 899–906. https://doi.org/10.1176/ajp.2006.163.5.899

Stevens, N. R., Heath, N. M., Lillis, T. A., McMinn, K., Tirone, V., & Sha'ini, M. (2018). Examining the effectiveness of a coordinated perinatal mental health care model using an intersectional-feminist perspective. *Journal of Behavioral Medicine, 41*, 627–640.

Sutter-Dallay, A. L., & Riecher-Rössler, A. (2016). Psychotropic drugs and the perinatal period. In A. Sutter-Dallay, N. Glangeaud-Freudenthal, A. Guedeney, & A. Richer-Rössler (Eds.), *Joint care of parents and infants in perinatal psychiatry* (pp. 79–92). Springer.

Trapolini, T., Ungerer, J. A., & McMahon, C. A. (2008). Maternal depression: Relations with maternal caregiving representations and emotional availability during the preschool years. *Attachment and Human Development, 10*(1), 73–90.

Tronick, E., & Reck, C. (2009). Infants of depressed mothers. *Harvard Review of Psychiatry, 17*(2), 147–156.

Tsivos, Z., Calam, R., Sanders, M. R., & Wittkowski, A. (2015a). A pilot randomised controlled trial to evaluate the feasibility and acceptability of the Baby Triple P positive parenting programme in mothers with postnatal depression. *Clinical Child Psychology and Psychiatry, 20*(4), 532–554. https://doi.org/10.1177/1359104514531589

Tsivos, Z. L., Calam, R., Sanders, M. R., & Wittkowski, A. (2015b). Interventions for postnatal depression assessing the mother–infant relationship and child developmental outcomes: A systematic review. *International Journal of Women's Health, 2015*, 429–447.

Van Ee, E., Kleber, R. J., & Jongmans, M. J. (2016). Relational patterns between caregivers with PTSD and their nonexposed children: A review. *Trauma, Violence, and Abuse, 17*(2), 186–203.

Van Ijzendoorn, M. H., Juffer, F., & Duyvesteyn, M. G. (1995). Breaking the intergenerational cycle of insecure attachment: A review of the effects of attachment-based interventions on maternal sensitivity and infant security. *Journal of Child Psychology and Psychiatry, 36*(2), 225–248.

Van Ijzendoorn, M. H., Schuengel, C., & Bakermans-Kranenburg, M. J. (1999). Disorganized attachment in early childhood: Meta-analysis of precursors, concomitants, and sequelae. *Development and Psychopathology, 11*(2), 225–250.

Vaughn, B. E., & Waters, E. (1990). Attachment behavior at home and in the laboratory: Q-sort observations and strange situation classifications of one-year-olds. *Child Development, 61*(6), 1965–1973.

Wan, M. W., & Green, J. (2009). The impact of maternal psychopathology on child-mother attachment. *Archives of Women's Mental Health, 12*(3), 123–134. https://doi.org/10.1007/s00737-009-0066-5

Waters, E., Weinfield, N. S., & Hamilton, C. E. (2000). The stability of attachment security from infancy to adolescence and early adulthood: General discussion. *Child Development, 71*(3), 703–706.

Weatherston, D. J., Ribaudo, J., Michigan Collaborative for Infant Mental Health Research. (2020). The Michigan infant mental health home visiting model. *Infant Mental Health Journal, 41*(2), 166–177. https://doi.org/10.1002/imhj.21838

Willems, Y. E., Boesen, N., Li, J., Finkenauer, C., & Bartels, M. (2019). The heritability of self-control: A meta-analysis. *Neuroscience and Biobehavioral Reviews, 100*, 324–334.

Wong, O., Nguyen, T., Thomas, N., Thomson-Salo, F., Handrinos, D., & Judd, F. (2016). Perinatal mental health: Fathers—The (mostly) forgotten parent. *Asia-Pacific Psychiatry, 8*(4), 247–255.

Wright, B., Fearon, P., Garside, M., Tsappis, E., Amoah, E., Glaser, D., Allgar, V., Minnis, H., Woolgar, M., Churchill, R., McMillan, D., Fonagy, P., O'Sullivan, A., & McHale, M. (2023). Routinely used interventions to improve attachment in infants and young children: A national survey and two systematic reviews. *Health Technology Assessment, 27*(2), 1–226. https://doi.org/10.3310/IVCN8847

Zelkowitz, P., Feeley, N., Shrier, I., Stremler, R., Westreich, R., Dunkley, D., Steele, R., Rosberger, Z., Lefebvre, F., & Papageorgiou, A. (2011). The cues and care randomized controlled trial of a neonatal intensive care unit intervention: Effects on maternal psychological distress and mother–infant interaction. *Journal of Developmental and Behavioral Pediatrics, 32*(8), 591–599. https://doi.org/10.1097/DBP.0b013e318227b3dc

Clinical Psychology & Psychotherapy

WILEY

RESEARCH ARTICLE OPEN ACCESS

A Novel Transdiagnostic Approach to the Prevention of Eating Disorders Using Virtual Reality: Preliminary Evaluation of the H.O.M.E. Intervention

Valentina Gardini | Silvana Grandi | Elena Tomba

Department of Psychology, University of Bologna, Bologna, Italy

Correspondence: Elena Tomba (elena.tomba@unibo.it)

Received: 17 October 2024 | **Revised:** 20 December 2024 | **Accepted:** 16 January 2025

Keywords: eating disorder prevention | emotion regulation | experiential avoidance | psychological flexibility | transdiagnostic approach | virtual reality

ABSTRACT

"Virtual reality" (VR) has been used effectively in clinical psychology to improve existing treatments and prevention protocols for many psychopathologies, including eating disorders (EDs). However, no VR software was developed to concurrently tackle dysfunctional eating behaviours and three third wave cognitive–behavioural transdiagnostic factors linked to EDs: psychological inflexibility, emotion dysregulation and experiential avoidance. This preliminary study is aimed at evaluating potential effects of a new VR-based preventive intervention (H.O.M.E.—How to Observe and Modify Emotions) in improving selected outcomes of transdiagnostic factors and dysfunctional eating behaviours in the general population (GP) with ED risk compared to a waiting list (WL). $N = 40$ GP participants with ED risk were screened (using the SCOFF and Eating Disorder Examination Questionnaire (EDE-Q)) and randomised into VR ($n = 20$) or WL (n = 20) groups. Before and after intervention and at 3- and 6-month follow-up, participants completed EDE-Q, Difficulties in Emotion Regulation Scale—brief version (DERS-16) for emotion dysregulation, Acceptance and Action Questionnaire II (AAQ-II) for psychological inflexibility and Multidimensional Psychological Flexibility Inventory (MPFI) experiential avoidance scale. H.O.M.E. improved scores in all EDE-Q subscales (EDE-Q-total: $p = 0.003$; EDE-Q-restraint: $p = 0.028$; EDE-Q-eating concerns: $p = 0.035$; EDE-Q-shape concerns: $p = 0.003$; EDE-Q-weight concerns: $p = 0.023$), AAQ-II ($p = 0.005$), DERS-16-total ($p = 0.006$), DERS-16-difficulty in engaging in goal-directed behaviours ($p = 0.008$), and DERS-16-limited access to emotion regulation strategies ($p = 0.001$), with results greater than WL and maintained at follow-ups. Results showed potential for H.O.M.E. to represent a feasible tool to prevent EDs. Given the similarity between VR and technologies used in everyday life, H.O.M.E. may help in engaging young individuals with ED risk towards psychological support before ED onset.

1 | Introduction

"Virtual reality" (VR) includes an array of technologies that allow individuals to experience and interact with computer-generated three-dimensional environments and objects through a head-mounted display (Maples-Keller et al. 2017). The application of VR in clinical psychological contexts as an instrument to offer patients a systematic and controlled exposure therapy without the complications of in vivo exposure (Wiederhold and Wiederhold 2005) has been supported and VR showed to improve existing cognitive–behavioural treatments (CBTs) and prevention protocols especially in anxiety, psychotic, substance-related and eating disorders (EDs) (Emmelkamp and Meyerbröker 2021; Freeman et al. 2017). Recently, VR programmes capable of being administered across several psychiatric diagnoses have been designed (Colombo et al. 2021). According to the transdiagnostic approach, many psychological disorders and comorbid psychiatric diagnoses

Summary

- Virtual reality (VR) demonstrated to be able to enhance treatments for psychological disorders, including eating disorders (EDs), but transdiagnostic VR software are scarce.
- H.O.M.E. (How to Observe and Modify Emotions) is the first VR-based intervention designed to target dysfunctional eating behaviours and three transdiagnostic factors linked to EDs: psychological inflexibility, emotion dysregulation and experiential avoidance.
- This preliminary study demonstrated that H.O.M.E. managed to improve some of the selected outcomes (i.e., dysfunctional eating behaviours, psychological flexibility and emotion regulation difficulties) in participants with ED risk.
- Preliminary results of the H.O.M.E. VR-based intervention were greater than the waiting list and maintained at 3- and 6-month follow-ups.
- These preliminary results show that the H.O.M.E. VR-based intervention has good feasibility, with no drop-out rates and promising changes on selected outcomes, but its efficacy needs to be tested in bigger clinical trials with active comparison groups (e.g., CBT-E).

are maintained or caused by similar cognitive and/or behavioural processes (Harvey et al. 2004; Mansell et al. 2008), defined as transdiagnostic factors. While research has yet to come to a consensus regarding a fixed number of specific transdiagnostic factors leading to the main psychological disorders, studies showed how, especially within the more recent third wave cognitive behavioural theory of mental disorders (Hayes and Hofmann 2017), several psychiatric diagnoses (e.g., mood and anxiety disorders, substance use disorders and EDs) are all characterised by psychological inflexibility (Hayes et al. 2006), emotion dysregulation (Aldao, Nolen-Hoeksema, and Schweizer 2010; Fairholme et al. 2010) and experiential avoidance (Chawla and Ostafin 2007; Hayes et al. 1996).

One of the clinical fields in which transdiagnostic approaches found the most evidence is EDs (Atwood and Friedman 2020). According to the main transdiagnostic cognitive–behavioural theoretical model for EDs (Fairburn, Cooper, and Shafran 2003), all EDs are characterised by similar ED features and maintained by shared nonspecific ED factors. Third wave cognitive–behavioural transdiagnostic factors (i.e., psychological inflexibility, emotion dysregulation and experiential avoidance) would also play a role in the onset and maintenance of EDs both in the clinical (Rawal, Park, and Williams 2010; Westwood et al. 2017) and general population (GP) (Della Longa and De Young 2018; Han and Pistole 2014; Masuda et al. 2010). Therefore, their reduction could help improve dysfunctional eating behaviours (Bluett et al. 2016; Holmqvist Larsson et al. 2020; Skinner, Rojas, and Veilleux 2017).

Despite researches supporting both transdiagnostic treatments (Atwood and Friedman 2020) and VR-based interventions for EDs (Riva, Malighetti, and Serino 2021), a systematic review (Gardini et al. 2022) found that so far no VR software has been designed to concurrently tackle psychological inflexibility, experiential avoidance and emotion regulation (ER). Moreover, no research has focused on the use of VR in the prevention of EDs by targeting these three transdiagnostic factors. Indeed, existing prevention strategies often focus on psychoeducation, cognitive–behavioural techniques and school-based interventions aimed at addressing body image concerns and reducing eating disorder behaviours (Le et al. 2017; Shaw, Stice, and Becker 2009; Stice, Onipede, and Marti 2021). While traditional prevention approaches have shown efficacy in some populations, they often face limitations, including low engagement rates and motivational barriers (Atkinson and Wade 2013; Harrer et al. 2020; Watson et al. 2016). Using VR in ED prevention may help solve these issues and motivate at-risk individuals towards change, similar to what has been observed in other populations (Riva 2022). Immersivity and realism offered by VR allow individuals to experience a sense of "emotional presence" comparable to reality in inducing emotional responses (Gorini et al. 2010; Riva, Molinari, and Vincelli 2002) which can be used to provide immersive experiences to at-risk individuals that could improve the aforementioned transdiagnostic factors linked to EDs.

Therefore, as part of the published H.O.M.E. (How to Observe and Modify Emotions) research protocol (Gardini et al. 2023), this preliminary study had three major aims: (1) to explore the feasibility of a VR-based intervention conducted using an innovative transdiagnostic VR software (the H.O.M.E. software) (Gardini et al. 2023) in improving dysfunctional eating behaviours and transdiagnostic factors (i.e., ER difficulties, psychological inflexibility and experiential avoidance) between pre- and postintervention in the GP at risk for EDs in order to evaluate whether its efficacy can be tested in bigger clinical trials, (2) to observe the potential effects of the H.O.M.E. VR-based intervention compared to a waiting list condition before moving forward to testing it when used against to or in combination with traditional active psychological interventions (e.g., CBT-E) and (3) to evaluate whether the changes achieved by the H.O.M.E. VR-based intervention in the current study were maintained at 3- and 6-month follow-up.

2 | Materials and Methods

2.1 | Ethical Approval

This project was initially approved on 09 October 2020 (Protocol No. 224366) by the Ethical Committee for Psychological Research of University of Bologna and subsequently revised on 14 December 2021 (Protocol No. 0314877).

2.2 | Recruitment Method

The presented preliminary study investigating the potential effects of the H.O.M.E. VR-based intervention protocol (Gardini et al. 2023) included individuals from the GP at risk for EDs. Participants were recruited via social media (e.g., Facebook, Instagram). Individuals interested in taking part in the research were informed about aims and characteristics

of the study, and informed consent was requested. Inclusion criteria were (a) 18–60 years of age, (b) screening at risk for EDs according to screening instruments reported below, (c) no prior diagnosis of EDs and (d) no health condition that better explains ED symptomatology. Exclusion criteria included (a) lack of capacity to consent for research, (b) current or former ED or other psychiatric diagnosis according to DSM-5-TR (American Psychiatric Association 2022) and (c) using medical devices (e.g., pacemaker, hearing aids) or having medical diagnoses (e.g., vertigo, vision impairments) interfering with VR use. Eligible participants were subsequently invited to take part in the study. Those agreeing underwent a brief clinical interview conducted by a clinical psychologist (V.G.) based on DSM-5-TR (American Psychiatric Association 2022) criteria to exclude presence of EDs or other psychiatric disorders. The clinical interview was performed before participants took part in the experimental phase of the study (see the Research Design section) and information was provided about their screening results, and if needed, they were informed about the psychological support options available within and outside the research and the contact details of the closest National Health Service ED Centre were provided.

2.3 | Research Design

The present preliminary study stems from the H.O.M.E. published research protocol (Gardini et al. 2023), which was subsequently divided in three studies conducted on GP participants with ED risk with different aims: (1) to preliminarily explore the potential effects and feasibility of the H.O.M.E. six-session VR-based intervention protocol in the improvement of the selected outcomes of dysfunctional eating behaviours and transdiagnostic factors in the GP at risk for EDs, also when compared to a waiting list; (2) to evaluate the quality of the H.O.M.E. software using quantitative measures (e.g., self-report questionnaires measuring user experience, sense of presence, motion sickness, possible elicited distress) and (3) to evaluate the quality of the H.O.M.E. VR-based intervention from the participants' point of view using qualitative methods (thematic analysis on interviews conducted on participants completing the protocols).

The presented preliminary study specifically focused on the first aim of the aforementioned research protocol.

To achieve this aim, suitable participants screened as at risk for EDs during a *screening phase* (see the Measures section) were subsequently contacted via e-mail or telephone within 7 days and invited to take part in the *experimental phase* of the study.

In the *experimental phase*, participants were randomised in the intervention (VR) or control (waiting list) group following a "block randomisation" method (Kim and Shin 2014). Participants in the intervention (VR) group underwent six 30-min sessions of the VR-based intervention administered using the software H.O.M.E. at the presence of a clinical psychologist (V.G.) (using the protocol described in Gardini et al. 2023). Controls (waiting list) did not receive any intervention during the study, but they received minimal attention from the clinical psychologists (V.G., E.T.) through phone check-up sessions monitoring symptoms and general well-being. They were also offered to receive the VR-based intervention after the end of the study.

To test whether changes were maintained over time, both groups were recontacted (V.G., E.T.) for *follow-ups* after 3 and 6 months after conclusion of the experimental phase.

Immediately before the first session of the experimental phase (T0), at conclusion of the experimental phase (immediately after the last VR session or WL check-up (T1)) and at 3- and 6-month follow-ups (T2, T3), participants of both groups were asked to complete online (using the platform Qualtrics) a series of psychometric self-report questionnaires (see the Measures section).

2.4 | VR Software: H.O.M.E.

H.O.M.E. is a transdiagnostic VR software developed by a group of clinical psychology researchers from the Department of Psychology of University of Bologna (Gardini et al. 2023). H.O.M.E. can be used by connecting a computer to an Oculus Quest headset. The software allows users to move around a virtual environment consisting of a house with four rooms (i.e., kitchen/living room, bathroom, bedroom and a study) and a garden (Figure 1) and to interact with objects relevant for specific mental health issues (e.g., comfort foods for people with eating-related issues; alcohol, pills and cigarettes for addictions; a computer with a gambling interface for gambling addictions; videogames and smartphones for technology-related addictions). The software includes two phases (Figure 2): an assessment phase where users can report the emotion they associate to each object and its intensity and an intervention phase based on the cognitive–behavioural model in which users can access a 'box of psychological resources' including pictures and descriptions of emotional, social and behavioural resources that can be used to face the emotions previously attributed to the objects. The software is meant to be transdiagnostic since it is designed to tackle the following transdiagnostic factors: ER strategies, psychological flexibility and emotional avoidance; and it is meant to be used in the prevention and treatment of several psychiatric disorders (e.g., EDs, addictions, substance use disorders) under the instructions of a clinical psychologist. In the present study, objects that could represent a trigger for people at risk for EDs (e.g., food in the kitchen area) were predominantly used.

The characteristics and aims of the H.O.M.E. software as well as the content of the H.O.M.E. six-session VR-based intervention used for the present study were published in a separate article (Gardini et al. 2023) and are summarised in Figure 3.

2.5 | Measures

To detect ED risk, recruited GP participants completed the following screening questionnaires:

- A nonstandardised *self-report questionnaire* to collect sociodemographic (age, marital status, educational level, occupational status) and clinical data (body mass index (BMI),

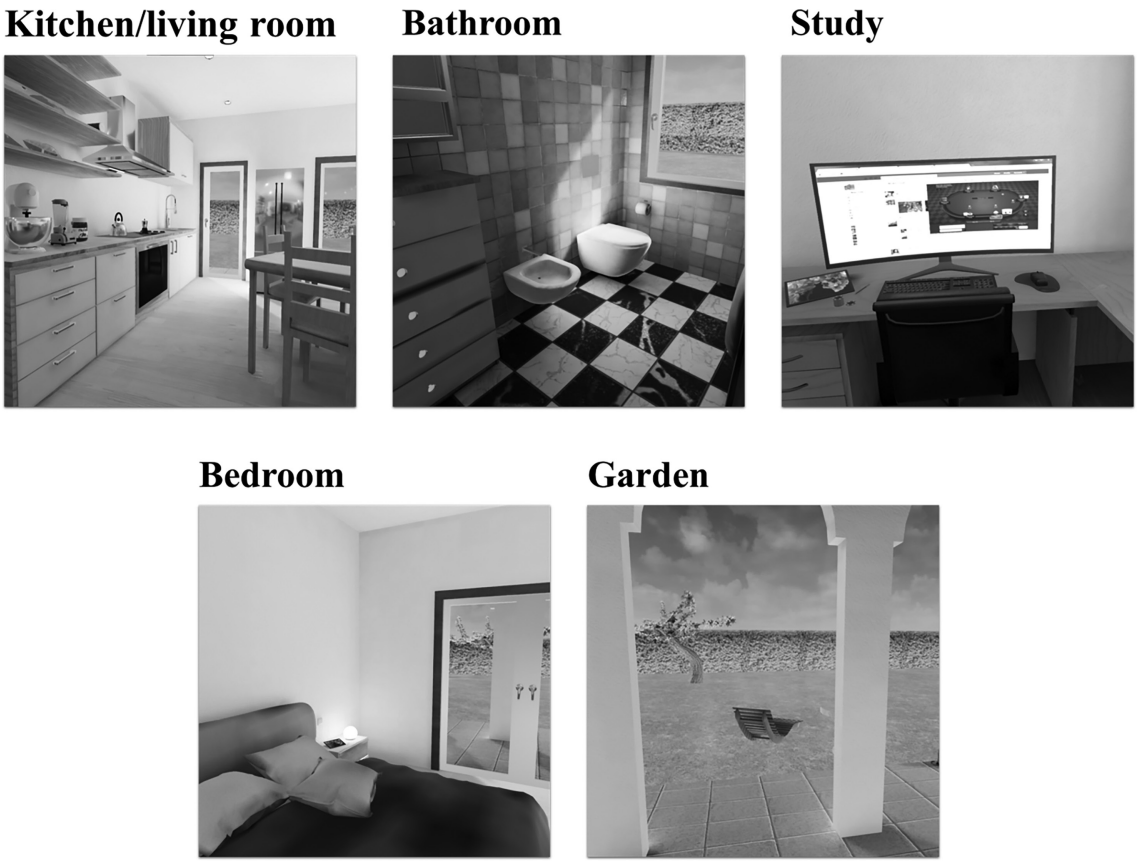


FIGURE 1 | H.O.M.E. software virtual environment (Gardini et al. 2023).

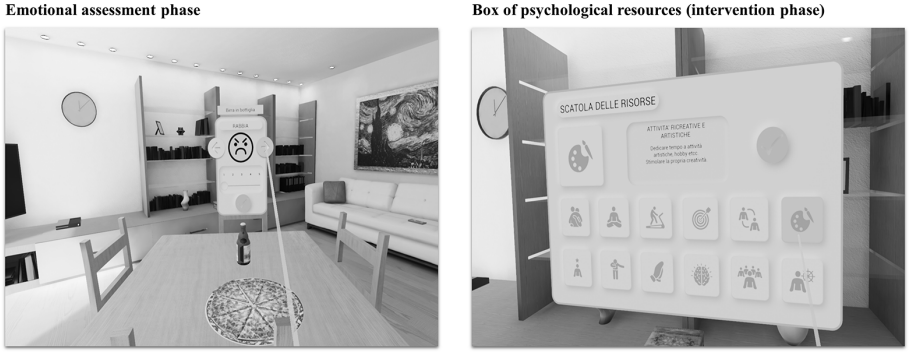


FIGURE 2 | Features available in the H.O.M.E. software (Gardini et al. 2023).

former ED or another DSM-5 diagnosis, clinical conditions and/or medical devices interfering with VR) and to investigate participants' daily use of technological devices (e.g., computer, videogames)

- The *SCOFF questionnaire* (Morgan, Reid, and Lacey 1999; Pannocchia et al. 2011), a psychometric five-item self-report screening questionnaire detecting ED risk. A cut-off of >2 indicates a possible ED case (Cronbach's $\alpha = 0.64$ in Italian version)
- The *Eating Disorder Examination Questionnaire* (EDE-Q 6.0) (Calugi et al. 2017; Fairburn and Beglin 1994). The EDE-Q produces a global score and four subscales for the past 28 days: eating concern, shape concern, weight concern and restraint (Cronbach's $\alpha = 0.70–0.83$ for subscale; ≥ 0.90 for the global score) (Peterson et al. 2007). Scores above 1.56 are associated with ED risk in the GP (Ekeroth and Birgegård 2014)

A brief clinical interview conducted by a clinical psychologist (V.G.) based on DSM-5-TR (American Psychiatric Association 2022) criteria was also performed before the experimental phase of the study on people screening as at risk at the SCOFF and EDE-Q to exclude presence of EDs or other DSM-5-TR psychiatric disorders.

At T0, T1, T2 and T3, recruited participants undergoing the VR-based interventions or waiting list filled in the following

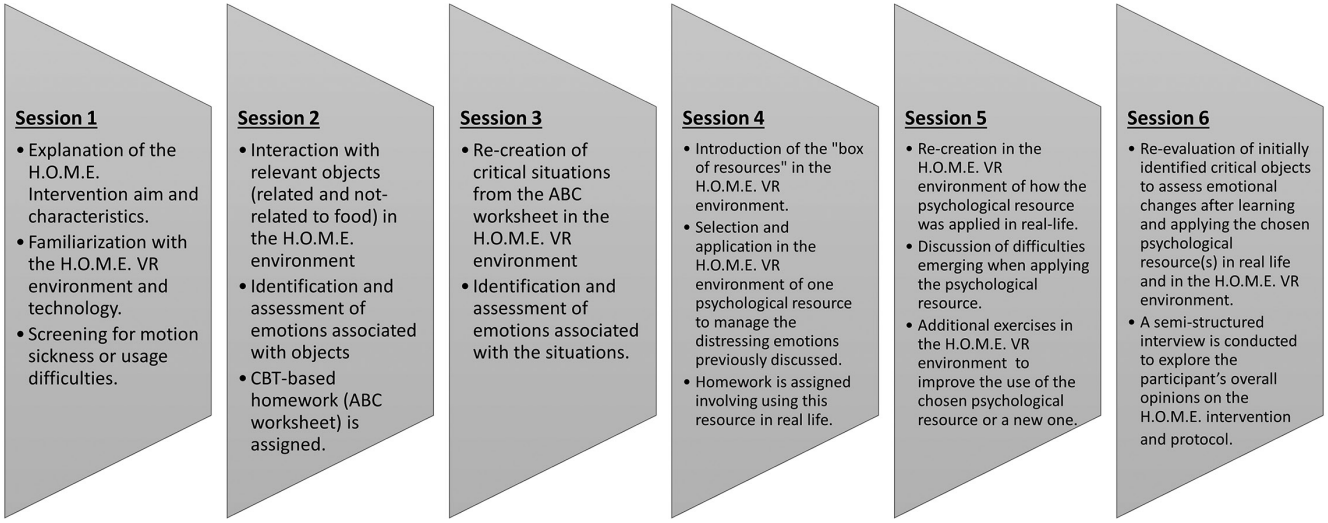


FIGURE 3 | Brief summary of the content and aims of the H.O.M.E. VR-based intervention sessions (Gardini et al. 2023).

psychometric questionnaires to evaluate changes in dysfunctional eating behaviours and transdiagnostic factors:

- The EDE-Q 6.0 (Calugi et al. 2017; Fairburn and Beglin 1994)
- The *Difficulties in Emotion Regulation Scale—brief version* (DERS-16) (Bjureberg et al. 2016), a 16-item self-report questionnaire evaluating five ER difficulties: nonacceptance of emotional responses (DERS-16-NA), difficulty in engaging in goal-directed behaviours (DERS-16-GO), impulse control difficulties (DERS-16-IMP), limited access to ER strategies (DERS-16-STR) and lack of emotional clarity (DERS-16-CL) (Cronbach's $\alpha = 0.87–0.96$) (Hallion et al. 2018)
- The *Acceptance and Action Questionnaire II* (AAQ-II) (Bond et al. 2011; Pennato et al. 2013), a seven-item self-report questionnaire assessing psychological flexibility (Cronbach's $\alpha = 0.83$) (Pennato et al. 2013)
- The *experiential avoidance scale* (MPFI-EA) from the *Multidimensional Psychological Flexibility Inventory* (MPFI) (Landi et al. 2021; Rolffs, Rogge, and Wilson 2018), a five-item self-report scale evaluating experiential avoidance (Cronbach's $\alpha = 0.91$ for Italian version)

2.6 | Data Analyses

Data analyses were performed using SPSS Statistic vrs.25. Descriptive statistics were run to analyse sociodemographic (age, marital status, educational level, occupational status) and clinical characteristics (weight history, former ED diagnosis, etc.), as well as mean SCOFF, EDE-Q, DERS-16, AAQ-II and MPFI-EA. The *t*-tests for continuous variables and chi-square tests for categorical variables were performed to test differences between those who completed the screening questionnaire and did not accept to take part in the research and those who accepted.

The *t*-tests and chi-square tests were also used to test preintervention/baseline differences between experimental (VR) and control (waiting list) groups.

To test the feasibility of the H.O.M.E. VR-based intervention, percentage of drop-out cases was calculated for the experimental (VR) group.

To test the potential effects of the H.O.M.E. VR-based intervention and to also compare its effects (in terms of changes produced in EDE-Q, DERS-16, AAQ-II and MPFI-EA between pre- and postintervention and between postintervention and follow-ups) to those of the control (waiting list) condition, repeated-measures ANOVAs were performed using time (T0, T1, T2, T3) as within-subject factor and group (VR or waiting list) as between-subject factor with post hoc contrast analyses. Given the preliminary and exploratory nature of the study, multiple comparison corrections (e.g., Bonferroni correction) were not performed (Rothman 1990) and effect sizes (Cohen's *d*) were reported to better understand the magnitude of the potential changes achieved.

3 | Results

3.1 | Description of the Sample

$N = 189$ people from the GP completed the online screening questionnaires on the platform Qualtrics. Of these, the number of people at risk for EDs was determined based on the *SCOFF Questionnaire* (Morgan, Reid, and Lacey 1999; Pannocchia et al. 2011) and EDE-Q 6.0 (Calugi et al. 2017; Fairburn and Beglin 2008) cut-offs of < 2 and < 1.56 (Ekeroth and Birgegård 2014), respectively. A total of $n = 104$ (55.03%) people screened with ED risk were all contacted to take part in the research. According to the brief clinical interview based on DSM-5-TR criteria, $n = 8$ (7.69%) people met criteria for an ED diagnosis and $n = 12$ (11.54%) for other DSM-5-TR diagnoses and were excluded from the present study. $N = 47$ people (45.19%), instead, refused to take part in the research for personal or logistical reasons. The final sample included $n = 40$ (38.46%) eligible people at risk for EDs that accepted to take part in the study and were randomised to the experimental (undergoing the H.O.M.E. VR-based preventive intervention protocol) ($n = 20$) or

control (waiting list) ($n = 20$) group using a block randomisation method.

Descriptive analyses revealed that $n = 64$ participants that did not take part in the study were females, with a mean age of 23.6 ± 8.08. Mean SCOFF score was 3.33 ± 0.99, while mean EDE-Q score was 3.20 ± 1.54. Independent sample *t*-test revealed that there was no significant difference between those deciding to take part in the research and those who did not participate (see Table S1).

When considering the sample that took part in the research, both the experimental (VR) and control (waiting list) groups were made up of only female participants, with similar mean ages (VR: 24.35 ± 6.02; waiting list: 24.75 ± 9.08). The majority of the sample in both groups were single (VR: $n = 18$, 90%; waiting list: $n = 18$, 90%) and university students (VR: $n = 15$, 75%; waiting list: $n = 16$, 80%). All participants in both groups were familiar with the use of technology (e.g., smartphones, computers, videogames) in their daily lives.

When considering dysfunctional eating behaviours, participants of both groups reported the highest scores in the EDE-Q-shape concerns (VR: 4.52 ± 0.93; waiting list: 4.27 ± 1.26) and EDE-Q-weight concerns (VR: 3.92 ± 1.11; waiting list: 3.30 ± 1.40) subscales.

Table 1 summarises sociodemographic and clinical characteristics of both groups, as well as the statistical indexes relative to the independent sample *t*-test and chi-square tests run to evaluate the presence of possible sociodemographic or clinical differences between groups at baseline (T0).

As it can be observed from Table 1, no significant difference was found between groups at baseline in any of the variables,

TABLE 1 | Sociodemographic and clinical characteristics of the VR and waiting list group samples and differences between groups.

	H.O.M.E. VR group ($n = 20$) $M \pm$ SD *or* % (n=)	**Waiting list group ($n = 20$) $M \pm$ SD *or* % (n=)**	***t* (df)**	**$\chi^2_{(df)}$**	***p***
Gender	100% ($n = 20$) females	100% ($n = 20$) females	—	—	—
Age	24.35 ± 6.02 years	24.75 ± 9.08 years	$-0.16_{(38.0)}$	—	0.870
Marital status	90.0% ($n = 18$) single 5% ($n = 1$) married 5% ($n = 1$) divorced	90.0% ($n = 18$) single 5% ($n = 1$) married 5% ($n = 1$) divorced	—	—	—
Educational level	35% ($n = 7$) high school diploma 65% ($n = 13$) university degree	50% ($n = 10$) high school diploma 50% ($n = 10$) university degree	—	$2.39_{(2)}$	0.303
Occupation	75% ($n = 15$) university students 20% ($n = 4$) employed 5% ($n = 1$) unemployed	80% ($n = 16$) university students 20% ($n = 4$) employed	—	$3.03_{(4)}$	0.552
BMI	25.72 ± 3.19	22.93 ± 3.28	$3.33_{(38.0)}$	—	0.087
SCOFF	3.65 ± 1.09	3.55 ± 0.95	$0.31_{(38.0)}$	—	0.758
EDE-Q-total	3.48 ± 1.07	3.18 ± 1.18	$0.85_{(38.0)}$	—	0.400
EDE-Q-restraint	2.79 ± 1.68	2.75 ± 1.38	$0.08_{(38.0)}$	—	0.935
EDE-Q-eating concerns	2.69 ± 1.44	2.36 ± 1.33	$0.75_{(38.0)}$	—	0.456
EDE-Q-shape concerns	4.52 ± 0.93	4.27 ± 1.26	$0.72_{(38.0)}$	—	0.479
EDE-Q-weight concerns	3.92 ± 1.11	3.3 ± 1.40	$1.48_{(38.0)}$	—	0.147
DERS-16-total	47.45 ± 15.77	46.95 ± 18.42	$0.09_{(38.0)}$	—	0.927
DERS-16-*clarity*	5.75 ± 2.34	5.15 ± 2.76	$0.74_{(38.0)}$	—	0.462
DERS-16-*goals*	10.35 ± 3.30	9.80 ± 3.82	$0.49_{(38.0)}$	—	0.629
DERS-16-*impulse*	7.55 ± 3.96	7.95 ± 3.76	$-0.33_{(38.0)}$	—	0.745
DERS-16-*strategies*	14.70 ± 5.56	15.10 ± 6.50	$-0.21_{(38.0)}$	—	0.836
DERS-16-*nonacceptance*	9.10 ± 3.14	8.95 ± 3.94	$0.13_{(38.0)}$	—	0.895
AAQ-II	28.15 ± 9.56	25.85 ± 10.97	$0.71_{(38.0)}$	—	0.484
MPFI-EA	3.28 ± 0.99	3.34 ± 1.10	$-0.18_{(38.0)}$	—	0.852

Abbreviations: AAQ-II, Acceptance and Action Questionnaire II; BMI, body mass index; DERS-16, Difficulties in Emotion Regulation Scale, short version; EDE-Q, Eating Disorder Examination Questionnaire; M, mean; MPFI-EA, Multidimensional Psychological Flexibility Inventory-experiential avoidance scale; SD, standard deviation.

including dysfunctional eating behaviours (SCOFF, EDE-Q) and transdiagnostic factors (DERS-16, AAQ-II, MPFI-EA).

3.2 | Feasibility and Effects of the H.O.M.E. Intervention and Comparison With Waiting List

The H.O.M.E. VR-based prevention intervention showed good feasibility, with no participant dropping out after undergoing at least one session of the intervention.

These preliminary results showed that the H.O.M.E. VR-based intervention was also capable of producing changes between pre- and postintervention (T0–T1 changes) in participants undergoing it. In particular, mixed-model repeated-measures ANOVAs with post hoc contrast revealed significant improvements and/or medium effect size differences between T0 and T1 in EDE-Q-total ($T0 = 3.48 \pm 1.07$, $T1 = 2.89 \pm 1.02$, $p = 0.003$, $d = 0.56$) and all EDEQ subscales (EDE-Q-restraint: $T0 = 2.79 \pm 1.68$, $T1 = 2.22 \pm 1.19$, $p = 0.028$, $d = 0.39$; EDE-Q-eating concerns: $T0 = 2.69 \pm 1.44$, $T1 = 2.20 \pm 1.22$, $p = 0.035$, $d = 0.37$; EDE-Q-shape concerns: $T0 = 4.52 \pm 0.93$, $T1 = 3.74 \pm 1.26$, $p = 0.003$, $d = 0.70$; EDE-Q-weight concerns: $T0 = 3.92 \pm 1.11$, $T1 = 3.38 \pm 1.18$, $p = 0.023$, $d = 0.47$), AAQ-II ($T0 = 28.15 \pm 9.56$, $T1 = 24.35 \pm 10.43$, $p = 0.005$, $d = 0.38$), DERS-16-total ($T0 = 47.45 \pm 15.77$, $T1 = 41.20 \pm 16.20$, $p = 0.006$, $d = 0.39$), DERS-16-difficulty in engaging in goal-directed behaviours ($T0 = 10.35 \pm 3.30$, $T1 = 9.10 \pm 3.68$, $p = 0.008$, $d = 0.36$) and DERS-16-limited access to ER strategies ($T0 = 14.70 \pm 5.56$, $T1 = 12.30 \pm 5.20$, $p = 0.001$, $d = 0.45$) which were observed. Changes achieved at postintervention (T1) were maintained at both 3- (T2) and 6-month (T3) follow-ups, with a lack of significant and/or medium effect size differences between T1–T2 and T1–T3 in all the improved aforementioned variables, with the only exception of DERS-16-difficulty in engaging in goal-directed behaviours ($T1 = 9.10 \pm 3.68$, $T3 = 7.85 \pm 3.31$, $p = 0.008$, $d = 0.36$) and DERS-16-nonacceptance of emotional responses ($T1 = 8.40 \pm 3.49$, $T3 = 7 \pm 3.28$, $p = 0.006$, $d = 0.41$) which improved between postintervention (T1) and 6-month follow-up (T3).

Table S2 shows mean levels of dysfunctional eating behaviours (EDE-Q) and transdiagnostic factors (DERS-16, AAQ-II and MPFI-EA) at each assessment point, separately for each group.

Results from repeated-measures mixed-model ANOVAs and statistical data are summarised in Tables 2 and 3.

Looking at the results from the ANOVAs, all the changes reported by the H.O.M.E. VR-based intervention also resulted to be greater than those of the waiting list condition (significant time × group interactions, p levels and effect sizes are reported in Tables 2 and 3).

The waiting list also did not report any significant improvement in dysfunctional eating behaviours or any of the transdiagnostic factors between T0 and T1, but levels of EDE-Q-restraint ($T1 = 2.85 \pm 1.27$, $T3 = 3.33 \pm 1.23$, $p = 0.046$, $d = -0.38$), DERS-16-total ($T1 = 47.75 \pm 15.13$, $T3 = 52.30 \pm 12.76$, $p = 0.025$, $d = -0.33$), DERS-16-difficulty in engaging in goal-directed behaviours ($T1 = 10.30 \pm 3.47$, $T3 = 11.40 \pm 2.70$, $p = 0.018$, $d = -0.35$) and DERS-16-limited access to ER strategies ($T1 = 15 \pm 5.91$, $T3 = 16.70 \pm 4.95$, $p = 0.019$, $d = -31$) even worsened between T1 and 6-month (T3) follow-up as represented by significant T1–T3 differences.

4 | Discussion

This preliminary study was part of the published H.O.M.E. research protocol (Gardini et al. 2023), and it is aimed at exploring the feasibility and potential effects of the transdiagnostic VR-based preventive intervention run through the H.O.M.E. software in improving the selected outcomes of transdiagnostic factors (i.e., psychological inflexibility, emotion dysregulation and experiential avoidance) and dysfunctional eating behaviours in the GP with ED risk also when compared to a waiting list. This study was conducted in order to explore and evaluate whether the H.O.M.E. VR-based intervention protocol has potential to be tested in bigger clinical trials and when used against to or in combination with active traditional psychological interventions (e.g., CBT-E).

Through a battery of online self-report screening questionnaires, people from the GP at risk of developing EDs were detected. However, a wide range of respondents that were contacted decided not to take part in the intervention. This is not surprising, as people suffering from EDs are often reluctant or motivated to seek help, due to the egosyntonic nature of these psychopathologies especially in the early stages (Roncero et al. 2013).

Regarding the sample of people who accepted to take part in the study, it consisted of mainly single women and university students, which is unsurprising considering that young women are more frequently affected by EDs (Jacobi et al. 2004). These sociodemographic characteristics are also consistent with samples of prevention interventions, which are often targeted at young adults (Watson et al. 2016). In fact, epidemiological data underline that it is precisely this segment of the population that is most at risk of developing EDs (Mastrobattista and Pacifici 2022) and, therefore, it is necessary to identify effective preventive intervention strategies that can be carried out on this population (Schwartz et al. 2019).

The importance of body and weight in today's society, which encourages young women in particular to pursue an ideal of thinness often unhealthy (Culbert, Racine, and Klump 2015), is reflected in the high levels of body shape concerns and weight concerns measured through the EDE-Q, which were the highest subscales in all samples. The levels of difficulties in ER and psychological inflexibility reported by participants were also higher than in other studies conducted on the GP without ED risk (McClure et al. 2022; Pennato et al. 2013). Indeed, as underlined by the literature (Morton et al. 2020; Prefit, Cândea, and Szentagotai-Tătar 2019), dysfunctional eating behaviours are often associated with difficulties in ER and psychological inflexibility. Levels of experiential avoidance, instead, were consistent with GP levels found in the literature (Landi et al. 2021).

Regarding the specific results obtained by this preliminary study, the H.O.M.E. VR-based preventive intervention showed

TABLE 2 | Results of the repeated-measures mixed-model ANOVAs investigating changes in the H.O.M.E. intervention and comparisons with waiting list (EDEQ, AAQ-II, MPFI-EA).

		EDEQ-total	EDEQ-restraint	EDEQ-eating concerns	EDEQ-shape concerns	EDEQ-weight concerns	AAQ-II-total	MPFI-EA-total
	Group ($F_{(df)}$; p; η^2_p)	$0.76_{(38.1)}$; 0.389; 0.02	$1.77_{(38.1)}$; 0.191; 0.04	$0.30_{(38.1)}$; 0.589; 0.01	$1.05_{(38.1)}$; 0.313; 0.03	$0.010_{(38.1)}$; 0.919; 0.00	$0.49_{(38.1)}$; 0.487; 0.01	$2.12_{(38.1)}$; 0.153; 0.05
	Time ($F_{(df)}$; p; η^2_p)	$2.32_{(38.1)}$; 0.079; 0.06	$1.39_{(38.1)}$; 0.249; 0.045	$2.48_{(38.1)}$; 0.065; 0.06	**$2.98_{(38.1)}$; 0.035; 0.07**	$0.91_{(38.1)}$; 0.440; 0.02	$1.31_{(38.1)}$; 0.275; 0.03	$1.15_{(38.1)}$; 0.333; 0.03
	Time*group ($F_{(df)}$; p; η^2_p)	**$5.37_{(38.1)}$; 0.002; 0.12**	**$2.00_{(38.1)}$; 0.017; 0.09**	**$3.84_{(38.1)}$; 0.012; 0.09**	**$3.58_{(38.1)}$; 0.016; 0.09**	**$3.54_{(38.1)}$; 0.017; 0.09**	**$5.90_{(38.1)}$; < 0.001; 0.13**	$0.88_{(38.1)}$; 0.452; 0.02
Post hoc								
VR MD (p; d)	T0×T1	**0.59 (0.003; 0.56)**	**0.57 (0.028; 0.39)**	**0.49 (0.035; 0.37)**	**0.78 (0.003; 0.70)**	**0.54 (0.023; 0.47)**	**3.80 (0.005; 0.38)**	−0.12 (0.679; −0.11)
	T0×T2	**0.67 (0.002; 0.54)**	0.42 (0.098; 0.27)	**0.71 (0.002; 0.47)**	**0.87 (0.006; 0.61)**	**0.67 (0.013; 0.47)**	**4.40 (0.005; 0.41)**	0.20 (0.545; 0.18)
	T0×T3	**0.62 (0.005; 0.54)**	0.49 (0.092; 0.33)	**0.69 (0.005; 0.49)**	**0.81 (0.004; 0.64)**	0.51 (0.062; 0.41)	**3.85 (0.017; 0.37)**	0.28 (0.367; 0.26)
	T1×T2	0.07 (0.657; 0.07)	−0.15 (0.507; −0.12)	0.22 (0.176; 0.16)	0.09 (0.679; 0.06)	0.13 (0.545; 0.09)	0.60 (0.551; 0.05)	0.32 (0.208; 0.27)
	T1×T3	0.03 (0.381; 0.03)	−0.08 (0.732; −0.07)	0.20 (0.310; 0.16)	0.03 (0.866; 0.02)	−0.03 (0.888; −0.02)	0.05 (0.968; 0.01)	0.40 (0.094; 0.35)
	T2×T3	−0.04 (0.683; −0.04)	0.07 (0.702; 0.05)	−0.02 (0.848; −0.01)	−0.06 (0.544; −0.04)	−0.16 (0.293; −0.10)	−0.55 (0.606; −0.05)	0.08 (0.428; 0.07)
Waiting listMD (p; d)	T0×T1	−0.13 (0.492; −0.11)	−0.10 (0.690; −0.08)	−0.10 (0.658; −0.08)	−0.06 (0.799; −0.05)	−0.26 (0.260; −0.20)	−1.40 (0.274; −0.13)	−0.45 (0.126; −0.41)
	T0×T2	−0.04 (0.848; −0.03)	−0.02 (0.936; −0.01)	0.02 (0.927; 0.02)	0.02 (0.950; 0.02)	−0.17 (0.512; −0.14)	−1.30 (0.385; −0.12)	−0.20 (0.545; −0.19)
	T0×T3	−0.31 (0.154; −0.28)	−58 (0.047; −0.44)	−0.19 (0.415; −0.15)	−0.09 (0.727; −0.07)	−0.37 (0.172; −0.29)	**−3.25 (0.042; −0.32)**	−0.31 (0.318; −0.29)
	T1×T2	0.09 (0.575; 0.09)	0.08 (0.723; 0.06)	0.22 (0.176; 0.10)	0.08 (0.720; 0.06)	0.09 (0.675; 0.06)	0.10 (0.921; 0.01)	0.25 (0.323; 0.24)
	T1×T3	−0.18 (0.296; −0.17)	**−0.48 (0.046; −0.38)**	−0.35 (0.389; −0.07)	−0.03 (0.866; −0.02)	−0.11 (0.607; −0.08)	−1.85 (0.143; −0.19)	0.14 (0.552; 0.13)
	T2×T3	**−0.27 (0.014; −0.26)**	**−0.56 (0.004; −0.42)**	−0.57 (0.180; −0.18)	−0.11 (0.277; −0.08)	−0.20 (0.190; −0.14)	−1.95 (0.073; −0.20)	−0.11 (0.278; −0.11)

Note: Bold: $p < 0.05$.

Abbreviations: AAQ-II, Acceptance and Action Questionnaire II; EDE-Q, Eating Disorder Examination Questionnaire; MD, mean difference; MPFI-EA, Multidimensional Psychological Flexibility Inventory-experiential avoidance scale.

TABLE 3 | Results of the repeated-measures mixed-model ANOVAs investigating changes in the H.O.M.E. intervention and comparisons with waiting list (DERS-16).

		DERS-16-total	DERS-16-*lack of emotional clarity*	DERS-16-*difficulty in engaging in goal directed behaviours*	DERS-16-*impulse control difficulties*	DERS-16-*limited access to emotion regulation strategies*	DERS-16-*nonacceptance of emotional responses*
	Group ($F_{(df)}$; p; η^2_p)	$2.41_{(38.1)}$; 0.129; 0.06	$0.23_{(38.1)}$; 0.632; 0.01	$2.42_{(38.1)}$; 0.128; 0.06	$1.62_{(38.1)}$; 0.211; 0.04	$2.70_{(38.1)}$; 0.108; 0.07	$1.90_{(38.1)}$; 0.176; 0.05
	Time ($F_{(df)}$; p; η^2_p)	**$2.92_{(38.1)}$; 0.037; 0.07**	$0.80_{(38.1)}$; 0.495; 0.02	$1.88_{(38.1)}$; 0.137; 0.05	$2.40_{(38.1)}$; 0.072; 0.06	**$2.81_{(38.1)}$; 0.043; 0.07**	**$2.81_{(38.1)}$; 0.043; 0.07**
	Time*group ($F_{(df)}$; p; η^2_p)	**$9.63_{(38.1)}$; <0.001; 0.20**	$2.40_{(38.1)}$; 0.071; 0.06	**$13.52_{(38.1)}$; <0.001; 0.26**	$2.53_{(38.1)}$; 0.061; 0.06	**7.64; <0.001; 0.17**	**$7.50_{(38.1)}$; <0.001; 0.17**
Post hoc							
VR MD (p; d)	T0×T1	**6.25 (0.006; 0.39)**	0.85 (0.080; 0.38)	**1.25 (0.008; 0.36)**	1.05 (0.072; 0.26)	**2.40 (0.001; 0.45)**	0.70 (0.204; 0.21)
	T0×T2	**8.70 (<0.001; 0.53)**	0.75 (0.844; 0.30)	**2.05 (<0.001; 0.61)**	**1.35 (0.023; 0.36)**	**2.60 (0.005; 0.44)**	**1.95 (<0.001; 0.60)**
	T0×T3	**10.00 (<0.001; 0.62)**	0.55 (0.249; 0.23)	**2.50 (<0.001; 0.76)**	**1.60 (0.004; 0.43)**	**3.25 (<0.001; 0.56)**	**2.10 (<0.001; 0.65)**
	T1×T2	2.45 (0.220; 0.15)	−0.10 (0.805; −0.04)	0.80 (0.056; 0.23)	0.30 (0.537; 0.08)	0.20 (0.759; 0.04)	**1.25 (0.014; 0.36)**
	T1×T3	3.75 (0.061; 0.23)	−0.30 (0.426; −0.13)	**1.25 (0.008; 0.36)**	0.55 (0.319; 0.14)	0.85 (0.229; 0.15)	**1.40 (0.006; 0.41)**
	T2×T3	1.30 (0.283; 0.08)	−0.20 (0.400; −0.08)	0.45 (0.148; 0.13)	0.25 (0.328; 0.07)	0.65 (0.137; 0.11)	0.15 (0.591; 0.05)
Waiting listMD (p; d)	T0×T1	−0.80 (0.712; −0.05)	−0.45 (0.347; −0.18)	−0.50 (0.269; −0.14)	0.20 (0.726; 0.06)	0.10 (0.885; 0.02)	−0.15 (0.783; −0.04)
	T0×T2	−0.30 (0.901; −0.02)	−0.25 (0.604; −0.10)	−0.50 (0.181; −0.14)	0.50 (0.385; 0.14)	0.00 (1.00; <0.001)	−0.05 (0.923; −0.01)
	T0×T3	**−5.35 (0.028; −0.34)**	−0.85 (0.078; −0.34)	**−0.1.60 (0.008; −0.49)**	−0.35 (0.511; −0.10)	−1.60 (0.075; −0.28)	−0.95 (0.102; −0.28)
	T1×T2	0.50 (0.801; 0.03)	0.20 (0.622; 0.09)	0.00 (1.00; <0.001)	0.30 (0.537; 0.10)	−0.10 (0.878; −0.02)	0.10 (0.837; 0.03)
	T1×T3	**−4.55 (0.025; −0.33)**	−0.40 (0.583; −0.18)	**−1.10 (0.018; −0.35)**	−0.55 (0.319; −0.18)	**−1.70 (0.019; −0.31)**	−0.80 (0.103; −0.27)
	T2×T3	**−5.05 (<0.001; −0.36)**	**−0.60 (0.015; −0.27)**	**−3.10 (0.004; −0.36)**	**−0.85 (0.002; −0.28)**	**−1.60 (<0.001; −0.30)**	**−0.90 (0.002; −0.31)**

Note: Bold: $p < 0.05$.

Abbreviations: DERS-16, Difficulties in Emotion Regulation Scale, short version; MD, mean difference.

promising results in improving all dysfunctional eating behaviours in the sample with significant differences and/or medium effect size differences between pre- and postintervention in the VR group. Levels of concerns regarding body shape were the variables that showed the greatest improvement (in terms of the greatest mean difference and effect size difference reported between pre- and postintervention scores). This scale was also the ED-related symptom with the highest levels in the sample at baseline. The improvement reported after the H.O.M.E. intervention is in line with results from other studies, which found that body shape concerns represent one of the ED-related symptoms which can more easily be improved by VR-based interventions, especially when VR is offered in addition to CBT-based protocols (Freeman et al. 2017). Similar to the presented preliminary results, reductions in the levels of concerns about eating and weight have also been found in the literature about VR (Irvine et al. 2020). For example, in a study by Irvine et al. (2020), a similar sample of young women with high levels of body image concerns (but without any ED diagnosis) reported improvements in concerns about weight, shape and eating after undergoing a VR-based training intervention programme through the use of an avatar. Even though the H.O.M.E. software does not offer an avatar feature, our VR-based intervention was still capable of producing beneficial results, which highlights that H.O.M.E. could have the potential to represent a helpful tool for the prevention of EDs and the reduction of dysfunctional eating behaviours in the GP at risk. Future research is, however, needed in order to confirm these results using more robust methodological approaches and bigger sample sizes.

Concerning the preliminary results about the transdiagnostic factors, as expected, difficulties in ER and psychological inflexibility also showed a significant improvement although with lower effect size differences at postintervention. Instead, no significant change and/or medium effect size differences were found for experiential avoidance.

Starting from the difficulties in ER, this transdiagnostic factor significantly improved both on a general level and in terms of improvements in participants' specific abilities to engage in goal-directed behaviours and to access their ER strategies. Although other difficulties in ER subscales did not show any improvements, possibly due to the small number of sessions of the intervention or the specific characteristics and targets of the intervention protocol, which mainly focused on helping people select strategies to face distressing emotions, this result brings up some potential beneficial effects of the H.O.M.E. VR-based preventive intervention. The ability of VR interventions to improve ER in a wide variety of clinical and nonclinical populations is also supported by studies using other VR software and protocols (Hadley et al. 2019). This may depend on the emotional involvement that VR is able to elicit (Colombo et al. 2021). In fact, during the H.O.M.E. VR-based intervention, participants were exposed to emotional stimuli, such as food, and this allowed an introspective observation of the emotions experienced and their association with dysfunctional eating behaviours. While these exploratory results are encouraging and pave the way for future research, the lack of large effect size differences underlines the need for future more robust studies to truly test the H.O.M.E. VR-based intervention protocol efficacy and to also improve the current intervention protocol.

Unfortunately, to the best of our knowledge, this preliminary study was also the first in the literature to test the application of a VR intervention for the improvement of psychological inflexibility and experiential avoidance. Even though VR was applied in EDs to improve *body image flexibility*, producing significant changes (Marco, Perpiñá, and Botella 2013), these results showed that VR also holds potential to produce significant improvements in psychological inflexibility but not in experiential avoidance, which could be due both to limitations present in the tool used to assess changes (the MPFI-EA, which is composed of only five items) and in characteristics of the H.O.M.E. intervention protocol (e.g., short duration). Indeed, these preliminary results highlight that the H.O.M.E. VR-based intervention protocol could benefit from some revision in order to better target psychological inflexibility and experiential avoidance more consistently. The lack of VR-based interventions capable of targeting these two transdiagnostic factors in the prevention of EDs or in other clinical and nonclinical populations (Gardini et al. 2022) does not allow authors to draw conclusions on specific aspects that could be improved in H.O.M.E. VR-based intervention protocol and highlights the necessity of incorporating qualitative evaluations from both participants and clinicians using the H.O.M.E. software and intervention. These qualitative evaluations will help to identify areas for improvement, refine the intervention's design and further explore its effects on transdiagnostic as well as ED-related factors. A higher number of sessions could also help produce greater changes in this aspect, as traditional psychotherapeutic interventions in similar populations usually entail a longer duration (Manlick, Cochran, and Koon 2013).

In line with recommendations for developing and testing VR intervention protocols and clinical trials (Birckhead et al. 2019), this preliminary study is also aimed at comparing the effects achieved by the H.O.M.E. VR-based intervention to those achieved by a waiting list condition, to establish whether the obtained changes could be attributed to the intervention and thus determine whether it could be tested when used against to or in combination with active traditional psychological interventions (e.g., CBT-E) as well. Results managed to establish the H.O.M.E. VR-based intervention promising effects compared to the waiting list. Indeed, all the aforementioned changes on targeted outcomes achieved by the VR group were greater than those achieved by the waiting list group, who also did not report any significant difference between pre- and postintervention levels of dysfunctional eating behaviours or transdiagnostic factors. In the literature, VR-based interventions have often proven their efficacy against nonactive control groups for the treatment of several psychopathologies (including EDs) (Ferrer-García and Gutiérrez-Maldonado 2012; Freeman et al. 2017; Geraets, Wallinius, and Sygel 2022; Riva, Malighetti, and Serino 2021). Although to date no preventive VR-based interventions are available for EDs specifically, the results achieved by the H.O.M.E. intervention can pave the way towards future studies to further confirm the potential benefits of this protocol and VR as a technology for the prevention of these psychopathologies, especially when also targeting transdiagnostic factors and when compared to or used in addition to traditional psychological interventions.

Moreover, the potential preventive utility of H.O.M.E. was also proven by the maintenance of the achieved improvements at

follow-ups of 3 and 6 months. Indeed, while participants in the waiting list condition seemed to experience the same levels of symptoms after 3 months and even a worsening of dysfunctional eating behaviours and ER difficulties after 6 months, improvements achieved by the VR group remained unchanged in the months following the intervention. Although longer follow-ups are needed to further prove the maintenance of these results, they confirm the potential of H.O.M.E. and VR in general to make long-lasting positive changes in the GP at risk for EDs.

Another important and promising element to underline is that no participant dropped out during the H.O.M.E. intervention and the entirety of the VR group underwent all the six sessions. This is of clinical relevance as drop-out rates in treatments targeting people with EDs are high (Dalle Grave 2014) and motivation of participants undergoing ED prevention protocols can be low (Harrer et al. 2020; Shaw, Stice, and Becker 2009): the interactive component of VR may have indeed contributed to engaging participants in completing the intervention, as it has oftentimes been observed in the literature (Riva et al. 2016).

4.1 | Limitations

The promising results achieved by the presented study should be observed in light of the methodological limitations, such as the small sample size, which might be due to participants' personal availability to take part in the study, and also to the egosyntonic nature of ED symptoms which may reduce people's motivation to change (Roncero et al. 2013). The small sample size, as well as its homogeneous characteristics (that is, GP at risk for EDs and predominantly young female university students), also limits the generalisability of these potential results to larger populations, other age and gender groups, and clinical samples with a full ED or other DSM-5-TR diagnoses. Future studies already plan to test the H.O.M.E. intervention protocol in more heterogeneous and clinical ED population to further test its potential in the prevention and treatment of EDs. The use of a waiting list condition as a nonactive control group for comparison also represents a limitation in assessing the true efficacy and effectiveness of the H.O.M.E. VR-based intervention protocol. However, these promising results also pave the way to future research testing H.O.M.E. when used against or in addition to evidence-based treatments (e.g., CBT-E).

Moreover, the self-report measures used for the evaluation of the variables include several limitations, such as the tendency participants may have to answer questions based on social desirability and the difficulties people may encounter in answering questions with a reliable level of insight, especially when they concern complex psychological factors such as psychological inflexibility and experiential avoidance (Chan 2008). Moreover, the MPFI-EA (Landi et al. 2021; Rolffs, Rogge, and Wilson 2018), used to evaluate experiential avoidance in the study, consisted of only five items, which may have led to methodological biases. Studies evaluating participants' opinions about the H.O.M.E. software and intervention using semistructured interviews are being conducted in order to better assess the overall quality of H.O.M.E. as well as its impact on a more subjective level.

Finally, the fact that the H.O.M.E. software needs to be used at the presence of a clinical psychologist guiding the patient through the environment allowed the intervention to be tailored to the participants' needs, but it also limits the applicability of the intervention outside of the psychologist's office. Therefore, authors also plan to develop a standalone version of the H.O.M.E. protocol using different technologies, such as augmented reality or mobile apps, which will allow the H.O.M.E. VR-based intervention to be offered to an even wider range of people (Cipresso et al. 2018; Riva et al. 2016; Vinci et al. 2020), thus making the protocol more accessible and useful in preventing the onset of EDs and other mental health conditions in more individuals at risk.

4.2 | Conclusion

While there are numerous studies in the literature that attest to the usefulness of VR systems in the treatment of EDs (Ferrer-García and Gutiérrez-Maldonado 2012; Riva, Malighetti, and Serino 2021), those investigating the use of VR-based interventions in the prevention of EDs in the GP at risk, especially through the improvement of transdiagnostic factors, are lacking (Ciao, Loth, and Neumark-Sztainer 2014).

These preliminary results highlighted the potential role that the H.O.M.E. VR-based transdiagnostic intervention protocol could play in the prevention of EDs, as it reduced dysfunctional eating behaviours, such as preoccupation with eating, shape and weight, and improved some of the transdiagnostic factors that can lead to EDs, in particular difficulties in ER and psychological inflexibility (Levin et al. 2014; Morton et al. 2020; Prefit, Cândea, and Szentagotai-Tătar 2019).

The present exploratory study also showed good feasibility of the H.O.M.E. protocol and its applicability in bigger and more robust research protocols as well as in the clinical practice. This could prove to be beneficial since, despite the use of VR is gradually spreading in clinical psychology and, in particular, in the assessment, treatment and prevention of EDs (Brown et al. 2020; Rizzo, Thomas Koenig, and Talbot 2019), many issues still hinder the applicability of this technology, such as the high costs (in terms of software, hardware and training) and the predominantly disorder-specific nature of VR software and intervention protocols available to date (Brown et al. 2020; Emmelkamp and Meyerbröker 2021; Freeman et al. 2017; Gardini et al. 2022). Working on developing and testing transdiagnostic VR software such as H.O.M.E. would improve VR applicability in the clinical practice and in the prevention of psychiatric disorders.

Acknowledgements

The authors have nothing to report. Open access publishing facilitated by Universita degli Studi di Bologna, as part of the Wiley - CRUI-CARE agreement.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

References

Aldao, A., S. Nolen-Hoeksema, and S. Schweizer. 2010. "Emotion-Regulation Strategies Across Psychopathology: A Meta-Analytic Review." *Clinical Psychology Review* 30, no. 2: 217–237. https://doi.org/10.1016/j.cpr.2009.11.004.

American Psychiatric Association. 2022. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed., text rev. ed. American Psychiatric Association.

Atkinson, M. J., and T. D. Wade. 2013. "Enhancing Dissemination in Selective Eating Disorders Prevention: An Investigation of Voluntary Participation Among Female University Students." *Behaviour Research and Therapy* 51, no. 12: 806–816. https://doi.org/10.1016/j.brat.2013.09.012.

Atwood, M. E., and A. Friedman. 2020. "A Systematic Review of Enhanced Cognitive Behavioral Therapy (CBT-E) for Eating Disorders." *International Journal of Eating Disorders* 53, no. 3: 311–330. https://doi.org/10.1002/eat.23206.

Birckhead, B., C. Khalil, X. Liu, et al. 2019. "Recommendations for Methodology of Virtual Reality Clinical Trials in Health Care by an International Working Group: Iterative Study." *JMIR Mental Health* 6, no. 1: e11973. https://doi.org/10.2196/11973.

Bjureberg, J., B. Ljótsson, M. T. Tull, et al. 2016. "Development and Validation of a Brief Version of the Difficulties in Emotion Regulation Scale: The DERS-16." *Journal of Psychopathology and Behavioral Assessment* 38, no. 2: 284–296. https://doi.org/10.1007/s10862-015-9514-x.

Bluett, E. J., E. B. Lee, M. Simone, et al. 2016. "The Role of Body Image Psychological Flexibility on the Treatment of Eating Disorders in a Residential Facility." *Eating Behaviors* 23: 150–155. https://doi.org/10.1016/j.eatbeh.2016.10.002.

Bond, F. W., S. C. Hayes, R. A. Baer, et al. 2011. "Preliminary Psychometric Properties of the Acceptance and Action Questionnaire-II: A Revised Measure of Psychological Inflexibility and Experiential Avoidance." *Behavior Therapy* 42, no. 4: 676–688. https://doi.org/10.1016/j.beth.2011.03.007.

Brown, T., E. Nauman Vogel, S. Adler, et al. 2020. "Bringing Virtual Reality From Clinical Trials to Clinical Practice for the Treatment of Eating Disorders: An Example Using Virtual Reality Cue Exposure Therapy." *Journal of Medical Internet Research* 22, no. 4: e16386. https://doi.org/10.2196/16386.

Calugi, S., C. Milanese, M. Sartirana, et al. 2017. "The Eating Disorder Examination Questionnaire: Reliability and Validity of the Italian Version." *Eating and Weight Disorders: EWD* 22, no. 3: 509–514. https://doi.org/10.1007/s40519-016-0276-6.

Chan, D. 2008. "So Why Ask Me? Are Self-Report Data Really That Bad?" In *Statistical and Methodological Myths and Urban Legends*. Routledge.

Chawla, N., and B. Ostafin. 2007. "Experiential Avoidance as a Functional Dimensional Approach to Psychopathology: An Empirical Review." *Journal of Clinical Psychology* 63, no. 9: 871–890. https://doi.org/10.1002/jclp.20400.

Ciao, A. C., K. Loth, and D. Neumark-Sztainer. 2014. "Preventing Eating Disorder Pathology: Common and Unique Features of Successful Eating Disorders Prevention Programs." *Current Psychiatry Reports* 16, no. 7: 453. https://doi.org/10.1007/s11920-014-0453-0.

Cipresso, P., I. A. C. Giglioli, M. A. Raya, and G. Riva. 2018. "The Past, Present, and Future of Virtual and Augmented Reality Research: A Network and Cluster Analysis of the Literature." *Frontiers in Psychology* 9. https://doi.org/10.3389/fpsyg.2018.02086.

Colombo, D., A. Díaz-García, J. Fernandez-Álvarez, and C. Botella. 2021. "Virtual Reality for the Enhancement of Emotion Regulation." *Clinical Psychology & Psychotherapy* 28, no. 3: 519–537. https://doi.org/10.1002/cpp.2618.

Culbert, K. M., S. E. Racine, and K. L. Klump. 2015. "Research Review: What We Have Learned About the Causes of Eating Disorders - A Synthesis of Sociocultural, Psychological, and Biological Research." *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 56, no. 11: 1141–1164. https://doi.org/10.1111/jcpp.12441.

Dalle Grave, R. 2014. *Disturbi Dell'alimentazione: Una Guida Pratica per I Familiari*. Positive Press.

Della Longa, N. M., and K. P. De Young. 2018. "Experiential Avoidance, Eating Expectancies, and Binge Eating: A Preliminary Test of an Adaption of the Acquired Preparedness Model of Eating Disorder Risk." *Appetite* 120: 423–430. https://doi.org/10.1016/j.appet.2017.09.022.

Ekeroth, K., and A. Birgegård. 2014. "Evaluating Reliable and Clinically Significant Change in Eating Disorders: Comparisons to Changes in DSM-IV Diagnoses." *Psychiatry Research* 216, no. 2: 248–254. https://doi.org/10.1016/j.psychres.2014.02.008.

Emmelkamp, P. M. G., and K. Meyerbröker. 2021. "Virtual Reality Therapy in Mental Health." *Annual Review of Clinical Psychology* 17: 495–519. https://doi.org/10.1146/annurev-clinpsy-081219-115923.

Fairburn, C. G., and S. J. Beglin. 1994. "Assessment of Eating Disorders: Interview or Self-Report Questionnaire?" *International Journal of Eating Disorders* 16, no. 4: 363–370.

Fairburn, C. G., and S. J. Beglin. 2008. "Eating Disorder Examination Questionnaire." *Cognitive Behavior Therapy and Eating Disorders* 309: 313.

Fairburn, C. G., Z. Cooper, and R. Shafran. 2003. "Cognitive Behaviour Therapy for Eating Disorders: A «Transdiagnostic» Theory and Treatment." *Behaviour Research and Therapy* 41, no. 5: 509–528. https://doi.org/10.1016/s0005-7967(02)00088-8.

Fairholme, C. P., C. L. Boisseau, K. K. Ellard, J. T. Ehrenreich, and D. H. Barlow. 2010. "Emotions, Emotion Regulation, and Psychological Treatment: A Unified Perspective." In *Emotion Regulation and Psychopathology: A Transdiagnostic Approach to Etiology and Treatment*, 283–309. Guilford Press.

Ferrer-García, M., and J. Gutiérrez-Maldonado. 2012. "The Use of Virtual Reality in the Study, Assessment, and Treatment of Body Image in Eating Disorders and Nonclinical Samples: A Review of the Literature." *Body Image* 9, no. 1: 1–11. https://doi.org/10.1016/j.bodyim.2011.10.001.

Freeman, D., S. Reeve, A. Robinson, et al. 2017. "Virtual Reality in the Assessment, Understanding, and Treatment of Mental Health Disorders." *Psychological Medicine* 47, no. 14: 2393–2400. https://doi.org/10.1017/S003329171700040X.

Gardini, V., C. Ruini, E. Tossani, S. Grandi, and E. Tomba. 2023. "Protocol for a Randomized Controlled Trial Testing the Efficacy of a Transdiagnostic Virtual Reality-Based Intervention for the Reduction of Unhealthy Lifestyles and Behaviors in the General Population." *Journal of Clinical Medicine* 12, no. 23: 7470. https://doi.org/10.3390/jcm12237470.

Gardini, V., G. Gamberini, S. Müller, S. Grandi, and E. Tomba. 2022. "The Effects of Virtual Reality in Targeting Transdiagnostic Factors for Mental Health: A Systematic Review of the Literature." *Journal of Clinical Medicine* 11, no. 21: 6463. https://doi.org/10.3390/jcm11216463.

Geraets, C. N. W., M. Wallinius, and K. Sygel. 2022. "Use of Virtual Reality in Psychiatric Diagnostic Assessments: A Systematic Review." *Frontiers in Psychiatry* 13: 828410. https://doi.org/10.3389/fpsyt.2022.828410.

Gorini, A., E. Griez, A. Petrova, and G. Riva. 2010. "Assessment of the Emotional Responses Produced by Exposure to Real Food, Virtual Food and Photographs of Food in Patients Affected by Eating Disorders." *Annals of General Psychiatry* 9: 30. https://doi.org/10.1186/1744-859X-9-30.

Hadley, W., C. Houck, L. K. Brown, J. S. Spitalnick, M. Ferrer, and D. Barker. 2019. "Moving Beyond Role-Play: Evaluating the Use of Virtual Reality to Teach Emotion Regulation for the Prevention of Adolescent Risk Behavior Within a Randomized Pilot Trial." *Journal of Pediatric Psychology* 44, no. 4: 425–435. https://doi.org/10.1093/jpepsy/jsy092.

Hallion, L. S., S. A. Steinman, D. F. Tolin, and G. J. Diefenbach. 2018. "Psychometric Properties of the Difficulties in Emotion Regulation Scale (DERS) and Its Short Forms in Adults With Emotional Disorders." *Frontiers in Psychology* 9: 539. https://doi.org/10.3389/fpsyg.2018.00539.

Han, S., and M. C. Pistole. 2014. "College Student Binge Eating: Insecure Attachment and Emotion Regulation." *Journal of College Student Development* 55, no. 1: 16–29. https://doi.org/10.1353/csd.2014.0004.

Harrer, M., S. H. Adam, E.-M. Messner, et al. 2020. "Prevention of Eating Disorders at Universities: A Systematic Review and Meta-Analysis." *International Journal of Eating Disorders* 53, no. 6: 813–833. https://doi.org/10.1002/eat.23224.

Harvey, A., E. Watkins, W. Mansell, and R. Shafran. 2004. *Cognitive Behavioural Processes Across Psychological Disorders. A Transdiagnostic Approach to Research and Treatment*. Oxford University Press.

Hayes, S. C., and S. G. Hofmann. 2017. "The Third Wave of Cognitive Behavioral Therapy and the Rise of Process-Based Care." *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)* 16, no. 3: 245–246. https://doi.org/10.1002/wps.20442.

Hayes, S. C., J. B. Luoma, F. W. Bond, A. Masuda, and J. Lillis. 2006. "Acceptance and Commitment Therapy: Model, Processes and Outcomes." *Behaviour Research and Therapy* 44, no. 1: 1–25. https://doi.org/10.1016/j.brat.2005.06.006.

Hayes, S. C., K. G. Wilson, E. V. Gifford, V. M. Follette, and K. Strosahl. 1996. "Experiential Avoidance and Behavioral Disorders: A Functional Dimensional Approach to Diagnosis and Treatment." *Journal of Consulting and Clinical Psychology* 64, no. 6: 1152–1168. https://doi.org/10.1037/0022-006X.64.6.1152.

Holmqvist Larsson, K., A. Lowén, L. Hellerstedt, L. Bergcrona, M. Salerud, and M. Zetterqvist. 2020. "Emotion Regulation Group Skills Training: A Pilot Study of an Add-On Treatment for Eating Disorders in a Clinical Setting." *Journal of Eating Disorders* 8, no. 1: 12. https://doi.org/10.1186/s40337-020-00289-1.

Irvine, K. R., A. R. Irvine, N. Maalin, et al. 2020. "Using Immersive Virtual Reality to Modify Body Image." *Body Image* 33: 232–243. https://doi.org/10.1016/j.bodyim.2020.03.007.

Jacobi, C., C. Hayward, M. de Zwaan, H. C. Kraemer, and W. S. Agras. 2004. "Coming to Terms With Risk Factors for Eating Disorders: Application of Risk Terminology and Suggestions for a General Taxonomy." *Psychological Bulletin* 130, no. 1: 19–65. https://doi.org/10.1037/0033-2909.130.1.19.

Kim, J., and W. Shin. 2014. "How to Do Random Allocation (Randomization)." *Clinics in Orthopedic Surgery* 6, no. 1: 103. https://doi.org/10.4055/cios.2014.6.1.103.

Landi, G., K. I. Pakenham, A. M. Giovannetti, et al. 2021. "Italian Validation of the Italian Multidimensional Psychological Flexibility Inventory (MPFI)." *Journal of Contextual Behavioral Science* 21: 57–65. https://doi.org/10.1016/j.jcbs.2021.05.007.

Le, L. K.-D., J. J. Barendregt, P. Hay, and C. Mihalopoulos. 2017. "Prevention of Eating Disorders: A Systematic Review and Meta-Analysis." *Clinical Psychology Review* 53: 46–58. https://doi.org/10.1016/j.cpr.2017.02.001.

Levin, M. E., C. MacLane, S. Daflos, et al. 2014. "Examining Psychological Inflexibility as a Transdiagnostic Process Across Psychological Disorders." *Journal of Contextual Behavioral Science* 3, no. 3: 155–163. https://doi.org/10.1016/j.jcbs.2014.06.003.

Manlick, C. F., S. V. Cochran, and J. Koon. 2013. "Acceptance and Commitment Therapy for Eating Disorders: Rationale and Literature Review." *Journal of Contemporary Psychotherapy* 43, no. 2: 115–122. https://doi.org/10.1007/s10879-012-9223-7.

Mansell, W., A. Harvey, E. Watkins, and R. Shafran. 2008. "Cognitive Behavioral Processes Across Psychological Disorders: A Review of the Utility and Validity of the Transdiagnostic Approach." *International Journal of Cognitive Therapy* 1: 181–191. https://doi.org/10.1521/ijct.2008.1.3.181.

Maples-Keller, J. L., B. E. Bunnell, S.-J. Kim, and B. O. Rothbaum. 2017. "The Use of Virtual Reality Technology in the Treatment of Anxiety and Other Psychiatric Disorders." *Harvard Review of Psychiatry* 25, no. 3: 103–113. https://doi.org/10.1097/HRP.0000000000000138.

Marco, J. H., C. Perpiñá, and C. Botella. 2013. "Effectiveness of Cognitive Behavioral Therapy Supported by Virtual Reality in the Treatment of Body Image in Eating Disorders: One Year Follow-Up." *Psychiatry Research* 209, no. 3: 619–625. https://doi.org/10.1016/j.psychres.2013.02.023.

Mastrobattista, L., and R. Pacifici. 2022. *Guida ai Servizi Territoriali per la Cura dei Disturbi della Nutrizione e Dell'alimentazione*. Roma: Istituto Superiore di Sanità.

Masuda, A., M. Price, P. L. Anderson, and J. W. Wendell. 2010. "Disordered Eating-Related Cognition and Psychological Flexibility as Predictors of Psychological Health Among College Students." *Behavior Modification* 34, no. 1: 3–15. https://doi.org/10.1177/0145445509351569.

McClure, Z., M. Messer, C. Anderson, C. Liu, and J. Linardon. 2022. "Which Dimensions of Emotion Dysregulation Predict the Onset and Persistence of Eating Disorder Behaviours? A Prospective Study." *Journal of Affective Disorders* 310: 123–128. https://doi.org/10.1016/j.jad.2022.05.041.

Morgan, J. F., F. Reid, and J. H. Lacey. 1999. "The SCOFF Questionnaire: Assessment of a New Screening Tool for Eating Disorders." *BMJ (Clinical Research Ed.)* 319, no. 7223: 1467–1468. https://doi.org/10.1136/bmj.319.7223.1467.

Morton, C., T. A. Mooney, L. L. Lozano, E. A. Adams, H. M. Makriyianis, and M. Liss. 2020. "Psychological Inflexibility Moderates the Relationship Between Thin-Ideal Internalization and Disordered Eating." *Eating Behaviors* 36: 101345. https://doi.org/10.1016/j.eatbeh.2019.101345.

Pannocchia, L., M. Di Fiorino, M. Giannini, and J. Vanderlinden. 2011. "A Psychometric Exploration of an Italian Translation of the SCOFF Questionnaire." *European Eating Disorders Review: The Journal of the Eating Disorders Association* 19: 371–373. https://doi.org/10.1002/erv.1105.

Pennato, T., C. Berrocal, O. Bernini, and T. Rivas. 2013. "Italian Version of the Acceptance and Action Questionnaire-II (AAQ-II): Dimensionality, Reliability, Convergent and Criterion Validity." *Journal of Psychopathology and Behavioral Assessment* 35, no. 4: 552–563. https://doi.org/10.1007/s10862-013-9355-4.

Peterson, C. B., R. D. Crosby, S. A. Wonderlich, et al. 2007. "Psychometric Properties of the Eating Disorder Examination-Questionnaire: Factor Structure and Internal Consistency." *International Journal of Eating Disorders* 40, no. 4: 386–389. https://doi.org/10.1002/eat.20373.

Prefit, A.-B., D. M. Cândea, and A. Szentagotai-Tătar. 2019. "Emotion Regulation Across Eating Pathology: A Meta-Analysis." *Appetite* 143: 104438. https://doi.org/10.1016/j.appet.2019.104438.

Rawal, A., R. J. Park, and J. M. G. Williams. 2010. "Rumination, Experiential Avoidance, and Dysfunctional Thinking in Eating Disorders." *Behaviour Research and Therapy* 48, no. 9: 851–859. https://doi.org/10.1016/j.brat.2010.05.009.

Riva, G. 2022. "Virtual Reality in Clinical Psychology." *Comprehensive Clinical Psychology*: 91–105. https://doi.org/10.1016/B978-0-12-818697-8.00006-6.

Riva, G., C. Malighetti, and S. Serino. 2021. "Virtual Reality in the Treatment of Eating Disorders." *Clinical Psychology & Psychotherapy* 28, no. 3: 477–488. https://doi.org/10.1002/cpp.2622.

Riva, G., E. Molinari, and F. Vincelli. 2002. "Interaction and Presence in the Clinical Relationship: Virtual Reality (VR) as Communicative Medium Between Patient and Therapist." *IEEE Transactions on Information Technology in Biomedicine: A Publication of the IEEE Engineering in Medicine and Biology Society* 6, no. 3: 198–205. https://doi.org/10.1109/titb.2002.802370.

Riva, G., R. M. Baños, C. Botella, F. Mantovani, and A. Gaggioli. 2016. "Transforming Experience: The Potential of Augmented Reality and Virtual Reality for Enhancing Personal and Clinical Change." *Frontiers in Psychiatry* 7. https://doi.org/10.3389/fpsyt.2016.00164.

Rizzo, A., S. Thomas Koenig, and T. B. Talbot. 2019. "Clinical Results Using Virtual Reality." *Journal of Technology in Human Services* 37, no. 1: 51–74. https://doi.org/10.1080/15228835.2019.1604292.

Rolffs, J. L., R. D. Rogge, and K. G. Wilson. 2018. "Disentangling Components of Flexibility via the Hexaflex Model: Development and Validation of the Multidimensional Psychological Flexibility Inventory (MPFI)." *Assessment* 25, no. 4: 458–482. https://doi.org/10.1177/1073191116645905.

Roncero, M., A. Belloch, C. Perpiñá, and J. Treasure. 2013. "Ego-Syntonicity and Ego-Dystonicity of Eating-Related Intrusive Thoughts in Patients With Eating Disorders." *Psychiatry Research* 208, no. 1: 67–73. https://doi.org/10.1016/j.psychres.2013.01.006.

Rothman, K. J. 1990. "No Adjustments Are Needed for Multiple Comparisons." *Epidemiology* 1, no. 1: 43–46.

Schwartz, C., K. Drexl, A. Fischer, et al. 2019. "Universal Prevention in Eating Disorders: A Systematic Narrative Review of Recent Studies." *Mental Health & Prevention* 14: 200162. https://doi.org/10.1016/j.mph.2019.200162.

Shaw, H., E. Stice, and C. B. Becker. 2009. "Preventing Eating Disorders." *Child and Adolescent Psychiatric Clinics of North America* 18, no. 1: 199–207. https://doi.org/10.1016/j.chc.2008.07.012.

Skinner, K. D., S. M. Rojas, and J. C. Veilleux. 2017. "Connecting Eating Pathology With Risk for Engaging in Suicidal Behavior: The Mediating Role of Experiential Avoidance." *Suicide & Life-Threatening Behavior* 47, no. 1: 3–13. https://doi.org/10.1111/sltb.12249.

Stice, E., Z. A. Onipede, and C. N. Marti. 2021. "A Meta-Analytic Review of Trials That Tested Whether Eating Disorder Prevention Programs Prevent Eating Disorder Onset." *Clinical Psychology Review* 87: 102046. https://doi.org/10.1016/j.cpr.2021.102046.

Vinci, C., K. O. Brandon, M. Kleinjan, and T. H. Brandon. 2020. "The Clinical Potential of Augmented Reality." *Clinical Psychology: Science and Practice* 27, no. 3. https://doi.org/10.1111/cpsp.12357.

Watson, H. J., T. Joyce, E. French, et al. 2016. "Prevention of Eating Disorders: A Systematic Review of Randomized, Controlled Trials." *International Journal of Eating Disorders* 49, no. 9: 833–862. https://doi.org/10.1002/eat.22577.

Westwood, H., J. Kerr-Gaffney, D. Stahl, and K. Tchanturia. 2017. "Alexithymia in Eating Disorders: Systematic Review and Meta-Analyses of Studies Using the Toronto Alexithymia Scale." *Journal of Psychosomatic Research* 99: 66–81. https://doi.org/10.1016/j.jpsychores.2017.06.007.

Wiederhold, B. K., and M. D. Wiederhold. 2005. *Virtual Reality Therapy for Anxiety Disorders: Advances in Evaluation and Treatment (pp. viii, 225)*. American Psychological Association. https://doi.org/10.1037/10858-000.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Social Psychology of Education (2021) 24:609–634
https://doi.org/10.1007/s11218-021-09626-x

Girls suffer: the prevalence and predicting factors of emotional problems among adolescents during upper secondary school in Norway

Jan Arvid Haugan[1] · **Per Frostad**[1] · **Per-Egil Mjaavatn**[1]

Received: 7 September 2020 / Accepted: 26 February 2021 / Published online: 29 March 2021

Abstract
This longitudinal, quantitative survey examined factors predicting 1077 Norwegian adolescents` emotional problems during Upper Secondary School (grades I–III, approximately 16–19 years old) considering the following research question: "To what extent do students in Upper Secondary School experience emotional problems, and how are these problems predicted by gender, academic/social self-concept, coping beliefs, appearance pressure and school stress?". The mentioned variables were scrutinized through analysis of frequencies, zero order correlations and structural equation modeling. Results verified previous findings that there was an increase in emotional problems for adolescents, especially among girls. Furthermore, the study results indicate that coping beliefs is a crucial factor when it comes to the perception of pressure and stress, and the subsequent development of emotional health problems.

Keywords Emotional problems · Coping beliefs · Stress · Gender

1 Introduction

Mental health problems are a matter of major public health concern and are among the leading contributors to the health burden for children and adolescents globally (Erskine et al., 2015; Whiteford et al., 2013). The literature usually separates between mental *problems* and *disorders.* Mental health problems refer to symptoms

✉ Jan Arvid Haugan
jan.haugan@ntnu.no

Per Frostad
per.frostad@ntnu.no

Per-Egil Mjaavatn
per.egil.mjaavatn@ntnu.no

[1] Department of Education and Lifelong Learning, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway

that have significantly negative impact on well-being, learning, daily tasks and socializing with others, while they do not satisfy the criteria for a psychiatric diagnosis (Mykletun et al., 2009). A mental disorder can be assessed when the symptoms are in coherence with the criteria in a diagnostic system, such as ICD10 (World Health Organization, 2018) or DSM-5 (American Psychiatric Association, 2013). Mental health disorders are concluded by health personnel based on diagnostic interviews, while it is common to identify the extent of mental health problems through the use of questionnaires. In the following we refer to mental health problems when we discuss mental health among adolescents.

About 10–20% of children and adolescents in the world suffer from mental health problems (Henderson et al., 2017; Kieling et al., 2011). This can have serious consequences for the quality of life as mental health problems during adolescence are related to weak academic and social functioning (Balazs et al., 2013; Rasing et al., 2020). Studies of prevalence from the 80s into the 21st century suggest a global increase in both mental health problems and disorders, especially internalized symptoms among girls (Bor et al., 2014; Choi, 2018). This tendency is supported by studies from the USA (Merikangas et al., 2010), the UK (Collishaw, 2015) and other countries in the western world (Auerbach et al., 2018). These circumstances have led to strengthened emphasis on how teachers, counsellors and the school as a whole can reduce contextual stressors and facilitate adolescents` ability to cope with life challenges. Based on this, research on risk- and protective-factors connected to coping has increased in the last few decades.

Coping is defined as the "constantly changing cognitive and behavioral efforts to manage specific external and/or internal demands that are appraised as taxing or exceeding the resources of the person" (Lazarus and Folkman, 1984, p. 141). However, coping factors cannot be reduced to purely individual, psychological traits (Condly, 2006; Herrman et al., 2011). Other definitions emphasize the ecological influence on coping, stating that risk- and protective-factors include: "(…) both the capacity of individuals to navigate their way to health-sustaining resources (…) and a condition of the individual`s family, community and culture to provide these health resources and experiences in culturally meaningful ways" (Ungar, 2008, p. 225). This ecological perspective on coping is supported by research indicating that the quality of relationships with significant others is of considerable importance (Currie et al., 2012; Lerner, et al., 2005), in addition to broader environmental factors such as a safe and inclusive classroom environment (Skaalvik and Skaalvik, 2013; Theron, 2016).

In the present study, we will demarcate the scrutiny of risk- and protective-factors to the associations between self-concept, coping beliefs, appearance pressure, school stress, and emotional problems. These are all individual perceptions, and we recognize that coping encompasses a dynamic, bi-directional process in which the environment and the individual interacts constructively over time. In such, contextual factors in both proximal microsystems (e.g. family, school, friends) and in the more distal macrosystems (e.g. politics, culture) affect coping during adolescence (Lerner et al., 2005, Ungar, 2008). However, numerous findings also indicate that individuals` cognitive perceptions and interpretations are important when understanding the development of stress, perceptions of pressure, and mental health problems (e.g.

Fletcher and Sarkar, 2013; Jongen, McCalman and Bainbridge, 2019). From the perspective of Lazarus and Folkman's (1984) transactional stress model, adolescents` emotional problems depend not only on the objective challenges or stressors they are exposed to, but also on how they subjectively evaluate the situation and beliefs in their own ability to handle it. Thus, their transactional model of stress claims, in line with other research (e.g. Sowislo and Orth, 2013; Yeoh et al., 2017), that self-concept and coping beliefs are decisive constructs to predict experiences of stress and subsequent emotional problems. Individuals with high levels of coping beliefs are more likely to view difficult tasks as challenges rather than things to be avoided. Thus, they are more likely to experience less distress and, and subsequently, fewer emotional problems (e.g. Bandura et al., 1999; Orth et al., 2014).

In the following, we will review research on the associations between adolescents' perceptions of emotional problems, self-concept, coping beliefs, appearance pressure and school stress. This literature review will end with a formulation of a primary research question with six related hypotheses. In the next section we describe methodological aspects, including descriptions of the participants and our procedure, instruments and data analysis. Thereafter, we present findings and discuss the self-reported experience of Norwegian Upper Secondary School students in light of earlier research and theory. This is followed by a discussion of the limitations with the study, practical implications and future directions for research. The article ends with conclusive remarks.

2 Literature review

2.1 Self-concept

Numerous studies have found that mental health is linked to the experience of self-concept during adolescence (e.g. McCarty et al., 2007; Ohannessian et al., 1999). In the literature, self-concept is generally defined as a person`s perceptions and evaluations of personal qualities, abilities, and knowledge (Esnaola et al., 2020; Ohannessian et al., 2019), in academic, physical, and social domains (Hankin et al., 2015; Van Keyserlingket al., 2019). The literature expresses increasing evidence that supports a multidimensional perspective of self-concept, often based on Shavelson, Hubner and Stanton`s (1976) multidimensional, hierarchical, and domain-specific model of self-concept, divided into academic and non-academic components. Academic self-concept is further divided into self-concepts in particular subject areas (Mathematics, English, etc.) and non-academic self-concept is divided into social, emotional, and physical self-concepts. This is supported by Marsh and Craven (1997) who argue that "specific domains of self-concept are more useful than a general domain" (p. 191) in order to understand the complexity, predict a wide variety of behaviors, relate self-concept to other constructs, and provide outcome measures for diverse interventions in educational settings.

Previous research indicates that self-concept is an important mediator factor for psychological, social, behavioral and educational outcomes, such as personal development, socialization processes, motivation, coping beliefs, perceptions of stress,

and mental health (Craven and Marsh, 2008; Marsh and Martin, 2011). Taking into account the multidimensionality nature of the self-concept, it is necessary to explore the predictive effects on mental health across several dimensions as the developmental patterns may differ. In this study we will emphasize two of these dimensions: academic self-concept and social self-concept.

Academic self-concept is related to learners` knowledge and perceptions about themselves in overall academic domains (Marsh and Martin, 2011; Wigfield and Karpathian, 1991). Marsh (1989) found that the academic self-concept reaches its lowest point in grades 8 or 9 and subsequently increases through late adolescence, while other studies have found a consistent decline during secondary education (De Fraine et al., 2007; Nagy et al., 2010). Some studies conclude that males score higher than females (De Fraine et al., 2007; Pinxten, et al., 2013), while others find an advantage for females (Young and Mroczek, 2003), or report no significant differences (Esnaola et al., 2018; Marsh et al., 2005). Few studies have investigated the development of social self-concept. However, some studies have found that lower social self-concept is associated with increased emotional problems during mid-late adolescence (Kim-Spoon et al., 2012; Lee et al., 2010).

Based on these findings, it is relevant to underline that Shavelson et al., (1976) emphasized the importance of self-attributions, and suggested that forms of behavior and dimensions of self-concept have correlated and reciprocal relations. This is supported by Marsh and Craven (2006) who reviewed research on academic outcomes and academic self-concept. They found the correlation between academic self-concept and academic achievement (e.g. school grades) to have the strongest association. However, they could not find correlations between other non-academic components of self-concept and academic outcomes, underlining the multidimensionality of the construct (Marsh and Martin, 2011). In sum, the above-mentioned findings indicate that both academic self-concept and social self-concept are directly and positively related with coping beliefs, and negatively related to emotional problems (Craven and Marsh, 2008; Marsh and Martin, 2011).

2.2 Coping beliefs

In a systematic review and structural equation modeling meta-analysis of 19 studies, Groth et al. (2019) found that coping is predicted by coping beliefs and subsequently predicts mental health outcomes. Four longitudinal studies indicated that coping beliefs were significantly associated with eating problems (Halvarsson-Edlund et al., 2008), anxiety (Lopez and Little, 1996; Weigold and Robitschek, 2011), and depression (Pérez et al., 2009). Furthermore, in light of their findings, Groth et al. (2019) suggest that individuals who report low coping beliefs tend to underestimate the potential positive consequences of their behavior and the likelihood to be rewarded for this behavior. These findings can be interpreted in light of the theory of learned helplessness (Abramson et al., 1978), which claims that learned helplessness occurs when individuals do not experience that highly desired outcomes are contingent on their responses. This may lead to an attribution of challenges and demands as uncontrollable, prompt individuals to use maladaptive coping strategies, such as self-pity,

and thereby a predisposition toward reactive depression when experiencing failure or negative life events (Lazarus and Folkman, 1984; Rotter, 1966). In this study, we will use the concept of *coping competence* as an indicator of coping beliefs. Coping competence is defined as a belief in the capacity to "(…) effectively cope with failure and negative life events as indicated by a reduced likelihood of helplessness reactions and fast recovery from any occurring helplessness symptoms" (Schroder and Ollis, 2013, p. 288). From this perspective, coping beliefs (competence) is understood as a component of emotional well-being and resilience against learned helplessness dispositions and helplessness-based reactive depression.

In sum, coping beliefs seems to be crucial for both the individuals` interpretation of the threat level, their own abilities and possible coping strategies. When the expectation of coping is low, various challenges, demands and changes that we are exposed to will eventually be perceived as stressful, while high coping will make us more robust to cope with challenges we face in life and thus reduce the likelihood that we will experience challenging situations as stressful. Based on this, we expect that coping beliefs is directly and negatively related to perceptions of school stress, appearance pressure and emotional problems (Sowislo and Orth, 2013; Yeoh et al., 2017).

2.3 Appearance and school stress

This article investigates the associations between emotional problems and adolescents' perceptions of performance-related stress in two areas: appearance and school. According to a study that compared the student reports on the perceptions of school pressure between 1994 and 2010, the level of perceived pressure did not increase during the period, except for a temporary uptake between 2002 and 2006. However, recent research indicates that the prevalence of adolescent experience of school stress has increased during the last decade (Bakken et al., 2018; Löfstedt et al., 2019), and that there is a clear association between school stress and emotional problems (Låftman and Modin, 2012; Ringdal et al., 2020).

In a systematical review of 33 studies, Lillejord et al. (2017) report mixed findings concerning gender differences in school stress. Some studies find no differences (Coelho and Romão, 2016; Sotardi, 2016), while most studies find that girls report higher levels of stress compared with boys (Goldstein et al., 2015; Seiffge-Krenke, 2012; Sun et al., 2013; Yang and Cheng, 2016). To explain the differences, Lillejord et al. (2017) argue that girls undergo more pubertal changes than boys, emphasize social demands to a greater extent, become more stressed by unclear expectations and more frequently experience school-related burnout. In line with this, Samela-Aro and Tynkkynen (2012) point to research that has found that girls and boys experience stress differently. Girls internalize symptoms such as depression or fatigue and develop a feeling of not coping with pressure. In addition, girls are more likely to experience competitive situations and that they are more exposed to the negative effects of stress, while boys may tend to develop an externalized reaction pattern such as adapting a cynical attitude toward school. This finding is supported by Suldo and Shaunessy-Dedrick (2013) and Wahab et al. (2013). In sum, it seems that

girls are more influenced by the school context than boys (Dalen, 2014; Låftman et al., 2013). In the high-performing school context that Låftman et al. (2013) report from, students believe that there is a culture of stress at school and that the students, especially girls, influence each other and exaggerate the stress. This is confirmed by Sonmark et al. (2016), who found that being surrounded by classmates who feel pressured is more detrimental to the mental health of girls than boys.

Furthermore, there is an extensive research tradition on adolescents` attitudes toward their own body and appearance. Studies show that girls have a more problematic body image than boys (Bakken et al., 2018; Skaalvik and Federici, 2015). Also, these gender differences increase beyond adolescence (Bearman et al., 2006; Jónsdóttir et al., 2008), and are more closely related to self-concept among girls than boys (Alm and Låftman, 2018). Dissatisfaction with one`s own body and appearance has a stronger association with mental health problems among girls than boys (Hargreaves and Tiggeman, 2003; Bearman and Stice, 2008). In sum, based on these findings, we expect that both appearance pressure and school stress are directly and positively related with emotional problems.

2.4 The present study

The present study was designed to explore how students perceive life in school in general, and the extent of emotional problems in particular. Even though mental health and well-being are becoming a health priority in most western countries, evidence concerning the prevalence and predictive factors over time remains mixed (Cosma et al., 2020; Rasing et al., 2020). Thus, this longitudinal study is designed to contribute to this knowledge gap in the literature. Based on the literature review, the following research question was formulated: "To what extent do students in Upper Secondary School experience emotional problems, and how are these problems predicted by gender, academic/social self-concept, coping competence, appearance pressure and school stress?". Based on this research question and our findings from the literature review, six hypotheses were formulated, and a theoretical path model was specified:

- H1: The extent of emotional problems among adolescents has increased (Bor et al., 2014; Choi, 2018).
- H2: Academic self-concept is directly and positively related with coping competence, and negatively related to emotional problems (Craven and Marsh, 2008; Marsh and Martin, 2011).
- H3: Social self-concept is directly and positively related with coping competence, and negatively related to emotional problems (Kim-Spoon, Ollendick and Seligman, 2012; Lee, Hankin and Mermelstein, 2010).
- H4: Coping competence is directly and negatively related with school stress, appearance pressure and emotional problems (Abramson et al., 1978; Schroder and Ollis, 2013)
- H5: Appearance pressure is directly and positively related with emotional problems (Alm and Låftman, 2018; Bearman and Stice, 2008).

- H6: School stress is directly and positively related with emotional problems (Låftman and Modin, 2012; Ringdal et al., 2020).

3 Method

3.1 Participants and procedure

This study was part of a larger data collection and analysis of students` perceptions of their life in school. The sample for the present article comprises 1077 students (56% response rate), 606 girls and 471 boys, through Upper Secondary School (grades I–III, approximately 16–19 years old), from both rural and urban areas in thirteen schools from one county in Norway. Data can be described as a convenience sample (McQueen and Knussen, 2006), and were collected in the fall of 2015 (USI), spring of 2017 (USII), and spring of 2018 (USIII).

The participating students were informed in advance that the participation in the study was voluntary and that they were considered to have given their consent by filling in the questionnaire. The data were collected with paper-based questionnaires and administrated by members of the research team according to the guidelines of the Norwegian Social Data Services. The Norwegian Data Inspectorate approved the survey.

3.2 Instruments

The dependent variable was measured on a three-point scale ranging from "Not true" (0), "Somewhat true" (1), and "Certainly true" (2). The independent variables, except grades from Lower Secondary School (0–6) and Appearance Pressure (1–4), were measures with items answered on a six-point Likert scale ranging from 1= very untrue to 6= very true, and the reliability was measured by Cronbach`s alpha.

3.2.1 Emotional problems

The dependent variable *Emotional Problems* scale was constructed from the Strength and Difficulties Questionnaire (SDQ). SDQ (Goodman, 1997, 1999) is a brief assessment of mental health problems that consist of 25 questions covering emotional symptoms, conduct problems, hyperactivity-inattention, and peer problems as well as prosocial behavior. SDQ was originally constructed for children between 11 and 16 but has recently been used in older age groups (e.g. Bøe et al., 2016; Vugteveen et al., 2019). SDQ Emotional Problems were measured by five statements that each were scored from 0 to 2, making a possible score from 0 to 10. Examples of statements: "I worry a lot", "I am often unhappy, down-hearted or tearful". Cronbach`s alpha for the scale was good for both girls (.71) and boys (.73).

3.2.2 Academic self-concept

Academic self-concept was measured by four items based on the "Self-Description Questionnaire III (SDQIII), originally developed by Marsh (1992), and acknowledged as a leading multidimensional self-concept instrument for adolescents (e.g. Boyle, 1994; Hattie, 1992). Examples of statements: "I cope with the challenge of learning new material at school", "When I sit down to learn something really difficult, I can handle it". Cronbach`s alpha for the scale was good for both girls (.89) and boys (.86).

3.2.3 Social self-concept

Social self-concept was also measured by four items based on the SDQIII (Marsh, 1992). Examples of statements: "It is easy for me to make friends", "Most people like me". Cronbach`s alpha for the scale was good for both girls (.84) and boys (.87).

3.2.4 Coping beliefs

Coping beliefs was measured by five items from The Coping Competence Questionnaire (Schroder and Ollis, 2013). This instrument had originally 12 statements, and we used the 5 statements with the highest loadings from the factor analysis. All the statements are formulated negatively and the variable score were turned before we made the scale. Examples of statements: "When I do not succeed right away, I think I will never get it", "When I perform poorly at school, I begin to doubt my abilities". Cronbach`s alpha for the scale was good for both girls (.84) and boys (.87).

3.2.5 Appearance pressure

Appearance pressure was measured using two items based on a scale developed by Skaalvik and Federici (2015). Examples of statements: "I have felt a strong pressure to look a certain way in recent months", "I have felt a strong pressure to dress in a special way in recent months". Cronbach`s alpha for the scale was good for both girls (.66) and boys (.71)

3.2.6 School stress

School stress was measured using three items based on a scale developed for the international survey, Health Behavior in School-Aged Children (Samdal et al., 2016), organized and administrated by the World Health Organization. Examples of statements: "I feel exhausted because of schoolwork", "I`m stressed by schoolwork". Cronbach`s alpha for the scale was good for both girls (.79) and boys (.77).

3.3 Data analysis

At first, we compared the full SDQ Emotional problems scale data with British norm-data (Goodman et al., 2000), and compared this with a survey from 2000 that included 19 210 adolescents from Norway (Rødje, 2004).

After this, an exploratory factor analysis was conducted on six sum scales (SDQ emotional problems, academic self-concept, social self-concept, coping competence, appearance pressure, and school stress) with a total of 23 statements (see Method for details). Here we found that one item ("often complains of headaches, stomachaches or sickness") in the SDQ Emotional problem scale loaded on the School stress factor and decided to leave this factor out. This resulted in an exploratory factor analysis with a total of 22 statements. The Kaiser-Meyer-Olkin value was .89, exceeding the recommended value of .6 (Kaiser, 1970; 1972) and Bartlett`s Test of Sphericity (Bartlett, 1954) reached statistical significance, supporting the factorability of the correlation matrix. Principal component analysis revealed the presence of six components with eigenvalues exceeding 1. The six-component solution explained a total of 68% of the variance. To aid the interpretation of these three components, oblimin rotation was performed. The rotated solution revealed the presence of a simple structure (Thurstone, 1947) with all six components showing strong loadings and all variables loading substantially on only one component.

Furthermore, the hypothesized model of the connections between the variables presented in Figure 1 were tested statistically to explore to what degree it was coherent with the observed data. Structural equation modeling (SEM) in the AMOS 26 program was used to analyze the model with latent (unobserved) variables. None of the error terms were allowed to correlate. The coherence between observed data

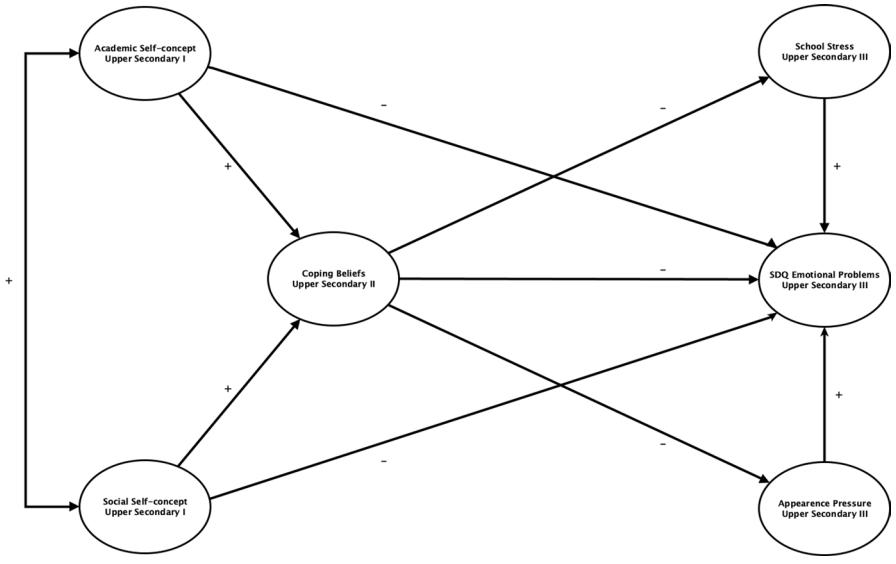


Fig. 1 Theoretical path model

and the hypothesized model is reported as the goodness of fit statistics. The goodness of fit indicators used to assess the model are Non-Normed Fit Index (NNFI, also known as TLI), Comparative Fit Index (CFI), and Root Means Square Error of Approximation (RMSEA), the study also includes chi square, RMSEA $\leq .07$, TLI $\geq .90$, CFI $\geq .90$ are considered as indicators of acceptable fit (Byrne, 2001; Hu and Bentler, 1999). This indicates a plausibility of the associations between the constructs. The model was tested with the whole sample, and then separate analyses were performed for boys and girls, because research indicates gender differences in the importance of the hypothesized relation between constructs.

4 Results

4.1 Distribution of SDQ emotional problems

Table 1 shows the distribution of SDQ Emotional Problems through Upper Secondary School for both girls and boys. These are interpreted based on the original three-band categorization from British norm-data (Goodman et al., 2000), and compared with a survey from 2000 that included 19 210 adolescents from Norway (Rødje et al., 2004).

British norm-data indicate that 90% of adolescents will score within the "Normal" (0–5), 5% within the "Borderline" (6), and 5% within the "Abnormal" categories. The 2000 survey from Norway (Rødje et al., 2004), shows that girls (11%) reported abnormal levels of emotional problems almost three times as much as boys (4%). In addition, girls (8%) reported borderline levels of emotional problems over twice as much as boys (3%). Compared to the British norms we can conclude that boys reported less emotional problems than expected, and girls reported about twice as many abnormal and borderline levels of emotional problems.

Our data show that boys score about the same as expected from the British norm-data and is distributed with almost exactly the same percentages as in the study from Norway (Rødje et al., 2004). However, when interpreting the scores from the girls in our study we can see another development. The girls report about four times as many emotional problems as the British norms and as the boys in our study. In addition, we can see that the extent is almost doubled compared with the 2004 study and that the emotional problems seem to increase as the girls grow older with 22% of

Table 1 Distribution of dependent variable SDQ Emotional problems in US1, US2 and US3 compared with the Akershus study in 2000 and British norms by percentage

Three-band categorization	British norms	Boys				Girls			
		2000	US1	US2	US3	2000	US1	US2	US3
Normal (0–5)	90	93	94	94	91	81	72	71	67
Borderline (6)	5	3	3	3	5	8	9	9	11
Abnormal (7-10)	5	4	3	3	4	11	19	20	22

girls reporting abnormal levels of emotional problems in their last spring of Upper Secondary School. In sum, these findings support H1.

4.2 Zero order correlations between observed variables

Zero order correlations between the study variables as well as N, range, number of items, Cronbach`s alpha, statistical means, standard deviations and Cohen`s D for boys and girls are shown in Table 2.

All correlations between emotional problems and the independent variables were significant for both girls and boys. The correlations between the dependent variable emotional problems and the independent variables academic self-concept, social self-concept and coping competence were negative, ranging from r = −.20 to r = −.57. This indicates that the higher the students perceive emotional problems, the lower they perceive academic self-concept, social self-concept, and coping competence. The correlation between emotional problems and the independent variables appearance pressure and school pressure was positive, indicating that the higher students perceive appearance pressure, the higher they perceive emotional problems. Most of the correlations between the observed variables are similar between girls and boys, however the negative correlation between emotional problems and coping competence seems to be stronger for girls (−.57) than for boys (−.38).

In addition, as the distribution of SDQ emotional problems in Table 1 indicated, the means of the dependent variable were significantly higher for girls (3.53) compared with boys (2.01). A similar difference is the mean for appearance pressure (1.94/1.56) and school stress (4.32/3.58). An opposite pattern is shown for academic self-concept (4.57/4.80), social self-concept (4.43/4.63) and coping competence (4.07/4.72), with higher scores for boys compared with girls. Considering effect size, Cohen (1988) claims that an effect of 0.2 is small, 0.5 is medium, and 0.8 is high. The Cohen's D measure thus indicates that the difference effect size is small for academic self-concept (.29) and social self-concept (.22). Medium for coping competence (.62), appearance pressure (.54), and school stress (.65). The effect size is close to strong for the dependent variable emotional problems (.77).

4.3 Structural equation model

The relations between the variables were further analyzed by means of SEM analysis for latent variables using the AMOS 26 program as shown in the theoretical model in Figure 1. The model had a satisfactory fit to data for both girls and boys: CFI = .945/933, TLI = .931/915, RMSEA = .050/.053. Figure 2 shows the SEM between academic self-concept (US I), social self-concept (US I), coping competence (US II), appearance pressure (US III), school stress (US III), and emotional problems (US III), and Table 3 shows the direct, indirect and total effects, for both girls and boys.

Figure 2 and Table 3 indicate that the correlation between academic self-concept and social self-concept are positive (r = .51) for both girls and boys. Academic self-concept and social self-concept in the first year of Upper Secondary School has a

Table 2 Zero order correlations between observed variables (boys/girls)

Variables	1	2	3	4	5	6
SDQ Emotional Problems US3	–	−.20**/−.26**	−.28**/−.33**	−.38**/-.57**	.33**/.40**	.38**/.44**
Academic self-concept US1		–	.42**/.42**	.32**/.40**	−.05/−.17**	−.19**/−.24**
Social self-concept US1			–	.25**/.32**	−.06/−.12**	−.04/−.12**
Coping competence US2				–	−.30**/−.36**	−.31**/−.41**
Appearance pressure US3					–	.23**/.30**
School Stress US3						–
N	339/561	469/605	466/596	469/603	451/585	403/558
Range	0-8	1-6	1-6	1-6	1-4	1-6
Number of items	4	4	4	5	2	3
Cronbach`s alpha	.71/.73	.80/.82	.89/.86	.84/.87	.66/.71	.79/.77
Mean	2.01/3.53	4.80/4.57	4.63/4.43	4.72/4.07	1.56/1.94	3.58/4.32
SD	1.76/2.13	.76/.83	.94/.87	.97/1.11	.62/.77	1.17/1.12
Cohen`s D	.77	.29	.22	.62	.54	.65

**p<.01

*p<.05

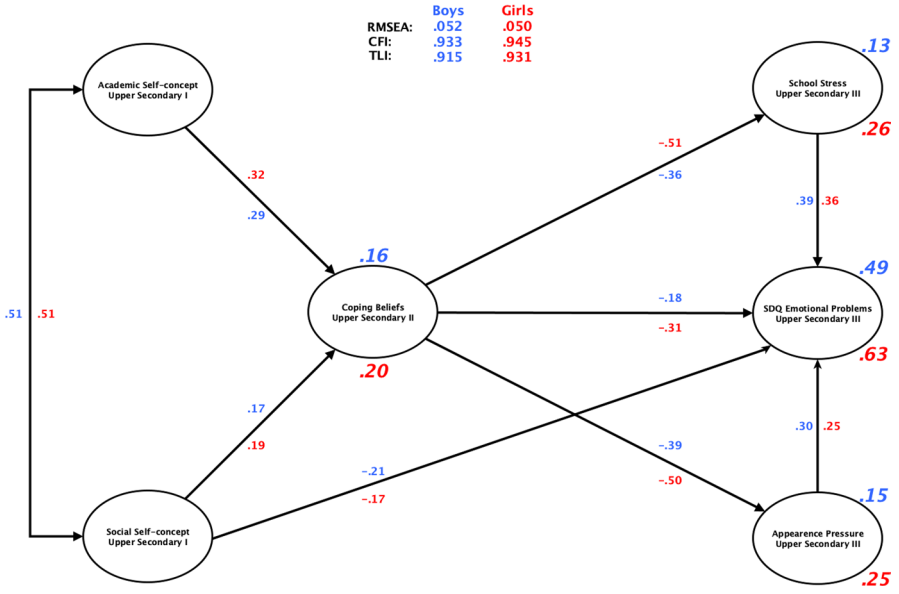


Fig. 2 Structural Equation Model between Grades, academic self-concept, social self-concept, coping competence, appearance pressure, school stress, and emotional problems divided by gender (boys/girls). (Color figure online)

Table 3 Direct, indirect and total effects in the Structural Equation Model for boys/girls

Variables	Effects		
	Direct	Indirect	Total
Academic Self-concept (USI)			
Coping Competence (USII)	29/.32	–	29/.32
Appearance Pressure (USIII)	–	−.11/−.16	−.11/−.16
School Stress (USIII)	–	−.10/−.16	−.10/−.16
Emotional Problems (USIII)	–	−.13/−.20	−.13/−.20
Social Self-concept (USI)			
Coping Competence (USII)	.17/.19	–	.17/.19
Appearance pressure (USIII)	–	−.07/−.10	−.07/−.10
School Stress (USIII)	–	−.06/−.10	−.06/−.10
Emotional Problems (USIII)	−.21/−.17	−.07/−.12	−.28/−.29
Coping Competence (USII)			
Appearance pressure (USIII)	−.36/−.50	–	−.36/−.50
School Stress (USIII)	−.36/−.51	–	−.36/−.51
Emotional Problems (USIII)	−.18/−.31	−.26/−.31	−.44/−.61
Appearance pressure (USIII)			
Emotional Problems (USIII)	.30/.25	–	.30/.36
School pressure (USIII)			
Emotional Problems (USIII)	.39/.36	–	.39/.25

significant direct positive association with coping competence in the second year of Upper Secondary School (β girls/boys = .32/.29 and .19/.17, respectively), and explains 20% and 16% (girls/boys) of the variance in the coping competence variable. Table 3 indicates that the two self-concept variables have indirect effects on appearance pressure, school stress and emotional problems. However, we could not find any direct significant relationships between the self-concept variables and the perception of appearance pressure and school stress. In addition, of the two self-concept variables, it is only social self-concept that has a direct negative relationship with emotional problems (β girls/boys = −.17/−21).

Coping competence has a significant direct negative relationship with appearance pressure (β girls/boys = −.50/−.39), school stress (β girls/boys = −.51/−.36), and emotional problems (β girls/boys =−.31/−.18) for both girls and boys, respectively. In addition, Table 3 indicates that coping competence also has a considerable indirect effect (β girls/boys =−.31/−.26), on emotional problems through appearance pressure and school stress, resulting in a total effect of β =−.61 for girls and β =−.44 for boys. The model explains 25% (girls) and 15% (boys) of the variance in appearance pressure, and 26% (girls) and 13% (boys) of the variance in school stress.

Both appearance pressure (β girls/boys =.25/.30) and school stress (β girls/boys =.36/.39) have a significant positive relationship with emotional problems for both girls and boys, respectively. In all, the structural equation model explains 63% of the variance in emotional problems for the girls, and 49% for the boys.

5 Discussion

The primary goal of this study was to further explore the prevalence of emotional problems, and a proposed pathway between emotional problems among Upper Secondary School students, and their perceptions of academic self-concept, social self-concept, coping competence, appearance pressure and school stress. The following research question guided the study: "To what extent do students in Upper Secondary School experience emotional problems, and how are these problems predicted by gender, academic/social self-concept, coping competence, appearance pressure and school stress?" Based on this research question and our findings from the literature review, six hypotheses were formulated, and a theoretical path model was specified. Our findings mainly confirmed previous studies.

The first main finding is that the prevalence of emotional problems among adolescents has increased considerably compared with the 2000 survey from Norway (Rødje et al., 2004). This supports H1 and previous research (Bor et al., 2014; Choi, 2018). However, even though it is expected, the large effect size difference between boys and girls is somewhat surprising. The zero-order correlations between the dependent and the independent variables reveal that the relationships are somewhat the same for both girls and boys, with the exception of the association between emotional problems and coping beliefs (competence). This relationship is stronger for girls compared to boys, indicating that girls may be more vulnerable for developing learned helplessness (Abramson et al., 1978; Schroeder and Ollis, 2013). This notion is supported when analyzing the Cohen's D effect size difference between the

genders. Girls perform significantly higher than boys in Lower Secondary School, but still they have lower academic self-concept and social self-concept. In addition, the coping belief, measured as coping competence, is considerably lower for girls compared to boys, while the perception of appearance pressure and school stress are higher to the same degree. This imbalance may explain the high effect size difference in emotional problems. In sum, this finding corresponds well with other research that has found that girls are at greater risk of developing emotional problems, reporting lower levels of self-concept, coping beliefs, appearance pressure and school stress (Caprara et al., 2010; Muris et al., 2015).

The other main finding is that the proposed pathway in the theoretical model presented in Figure 1 had a good fit with the data, especially for girls, as it explained 63% of the variance in emotional problems. This indicates that academic self-concept and social self-concept, coping competence, and the perceptions of appearance pressure and school stress predicts emotional problems among adolescents in a prominent manner. In support of earlier research, the correlation between the self-concept variables was high (Marsh and Craven 2006; Marsh and Martin, 2011). In addition, Shavelson et al.`s (1976) theory on the multidimensional nature of self-concept was underlined by our finding about the difference of whether, and with what strength, academic self-concept and social self-concept predicted the other variables in the model. As expected, the self-concept variables predicted the adolescents' perception of coping competence in USII in similar strength for both girls and boys. This supports H2, H3 and previous research (Craven and Marsh, 2008; Marsh and Martin, 2011). Furthermore, it was expected that social self-concept in USI was significant for both girls' and boys' perceptions of emotional problems in USIII. However, in contrast with previous findings on the longitudinal effects academic self-concept has on emotional problems (Craven and Marsh, 2008; Marsh and Martin, 2011), we could not find any significant direct negative association to emotional problems, even though there were some indirect effects through coping beliefs. This finding is surprising and needs further scrutiny in future research

Another interesting finding from the structural equation modeling is that coping competence has a significant negative relationship with appearance pressure, school stress and emotional problems for both girls and boys. However, the strength of these direct and indirect associations is somewhat stronger for girls compared with boys, indicating that coping competence seems to be important for both genders, and most important for girls, when it comes to predicting perceptions of pressure and negative health outcomes. These findings support H4 and previous research (Sowislo and Orth, 2013; Yeoh et al., 2017). In addition, in support of H5, H6 and previous research (Alm and Låftman, 2018; Bearman and Stice, 2008; Låftman and Modin, 2012; Ringdal et al., 2020) appearance pressure and school stress had a direct and positive relationship with emotional problems. However, it was more surprising that these associations were somewhat stronger for boys compared with girls, even though the differences were small.

In sum, the findings can be discussed in light of other studies that have shown that the level of coping beliefs plays a significant role in how different kinds of perceived demands are experienced (e.g. appearance pressure and school stress), and affects the risk that stress leads to emotional problems (Caprara et al., 2004; Shelley

and Pakenham, 2004). Several studies conclude that increasing individuals` coping beliefs and expectations of coping may reduce the negative effect of the various stressors we are exposed to (e.g. McKay, Dempster and Byrne, 2014). In a longitudinal study, Burger and Samuel (2017) investigated how perceived stress and coping beliefs play a role in influencing adolescents` satisfaction with life. They found that perceived stress was negatively correlated with life satisfaction, whereas the opposite was true of coping beliefs. In addition, the individuals` basic expectation of coping further moderated the negative effects of stress. Thus, a high level of coping beliefs neutralized the negative effects of stress on adolescents` satisfaction with life. Furthermore, our coping beliefs today are related to the coping experiences we have had in the past (Groth et al., 2019; Samdal et al., 2017). A situation where we are exposed to stress stimuli can result in both positive and negative response outcomes. Positive outcomes represent situations in which individuals experience mastery. They have experienced being challenged, but also that they can cope with it. Negative outcomes are characterized by individuals experiencing that they do not cope. At worst, they will experience helplessness ("I can`t do anything about the situation"), and hopelessness ("things go wrong anyway and it`s my own fault") (cf. Abramson et al., 1978; Schroder and Ollis, 2013). A number of such negative outcomes contribute to both low coping beliefs and sustained stress activations, which in turn can lead to negative health outcomes.

5.1 Perceptions of internal/external demands/resources as predictors of emotional problems

Why do we observe this recent increase in mental health problems? How can this be explained? In light of Lazarus and Folkman's (1984) transactional model of stress the causes of mental health problems can be placed in two main groups: resources and demands. These can each be further divided in two: internal and external. To maintain emotional health, resources and demands have to be perceived as balanced by the individual. Thus, to reduce emotional health problems, we have to consider how we can enhance the adolescents' perceptions of internal and external resources, and reduce the perception of internal and external demands.

Considerable research has investigated how to enhance adolescents' internal and external resources, especially in a school context. The first meta-review over research on mental health promotion and problem prevention in schools was conducted by the European Union Dataprev project (Weare and Nind, 2011). This project included 52 systematic reviews and meta-analyses of mental health in schools and the results show that interventions have a wide variety of beneficial effects on children, families and communities on a range of social, emotional and educational outcomes. A conservative estimate from the London School of Economics and Public Health England shows that for every pound Britain invests in social and emotional learning (SEL) school programs to prevent depression among children and young people, the UK saves five times the investment cost in just two years (McDaid et al., 2017). These findings were supported by a meta-analysis of 81 studies compromising 31 974 school students (Werner-Seidler et al., 2017), and another

meta-analysis of 82 school-based social and emotional learning programs (Taylor et al., 2017). Although the effect sizes are small, these studies provide solid evidence that mental health can be strengthened, and emotional problems can be prevented, through school measures. The best-tested school programs are based on social and emotional learning, cognitive-behavioral, or multifamily therapy (MFT). A meta-analysis of 146 RCT-studies found that it is not sufficient to disseminate information only, e.g. through teaching materials or with educational or psycho-educational measures (Stockings et al., 2016). Research suggests that the measures should also contain a psychological component with an element of practice, such as training in cognitive techniques. In addition, Mendelson and Eaton's (2018) meta-analysis indicates that internet- or computer-based prevention strategies and mindfulness-based interventions are promising areas for further development and need further research.

On the other hand, there are a number of increased internal and external demands that twentifirst century adolescents experience that potentially can have negative consequences for their emotional well-being by contributing to the increasing prevalence of mental health problems (Bor et al., 2014; Seiffge-Krenke et al., 2009). For adolescents living in a world characterized by rapid changes and technological advancement, it is important to understand that different trends and factors can be related to their emotional well-being and ill-being. This does not mean that the fundamental elements of mental health have changed. However, as a result of various contexts and the environment that children face today, different trends arise, and accordingly various approaches exist to mitigate the negative effects. An example is the Programme for International Student Assessment (PISA), that was introduced for the first time in 2000. This was aimed to provide a "global" comparison across nations to determine the relative quality of each country`s educational system (OECD, 2001). This cross-national comparison led to considerable debate, especially among the countries with the lowest scores, and facilitated reforms in educational policy and practice, and enhanced expectations, demands and competition among schools and students (Cosma et al., 2020; Klinger et al., 2015). In addition, several theorists argue that too much responsibility and accountability has been placed on the internal and individual. Research has found a rise in internal perfectionism expectations and demands connected to both school and appearance that may contribute to increases in stress and subsequent mental health problems (Cosma et al., 2020; Curran and Hill, 2019). This finding is in line with the individualization thesis that is based on the idea that collective identities associated with class, family and gender are weakened in favor of individual aspirations that have become more important (Beck, 1992). This notion is relevant for the increase in emotional health problems in several ways. First, it is up to each individual to create their own future, find out "who one is", set goals, and reach them (Giddens, 1991; Illeris et al., 2009). Consequently, the responsibility for success and for one`s own destiny also increases, especially within the school system. Second, the individualization thesis is relevant in this context because mental health problems themselves may be understood to be more individualized than before. While conditions such as neurasthenia, melancholy, sadness and nervousness were perceived as a consequence of culture a hundred years ago, it is claimed that similar conditions today are to a greater extent interpreted as a weakness of the individual.

Overall, our analyses of the two pressure areas: school and appearance, indicate that girls are more vulnerable than the boys. In both of these areas, the pressure is more closely associated with emotional problems among girls than boys. Thus, this indicates that school stress, pressure related to appearance, and coping beliefs is potentially a greater health risk for girls than for boys. One possible explanation is that girls, to a higher degree than boys, reflect their intrinsic value on how they succeed in school (Esnaola et al., 2018; Låftman and Modin, 2012). There is also reason to believe that the perception of personal appearance is more important for the girls` self-image than for boys (Alm and Låftman, 2018; Esnaola et al., 2018). Although many argue that the ideals for body and appearance also have changes for boys (Bassett-Gunter et al., 2017; Karazsia et al., 2017), these findings indicate that looking good is still more important for girls compared with boys. Thus, it may be that girls actually experience more pressure in these areas, and also interpret the pressure as more threatening to their own identity. Because external stressors are perceived and interpreted subjectively, based on personal values, the social roles we hold, and from past experiences, the same amount of actual demands and expectations can be experienced without problems for some and pose a health risk for to others (Avison, 2010; Thoits, 2010). At the same time, the potential for experiencing a situation with a lot of pressure such as stress may be explained by gender differences in the amount of resources available in the situation and to differences in the coping strategies available to girls and boys. However, other research indicates that the suicidal rates are far higher among men compared with women (Oliffe et al., 2016; Rasmussen et al., 2017). This may indicate that stress and emotional problems among adolescent boys are underreported. Thus, we need more research to investigate long-term effects of perceptions of demands and resources across genders through childhood, adolescence and adulthood.

5.2 Limitations and future directions

This study has some limitations that should be acknowledged. First, all data were based on students` self-reports. Future studies should employ more objective measures of the social environment (e.g. behavioral observation), triangulate subjective perceptions (e.g. different informants such as friends, parents, teachers, coaches etc.), and make use of both quantitative and qualitative methods. Second, the data from this study are limited to the school context. Future studies should test the associations between environmental factors from several micro-systems, subjective needs and adolescents` emotional health via longitudinal and experimental designs. Third, this study has only considered individual-level variables, without considered the effect of the differences between social classes and schools on perceived ecology with relation to emotional health among individuals. Several studies have revealed that differences in the classroom or school environment exist across school level (e.g. Danielsen et al., 2010; Wang et al., 2018). Future studies may consider using a multi-level modeling approach to study the impact of class-, school- and societal-level factors.

Despite these limitations, the results in our study may have important clinical implications for current interventions aimed at preventing and reducing mental health problems in school. Our findings suggest that targeting perceptions of self-concept and coping beliefs will reduce feelings of pressure, stress and emotional problems. Integrative interventions should focus on enhancing adolescents` self-concept and coping beliefs as early as possible to disrupt any potential detrimental cascading effect on mental health. This is in line with the findings of intervention studies that reveal that coping beliefs seems to be an important mechanism to improve mental health outcomes (Kendall et al., 2016; Ohannessian et al., 2019).

A comprehensive understanding of pressures and stress among young people today should include knowledge of who is the agent behind the significant amount of pressure young people experience in everyday life. Adolescents can be both passive recipients of experienced pressure and active agents that exercise resistance. Knowledge about the various actors is not least important. Thus, teachers, counsellors and schools should aim to find good methods and structures that reduce the amount of pressure or make youth more robust against the pressure they experience in everyday life. To what extent do adolescents themselves contribute to the pressure they experience, and to what extent are they actors in reducing the pressure? Is this the solution to teaching each one to prioritize better in their own lives? To what extent and in what arenas do adults, such as parents, teachers, coaches, school authorities and commercial advertisers have a responsibility? And to what extent does pressure from the adult community work together with peer pressure—both analog and digital—to increase the amount of pressure and the vulnerability of that pressure? These are important issues for the social debate and should be explored further in future research.

Several replication studies are needed that can confirm, disprove, or nuance findings from earlier studies. Studies with multiple data sources are needed, such as interviews with parents, teachers, counsellors and school leaders, which can be combined with self-reported data. Future research should also examine a longer-term follow-up of comprehensive data on prevalence and outcomes, encompass more high quality rigorous randomized control trial designs with comparable control groups, and also study the potential effectiveness and feasibility of teachers delivering these program sessions and training. The aspect of investigating teacher delivery of efforts to improve students' mental health is important. As several studies have pointed out, teachers can play a vital role in delivering such prevention programs. Future studies should also investigate the practicality of implementation in terms of costs, additional burden on teachers, and the necessary support that teachers need to deliver this training and programs. In order to prevent an increase in teachers' workloads, which can reduce a teacher's emotional well-being and in turn that of students, it will be important think about ways in which social and emotional learning can be incorporated into the curriculum and existing day-to-day learning activities.

6 Conclusion

Although this study has some limitations, the results in the present study underscore the need for research on self-concept, coping beliefs and environmental pressure when it comes to understanding adolescents' coping and emotional problems. In sum, our findings can be illuminated in light of other research that has suggested that girls perceive more pressure than boys and at the same time are more vulnerable to the effects of this perception (Bakken et al., 2018; Hankin et al., 2008). Thus, girls perceive their external and internal demands as higher than boys but their external and internal resources are perceived to be lower. In light of the transactional theory of stress (Lazarus and Folkman, 1984) and the theory of learned helplessness (Abramson et al., 1978), these unfortunate connections may explain why girls are developing higher levels of emotional problems. These findings are of great significance for practical implications and applications. Teachers, counsellors and the school as a whole should increase their effort to facilitate the development of psychological resilience as this seems to be both a decisive protective- and risk-factor when it comes to adolescents' sense of stress and emotional well-being, especially among girls.

Funding Open access funding provided by NTNU Norwegian University of Science and Technology (incl St. Olavs Hospital - Trondheim University Hospital). The authors received no financial support for the research, authorship, and/or publication of this article.

Declarations

Conflict of interest The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Informed consent Informed consent was obtained from all participants, Additionally, parents were sent an information letter with the option to opt their child out of the study. The study was approved by the Norwegian Data Inspectorate

References

Abramson, L. Y., Seligman, M. E., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology*, *87*(1), 49–74.

Alm, S., & Låftman, S. B. (2018). The gendered mirror on the wall: Satisfaction with physical appearance and its relationshop to global self-esteem and psychosomatic complaints among adolescent boys and girls. *Young, 26*(5), 525–541.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association.

Auerbach, R. P., Mortier, P., Bruffaerts, R., Alonso, J., Benjet, C., Cuijpers, P., Demyttenaere, K., Ebert, D. D., Green, J. G., Hasking, P., Murray, E., Nock, M. K., Pinder-Amaker, S., Sampson, N. A., Stein, D. J., Vilagut, G., Zaslavsky, A. M., & Kessler, R. C. (2018). WHO World Mental Health Surveys International College Student Project: Prevalence and Distribution of Mental Disorders. *Journal of Abnormal Psychology, 127*(7), 623–638.

Avison, W. R. (2010). Incorporating children`s lives into a life course perspective on stress and mental health. *Journal of Health and Social Behavior, 51*(4), 361–375.

Bakken, A., Sletten, M. A., & Eriksen, I. M. (2018). Generasjon prestasjon? Ungdoms opplevelse av press og stress [Generation achievement? Adolescents` experience of pressure and stress]. *Norwegian Journal of Youth Research, 18*(2), 46–76.

Balazs, J., Miklósi, M., Kereszteny, A., Hoven, C. W., Carli, V., Wasserman, C., & Cosman, D. (2013). Adolescent subthreshold-depression and anxiety: Psychopathology, functional impairment and increased suicide risk. *Journal for Child Psychology and Psychiatry, 54,* 670–677.

Bandura, A., Pastorelli, C., Barbaranelli, C., & Caprara, G. V. (1999). Self-efficacy pathways to childhood depression. *Journal of Personality and Social Psychology, 76*(2), 258–269.

Bartlett, M. S. (1954). A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society, 16,* 296–298.

Bassett-Gunter, R., McEwan, D., & Kamarhie, A. (2017). Physical activity and body image among men and boys: A meta-analysis. *Body Image, 22,* 114–128.

Bearman, S. K., & Stice, E. (2008). Testing a gender additive model: The role of body image in adolescent depression. *Journal of Abnormal Child Psychology, 36*(8), 1251–1263.

Bearman, S. K., Martinez, E., & Stice, E. (2006). The skinny on body dissatisfaction: A longitudinal study of adolescent girls and boys. *Journal of Youth and Adolescence, 35*(2), 217–229.

Beck, U. (1992). *Risk society: Towards a new modernity*. Sage.

Bøe, T., Hysing, M., Skogen, J. C., & Breivik, K. (2016). The Strengths and Difficulties Questionnaire (SDQ): Factor Structure and Gender Equivalence in Norwegian Adolescents. *PLoS One, 11*(5), 1–15.

Bor, W., Dean, A. J., Najman, J., & Hayatbakhsh, R. (2014). Are child and adolescent mental health problems increasing in the 21st century? A systematic review. *Australian and New Zealand Journal of Psychiatry, 48*(7), 606–616.

Boyle, G. J. (1994). Self-Description Questionnaire II: A review. *Test Critiques, 10,* 632–643.

Burger, K., & Samuel, R. (2017). The Role of Perceived Stress and Self-Efficacy in Young People`s Life Satisfaction: A Longitudinal Study. *Journal of Youth and Adolescents, 46*(1), 78–90.

Byrne, B. M. (2001). *Structural equation modelling with AMOS Basic concepts applications and programming*. Mahwah: Lawrence Erlbaum Ass.

Caprara, G. V., Barbaranelli, C., Pastorelli, C., & Cervone, D. (2004). The contribution of self-efficacy beliefs to psychosocial outcomes in adolescence: Predicting beyond global dispositional tendencies. *Personality and Individual Differences, 37*(4), 751–763.

Caprara, G. V., Alessandri, G., Di Giunta, L., Panerai, L., & Eisenberg, N. (2010). The constribution of agreeableness and self-efficacy beliefs to prosociality. *European Journal of Personality, 24,* 36–55.

Choi, A. (2018). *Emotional well-being of children and adolescents: Recent trends and relecvant factors*. OECD Education Working Paper No. 169. in Portuguese Middle School Transition: A Multilevel Analysis. *The Spanish Journal of Psychology, 19*(e61), 1-8.

Coelho, V. A., & Romão, A. M. (2016). Stress in Portuguese Middle School Transition: A Multilevel Analysis. *The Spanish Journal of Psychology, 19*(e61), 1–8.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale: L. Erlbaum Associates.

Collishaw, S. (2015). Annual Research Review: Secular trends in child and adolescent mental health. *Journal of Child Psychology and Psychiatry, 56,* 370–393.

Condly, S. J. (2006). Resilience in children: A review of literature with implications for educators. *Urban Education, 41*(3), 211–236.

Cosma, A., Stevens, G., Martin, G., Duinhof, E. L., Walsh, S., Garcia-Moya, I., Költö, A., Gobina, I., Canale, N., Catunda, C., Inchley, J., & de Looze, M. (2020). Cross-National Time Trends in

Adolescent Mental Well-Being from 2002 to 2018 and the Explanatory Role of Schoolwork Pressure. *Journal of Adolescent Health, 66*(6), 50–58.

Craven, R. G., & Marsh, H. W. (2008). The centrality of the self-concept construct for psychological wellbeing and unlocking human potential: Implications for child and educational psychologists. *Educational and Child Psychology, 25,* 104–118.

Curran, T., & Hill, A. P. (2019). Perfectionism is increasing over time: A meta-analysis of birth cohort differences from 1989 to 2016. *Psychological Bulletin, 145*(4), 410–429.

Currie, C., Zanotta, C., Morgan, A., Currie, D., de Looze, M., Roberts, C., Samdal, O., Smith, O. R. F., & Barnekow, V. (2012). *Social determinants of health and well-being among young people. Health behavior in school-aged children (HBSC) study: International report from the 2009/2010 survey* (Report nr.6).: WHO Regional Office for Europe.

Dalen, J. D. (2014). Gender differences in the relationship between school problems, school class context and psychological distress: results from the Young-HUNT 3 study. *Social Psychiatry and Psychiatric Epidemiology, 49*(2), 183–191.

Danielsen, A. G., Wiium, N., Wilhelmsen, B. U., & Wold, B. (2010). Perceived support provided by teachers and classmates and students` self-reported academic initiative. *Journal of School Psychology, 48*(3), 247–267.

De Fraine, B., Van Damme, J. V., & Onghena, P. (2007). A longitudinal analysis fo gender differences in academic self-concept and language achievement: A multivariate multilevel latent growth approach. *Contemporary Educational Psychology, 32,* 132–150.

Erskine, H. E., Moffit, T. E., Copeland, W., Costello, E. J., Ferrari, A. J., Patton, G., Degenhardt, L., Vos, T., Whiteford, H. A., & Scott, J. G. (2015). A heavy burden on young minds: the global burden of mental and substance use disorders in youth. *Psychological Medicine, 45,* 1551–1563.

Esnaola, I., Elosua, P., & Freeman, J. (2018). Internal structure of academic self-concept through the Self-Description Questionnaire II-Short (SDQII-S). *Learning and Individual Differences, 62,* 174–179.

Esnaola, I., Sesé, A., Antonio-Agirre, I., & Azpiazu, L. (2020). The development of multiple self-concept dimensions during adolescence. *Journal of Research on Adolescence, 30,* 100–114.

Fletcher, S., & Sarkar, M. (2013). Psychological resilience: A review and critique of defnitions, concepts, and theory. *European Psychologist, 18*(1), 12–23.

Giddens, A. (1991). *Modernity and Self-Identity. Self and Society in the Late Modern Age*. Stanford University Press.

Goldstein, S. E., Boxer, P., & Rudolph, E. (2015). Middle School Transition Stress: Links with Academic Performance, Motivation, and School Experience. *Contemporary School Psychology, 19*(1), 21–29.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry, 38,* 581–586.

Goodman, R. (1999). The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *Journal of Child Psychology and Psychiatry, 40,* 791–799.

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000). Using the Strength and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *The British Journal of Psychiatry, 177*(6), 534–539.

Groth, N., Schnyder, N., Kaess, M., Markovic, A., Rietschel, L., Moser, S., Michel, C., Schultze-Lutter, F., & Schmidt, S. J. (2019). Coping as a mediator between locus of control, competence beliefs, and mental health: A systematic review and structural equation modeling meta-analysis. *Behavior Research and Therapy, 121,* 1–16.

Halvarsson-Edlund, K., Sjödén, P.-O., & Lunner, K. (2008). Prediction of disturbed eating attitudes in adolescent girls: A 3-year longitudinal study of eating patterns, self-esteem and coping. *Eating and Weight Disorders, 13*(2), 87–94.

Hankin, B. L., Wetter, E., & Cheely, C. (2008). Sex differences in child and adolescent depression: A developmental psychopathological approach. In J. R. Z. Abela & B. L. Hankin (Eds.), *Handbook of child adolescent depression* (pp. 377–414). Guildford Press.

Hankin, B. L., Young, J. F., Abela, J. R. Z., Smolen, A., Jenness, J. L., Gulley, L. D., Technow, J. R., Gottlieb, A. B., Cohen, J. R., & Oppenheimer, C. W. (2015). Depression from childhood into late adolescence: Influence of gender, development, genetic susceptibility, and peer stress (2015). *Journal of Abnormal Psychology, 124*(4), 803–816.

Hargreaves, D., & Tiggemann, M. (2003). The effect of "thin ideal" television commercials on body dissatisfaction and schema activation during early adolescence. *Journal of Youth and Adolescence, 32*(5), 367–373.

Hattie, J. (1992). *Self-concept*. Erlbaum.

Henderson, J. L., Cheung, A., Cleverley, K., Chaim, G., Moretti, M. E., de Oliveira, C., Hawke, L. D., Willan, A. R., & O`Brien, D., Heffernan, O., Herzog, T., Courey, L., McDonald, H., Grant, E., & Szatmari, P. . (2017). Integrated collaborative care teams to enhance service delivery to youth with mental health and substance use challenges: protocol for a pragmatic randomised controlled trial. *BMJ Open, 7*(2), 1–11.

Herrman, H., Stewart, D. E., Diaz-Granados, N., Berger, E. L., Jackson, B., & Yuen, T. (2011). What Is Resilience? *The Canadian Journal of Psychiatry, 56*(5), 258–265.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance struture analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.

Illeris, K., Katznelson, N., Nielsen, J. C., Simonsen, B., & Sørensen, N. S. (2009). *Ungdomsliv: Mellem individualisering og standardisering. [Adolescent life: Between individualization and standardization]*. København: Samfundslitteratur.

Jongen, C. S., McCalman, J., & Bainbridge, R. G. (2019). A Systematic Scoping Review of the Resilience Intervention Literature for Indigenous Adolescents in CANZUS Nations. *Frontiers in Public Health, 7,* 1–16.

Jónsdóttir, S. R., Arnarson, E. Ö., & Smári, J. (2008). Body esteem, perceived competence and depression in Icelandic adolescents. *Nordic Psychology*, *60*(1), 58–71.

Kaiser, H. (1970). A second generation Little Jiffy. *Psychometrika, 35,* 401–415.

Kaiser, H. (1972). An index of factorial simplicity. *Psychometrika, 39,* 31–36.

Karazsia, B. T., Murnen, S. K., & Tylka, T. L. (2017). Is Body Dissatisfaction Changing Across Time? *A Cross-Temporal Meta-Analysis. Psychological Bulletin, 143*(3), 293–320.

Kendall, P. C., Cummings, C. M., Villabø, M. A., Narayanan, M. K., Treadwell, K., Birmaher, B., Compton, S., Piacentini, J., Sherill, J., Walkup, J., Gosch, E., Keeton, C., Ginsburg, G., Suveg, C., & Albano, A. M. (2016). Mediators of Change in the Child/Adolescent Anxiety Multimodal Treatment Study. *Journal of Consulting and Clinical Psychology, 84*(1), 1–14.

Kieling, C., Baker-Henningham, H., Belfer, M., Conti, G., Ertem, I., Omigbodun, O., Rohde, L. A., Srinath, S., Ulkuer, N., & Rahman, A. (2011). Child and adolescent mental health worldwide: Evidence for action. *The Lancet, 378*(9801), 1515–1525.

Kim-Spoon, J., Ollendick, T. H., & Seligman, L. D. (2012). Perceived competence and depressive symptoms among adolescents: The moderating role of attributional style. *Child Psychiatry and Human Development, 43,* 612–630.

Klinger, D. A., Freeman, J. G., Bliz, L., Liiv, K., Ramelow, D., Sebok, S. S., Samdal, O., Dür, W., & Rasmussen, M. (2015). Cross-national trends in perceived school pressure by gender and age from 1994 to 2010. *European Journal of Public Health, 25*(2), 51–56.

Låftman, S. B., & Modin, B. (2012). School-performance indicators and subjective health complaints: are there gender differences? *Sociology of Health and Illness, 34*(4), 608–625.

Låftman, S. B., Almquist, Y. B., & Östberg, V. (2013). Students` accounts of school-performance stress: a qualitative analysis of a high-achieving setting in Stockholm Sweden. *Journal of Youth Studies, 16*(7), 932–949.

Lazarus, R., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer.

Lee, A., Hankin, B. L., & Mermelstein, R. J. (2010). Perceived social competence, negative social interactions, and negative cognitive style predict depressive symptoms during adolescence. *Journal of Clinical Child and Adolescent Psychology, 39,* 603–615.

Lerner, R. M., Almerigi, J. B., Theokas, C., & Lerner, J. V. (2005). Positive youth development a view of the issues. *The Journal of Early Adolescence, 25*(1), 10–16.

Lillejord, S., Børte, K., Ruud, E., & Morgan, K. (2017). *Stress i skolen – en systematisk kunnskapsoversikt [Stress in school – A systematic review]*. Kunnskapssenter for utdanning.

Löfstedt, P., Eriksson, C., Potrebny, T., Välimaa, R., Thorsteinsson, E. B., Due, P., Damsgaard, M. T., Suominen, S., Rasmussen, M., & Torsheim, T. (2019). Trends in perceived school stress among adolescents in five Nordic countries 2002–2014. *Nordic Welfare Research, 4*(2), 101–112.

Lopez, D. F., & Little, T. D. (1996). Children`s action-control beliefs and emotional regulation in the social domain. *Developmental Psychology, 32*(2), 299–312.

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology, 81,* 417–430.

Marsh, H. W. (1992). *Self Description Questionnaire (SDQ) III: A theoretical and empirical basis for the measurement of multiple dimensions of late adolescent self-concept: An interim test manual and a research monograph*. University of Western Sydney, Faculty of Education.

Marsh, H. W., & Craven, R. (1997). Academic self-concept: Beyond the dustbowl. In G. Phye (Ed.), *Handbook of classroom assessment: Learning, achievement, and adjustment* (pp. 131–198). Academic Press.

Marsh, H. W., & Craven, R. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science, 1,* 133–163.

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology, 81,* 59–77.

Marsh, H. W., Ellis, L. A., Parada, R. H., Richards, G., & Heubeck, B. G. (2005). A short version of the Self Description Questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analyses. *Psychological Assessment, 17,* 81–102.

McCarty, C. A., Stoep, A. V., & McCauley, E. (2007). Cognitive features associated with depressive symptoms in adolescence: Directionality and specificity. *Journal of Clinical Child and Adolescent Psychology, 36,* 147–158.

McDaid, D., Park, A.-L., Knapp, M., Wilson, E., Rosen, B., & Beecham, J. (2017). *Commissioning cost-effective services for promotion of mental health and wellbeing and prevention of mental ill-health*. Public Health England.

McKay, M. T., Dempster, M., & Byrne, D. G. (2014). An examination of the relationship between self-efficacy and stress in adolescents: the role of gender and self-esteem. *Journal of Youth Studies, 17*(9), 1131–1151.

McQueen, R. A., & Knussen, C. (2006). *Introduction to Research Methods and Statistics in Psychology*. Pearson.

Mendelson, T., & Eaton, W. W. (2018). Recent advances in the prevention of mental disorders. *Social Psychiatry and Psychiatric Epidemiology, 53,* 325–339.

Merikangas, K. R., He, J.-P., Brody, D., Fisher, P. W., Bourdon, K., & Koretz, D. S. (2010). Prevalence and treatment of mental disorders among US children in the 2001–2004 NHANES. *Pediatrics, 125*(1), 75–81.

Muris, P., Meesters, C., Bouwman, L., & Notermans, S. (2015). Relations among behavioral inhibition, shame- and guild-proneness, and anxiety disorders symptoms in non-clinical children. *Child Psychiatry and Human Development, 46,* 209–216.

Mykletun, A., Knudsen, A. K., & Mathisen, K. S. (2009). *Psykiske lidelser i Norge: Et folkehelseperspektiv. [Mental disorders in Norway: A public health perspective]*. Oslo: Folkhehelseinstituttet.

Nagy, B., Watt, H. M. G., Eccles, J. S., Trautwein, U., Lüdtke, O., & Baumert, J. (2010). The development of students` mathematics self-concept in relation to gender. Different countries, different trajectories? *Journal of Research on Adolescence, 20,* 482–506.

OECD. (2001). *Knowledge and Skills for Life: First Results from the OECD Programme for International Student Assessment (PISA) 2000*. France: OECD.

Ohannessian, C. M., Lerner, R. M., Eye, A., & Lerner, J. V. (1999). Does self-competence predict gender differences in adolescent depression and anxiety? *Journal of Adolescence, 22,* 397–411.

Ohannessian, C. M., Vannucci, A., Lincoln, C. R., Flannery, K. M., & Trinh, A. (2019). Self-competence and depressive symptoms in middle-late adolescence: Disentangling the direction of effect. *Journal of Resarch on Adolescence, 29*(3), 736–751.

Oliffe, J. L., Ogrodniczuk, J. S., Gordon, S. J., Creighton, G., Kelly, M. T., Black, N., & Mackenzie, C. (2016). Stigma in Male Depression and Suicide: A Canadian Sex Comparison Study. *Community Mental Health Journal, 52,* 302–310.

Orth, U., Robins, R. W., Widamann, K. F., & Conger, R. D. (2014). Is low self-esteem a risk factor for depression? Findings from a longitudinal study of Mexican-origin youth. *Developmental Psychology, 50*(2), 622–633.

Pérez, J. E., Little, T. D., & Henrich, C. C. (2009). Spirituality and depressive symptoms in a school-based sample of adolescents: A longitudinal examination of mediated and moderated effects. *Journal of Adolescent Health, 44*(4), 380–386.

Pinxten, M., De Fraine, B., Van Damme, J., & D` Haenens, E. . (2013). Student achievement and academic self-concept among secondary students in Flanders: Gender and change over time. *Irish Educational Studies, 32,* 157–178.

Rasing, S. P., Braam, M. W., Brunwasser, S. M., Janssens, J. M., Creemers, D. H., & Scholte, R. H. (2020). Depression and anxiety symptoms in female adolescents: Relations with parental psychopathology and parenting behavior. *Journal of Research on Adolescence, 30*(1), 298–313.

Rasmussen, M. L., Hjelmeland, H., & Dieserud, G. (2017). Barriers toward help-seeking among young men prior to suicide. *Death Studies, 42*(2), 96–103.

Ringdal, R., Espnes, G. A., Eilertsen, M.-E.B., Bjørnsen, H. N., & Moksnes, U. K. (2020). Social support, bullying, school-related stress and mental health in adolescence. *Nordic Psychology, 72,* 1–18.

Rødje, K., Clench-Aas, J., van Roy, B., Holmboe, O., & Müller, A. M. (2004). *Helseprofil for barn og ungdom i Akershus: Ungdomsrapport [Health profile for children and adolescent in Akershus: Adolescent report] (2/2004)*. Nasjonalt kunnskapssenter for helsetjenesten.

Rotter, J. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*(1), 1–28.

Samdal, O., Mathisen, F. K. S, Torsheim, T., Diseth, Å. R., Fismen, A.-S., Larsen, T., Wold, B. & Årdal, E. (2016). *Helse og trivsel blant barn og unge: Resultater fra den landsrepresentative spørreundersøkelsen «Helsevaner blant skoleelever. En WHO-undersøkelse i flere land».* [Health and well-being among children and adolescents: Results from the national representative study "Health habits among pupils. A WHO survey in several countries"]. Universitetet i Bergen: HEMIL-senteret.

Samdal, O., Wold, A. Harris., & og T. Torsheim, . (2017). *Stress og mestring [Stress and Coping]*. Universitetet i Bergen.

Samela-Aro, K., & Tynkkynen, L. (2012). Gendered pathways in school burnout among adolescents. *Journal of Adolescence, 35*(4), 929–939.

Schroder, K. E. E., & Ollis, C. L. (2013). The Coping Competence Questionnaire: A measure of resilience to helplessness and depression. *Motivation & Emotion, 37,* 286–302.

Seiffge-Krenke, I. (2012). Competent youth in a «disorderly world": Findings from an eighteen-nation study. *New Directions for Student Leadership, 2012*(135), 107–117.

Seiffge-Krenke, I., Aunola, K., & Nurmi, J. E. (2009). Changes in stress perception and coping during adolescence: The role of situational and personal factors. *Child Development, 80*(1), 259–279.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46,* 407–441.

Shelley, M., & Pakenham, K. I. (2004). External health locus of control and general self-efficacy: Moderators of emotional distress among university students. *Australian Journal of Psychology, 56*(3), 191–199.

Skaalvik, E. M., & Federici, R. A. (2015). Prestasjonspresset i skolen. *Bedre skole, 3,* 11–15.

Skaalvik, E. M., & Skaalvik, S. (2013). School goal structure: Associations with students` perceptions of their teachers as emotionally supportive, academic self-concept, intrinsic motivation, effort, and help seeking behavior. *International Journal of Educational Research, 61,* 5–14.

Sonmark, K., Godeau, E., Augustine, L., Bygren, M., & Modin, B. (2016). Individual and Contextual Expressions of School Demands and their Relation to Psychosomatic Health: A comparative Study of Students in France and Sweden. *Child Indicators Research, 9*(1), 93–109.

Sotardi, V. A. (2016). Understanding Student Stress and Coping in Elementary School: A Mixed-Method. *Longitudinal Study. Psychology in the Schools, 53*(7), 705–721.

Sowislo, J. F., & Orth, U. (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychological Bulletin, 139*(1), 213–240.

Stockings, E. A., Degenhardt, L., Dobbins, T., Lee, Y. Y., Erskine, H. E., Whiteford, H. A., & Patton, G. (2016). Preventing depression and anxiety in young people: a review of the joint efficacy of universal, selecktive and indicative prevention. *Psychological Medicine, 46,* 11–26.

Suldo, S. M., & Shaunessy-Dedrick, E. (2013). The Psychosocial Functioning of High School Students in Academically Rigorous Programs. *Psychology in the Schools, 50*(8), 823–843.

Sun, J., Dunne, M. P., Hou, X. Y., & Xu, A. Q. (2013). Educational stress among Chinese adolescents: individual, familiy, school and peer influences. *Educational Review, 65*(3), 284–302.

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting Positive Youth Development Through School-Based Social and Emotional Learning Interventions: A Meta-Analysis of Follow-Up Effects. *Child Development, 88*(4), 1156–1171.

Theron, L. C. (2016). The everyday ways that school ecologies facilitate resilience: Implications for school psychologists. *School Psychology International, 37*(2), 87–103.
Thoits, P. A. (2010). Stress and health: Major findings and policy implications. *Journal of Health and Social Behavior, 51*(1), 41–51.
Thurstone, L. L. (1947). *Multiple factor analysis*. University of Chicago Press.
Ungar, M. (2008). Resilience across cultures. *British Journal of Social Work, 38*(2), 218–235.
Van Keyserlingk, L., Becker, M., & Jansen, M. (2019). Academic self-concept during the transition to upper secondary shool. *Contemporary Educational Psychology, 56,* 152–160.
Vugteveen, J., de Bildt, A., Theunissen, M., Reijneveld, M., & Timmerman, M. (2019). Validity Aspects of the Strengths and Difficulties Questionnaire (SDQ). Adolescent Self-Report and Parent-Report Versions Among Dutch Adolescents. *Assessment, 2019,* 1–16.
Wahab, S., Rahman, F. N. A., Hasan, W. M. H. W., Zamani, I. Z., Arbaiei, N. C., Khor, S. L., & Nawi, A. M. (2013). Stressors in secondary boarding school students: Association with stress, anxiety and depressive symptoms. *Asia-Pacific Psychiatry, 5*(S1), 82–89.
Wang, J., Hu, S., & Wang, L. (2018). Multilevel analysis of personality, family, and classroom influences on emotional and behavioral problems among Chinese adolescent students. *PLoS one, 13*(8), 1–16.
Weare, K., & Nind, M. (2011). Mental health promotion and problem prevention in schools: what does the evidence say? *Health Promotion International, 26*(1), 29–69.
Weigold, L. K., & Robitschek, C. (2011). Agentic personality characteristics and coping: Their relation to trait anxiety in college students. *American Journal of Orthopsychiatry, 81*(2), 255–264.
Werner-Seidler, A., Perry, Y., Calear, A. L., Newby, J. M., & Christensen, H. (2017). School-based depression and anxiety programs for young people: A systematic review and meta-analysis. *Clinical Psychology Review, 51,* 30–47.
Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., Charlson, F. J., Norman, R. E., Flaxman, A. D., Johns, N., Burstein, R., Murray, C. J. L., & Vos, T. (2013). Global burden of disease attibutable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet, 382,* 1575–1586.
Wigfield, A., & Karpathian, M. (1991). Who am I and what can I do? Children`s self-concepts and motivation in achievement situations. *Educational Psychologist, 26,* 233–261.
World Health Organization. (2018). *International Statistical Classification of Diseases and Related Health Problems(11th Revision)*. Switzerland: World Health Organization.
Yang, H., & Chen, J. (2016). Learning Perfectionism and Learning Burnout in a Primary School Student Sample: A Test of a Learning-Stress Mediation Model. *Journal of Child and Family Studies, 25*(1), 345–353.
Yeoh, S. H., Tam, C. L., Wong, C. P., & Bonn, G. (2017). Examining depressive symptoms and their predictors in Malaysia: Stress, locus of control, and occupation. *Frontiers in Psychology, 8*(8), 1–10.
Young, J. F., & Mroczek, D. K. (2003). Predicting intraindividual self-concept trajectories during adolescence. *Journal of Adolescence, 26,* 586–600.

Jan Arvid Haugan is an Associate Professor at the Norwegian University of Science andTechnology. His research interests are the development of social relationships, stress, copingand well-being within and across different microsystems in adolescents life-ecology,especially the microsystems related to sports and education.

Per Frostad is a professor at the Norwegian University of Science and Technology. Hisresearch interests are social inclusion, mental health and wellbeing in upper secondaryschools.

Per Egil Mjaavatn is an Associate Professor at the Norwegian University of Science and Technology. His research interests are everyday life and wellbeing among children and youth, students' opinions about their physical and mental health, social relations and their experiences in the education system.

The *Lancet Psychiatry* Commission on psychological treatments research in tomorrow's science

Emily A Holmes, Ata Ghaderi, Catherine J Harmer, Paul G Ramchandani, Pim Cuijpers, Anthony P Morrison, Jonathan P Roiser, Claudi L H Bockting, Rory C O'Connor, Roz Shafran, Michelle L Moulds, Michelle G Craske

Lancet Psychiatry 2018
5: 237–86

Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden (Prof E A Holmes PhD, Prof A Ghaderi PhD); **Department of Psychiatry, University of Oxford, Oxford, UK** (Prof C J Harmer DPhil); **Oxford Health NHS Trust Foundation, Warneford Hospital, Oxford, UK** (Prof C J Harmer); **Centre for Mental Health, Imperial College, London, UK** (Prof P G Ramchandani DPhil); **Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, Netherlands** (Prof P Cuijpers PhD); **Psychosis Research Unit, Greater Manchester Mental Heath Trust, Manchester, UK** (Prof A P Morrison ClinPsyD); **School of Psychological Sciences, University of Manchester, Manchester, UK** (Prof A P Morrison); **Institute of Cognitive Neuroscience, University College London, London, UK** (Prof J P Roiser PhD); **Academic Medical Center, Department of Psychiatry, University of Amsterdam, Amsterdam, Netherlands** (Prof C L H Bockting PhD); **Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK** (Prof R C O'Connor PhD); **University College London Great Ormond Street Institute of Child Health, London, UK** (Prof R Shafran PhD); **School of Psychology, The University of New South Wales, UNSW, Sydney, NSW, Australia** (Prof M L Moulds PhD); **and Department of Psychology and Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, CA, USA** (Prof M G Craske PhD)

Executive summary

Background

Psychological treatments occupy an important place in evidence-based mental health treatments. Now is an exciting time to fuel treatment research: a pressing demand for improvements is poised alongside new opportunities from closer links with sister scientific and clinical disciplines. The need to improve mental health treatment is great; even the best treatments do not work for everyone, treatments have not been developed for many mental disorders, and the implementation of treatments needs to address worldwide scalability. Psychological treatments have yet to benefit from numerous innovations that have occurred in science, particularly those that have emerged in the past 20 years, and arguably vice versa. This Commission comprises ten parts that each outline an area in which we see substantial opportunity and scope for advancements that will move psychological treatments research forward.

Part 1: How do existing treatments work? Making the case for the mechanisms of psychological treatments

Beyond knowing that an intervention is efficacious, research initiatives are needed that clarify the key mechanisms through which interventions work. An experimental psychopathological approach enables the identification of mechanisms. Research on these mechanisms has considerable scope to facilitate treatment innovation.

Part 2: Where can psychological treatments be deployed? Research to improve mental health worldwide

We outline a number of factors to facilitate worldwide access to psychological treatments. Future research initiatives need to continue to develop and assess the efficacy of brief and flexible interventions that can be adapted to meet the needs of individuals across cultural contexts, and delivered and disseminated in a sustainable way.

Part 3: With what? The potential for synergistic treatment effects—using and developing cross-modal treatment approaches

The combination of psychological and pharmacological treatments needs to be better understood, both in terms of the clinical effect and the underlying shared and different mechanisms. Efforts to develop and investigate the efficacy of novel cross-modal treatments could contribute to treatment innovation.

Part 4: When in life? Psychological science, prevention, and early intervention—getting the approach right from the start

The social and economic tolls of mental health problems early in life make the development of effective prevention and early intervention approaches a priority. A preventive focus and a developmental approach are needed to identify risk factors for psychopathology, and identification of the optimal time at which to offer prevention approaches is needed to increase the likelihood of vulnerable young people growing up with positive mental health.

Part 5: Technology—can we transform the availability and efficacy of psychological treatment through new technologies?

New technologies provide exciting and timely means by which to disseminate and extend the efficacy and global reach of evidence-based interventions. eHealth and mHealth approaches that use information technology (eg, the internet, virtual reality, serious gaming) and mobile and wireless applications (eg, text messaging, apps) are examples of how technology has been harnessed to innovate psychological treatments and their availability and evaluation.

Part 6: Trials to assess psychological treatments

The findings of randomised controlled trials that assess psychological therapies inform policy and practice. Accordingly, the design and conduct of these trials warrants scrutiny and ongoing efforts for quality improvement (eg, reporting standards, specification of protocols, inclusion and exclusion criteria, choice of outcome measures, measurement of adverse effects, and prevention of bias in design and analysis). We outline several opportunities for further improvement that should enhance the credibility and quality of future trials.

Part 7: Training—can we cultivate a vision for interdisciplinary training across mental health sciences to improve psychological treatments?

Early examples of collaboration between basic scientists and clinicians translated into historical steps in the innovation of psychological treatment. Such synergy has become less apparent in the past few years. The improvement in links between clinical psychology, psychiatry, and basic research has the potential to deliver more advances in psychological treatments. We propose opportunities to improve training in interdisciplinary mental health sciences. This training approach would be the first step toward forging links between scientists and

Correspondence to:
Prof Emily A Holmes, Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, SE-171 77 Stockholm, Sweden
emily.holmes@ki.se

clinicians in the next generation and bridging the gap between clinical practice and the basic research programmes that underpin psychological treatments.

Part 8: Whom should we treat, for what, and with what? Embracing the complexity of mental disorders from personalised models to universal approaches

Mental disorders are inherently complex (eg, heterogeneity in symptoms across disorders, high rates of comorbidity) and evidence-based treatments must address this complexity. Potential solutions include considering both highly individualised (ie, personalised) approaches and so-called universal or transdiagnostic approaches that target common mechanisms. A goal of future research will be to examine whether these approaches improve treatment effectiveness.

Part 9: Target: suicidal behaviour—protecting lives

Suicidal behaviour is one of many areas in which advances are needed. Despite developments in the understanding of risk factors that predict the likelihood of suicide attempts, and the treatment and prevention of suicidal behaviour, many questions remain. We specify areas for future research—eg, use of new technologies, the role of culture, input from individuals with lived experience of suicidal behaviour, and using a team-based approach in the development, assessment, and dissemination of prevention efforts.

Part 10: Active innovation and scrutiny of future psychological treatments research

The task of improving psychological treatments is an exciting prospect for scientists and clinicians with an interest in the so-called science of mental life. Clinicians, researchers, service users, carers, funders, commissioners, managers, policy planners, and change experts all have a part to play in improving psychological treatment. Some long-held ideas need examination, from the branding of psychological treatment types, to considering what people actually want treatment for. Scrutiny of new ideas should be rigorous and yet encourage innovation.

Introduction

Psychology and psychological treatments

Psychology from its inception was defined as "the science of mental life".[1] Psychological treatments have developed to occupy a key place in evidence-based treatments for mental health. Many pivotal techniques used in evidence-based psychological treatments arose from psychological research on processes in the 1950s and 1960s, with basic and clinical researchers often in the same department. During the past few decades, the psychological treatment field has drifted away from its scientific roots, while mechanistic studies have drifted away from treatment issues. Now is the time for greater synergy between basic and clinical researchers to invigorate psychological treatment research.[2] Psychological treatments offer great promise for continued innovation, not least because of the development of scientific methods and perspectives from many allied fields.

While researchers and industry struggle to produce new drugs for mental disorders, psychological treatments research might have the potential to deliver acceptable, effective, and safe treatment options more quickly.[3] Building bridges between psychological treatment and other modalities—eg, via combination approaches—could also benefit many service users, but will not be an easy task. New trials of psychological treatments are met with not only enthusiasm, but also controversy. Questions are constantly being asked about trial design, implementation, and interpretation. Do trial populations reflect real clinical populations? What is an appropriate control group? At what point should trial evidence be translated into day-to-day practice? How can an inter-vention be disseminated nationally and internationally? Existing assumptions are also being queried, for instance, is single-session therapy feasible? Is one consistent therapist an optimal or even necessary component of psychological treatment? How can new technologies best be harnessed?

We note that in the wider literature many terms are used, including mental health disorder, psychological disorder, psychiatric disorder, mental health problem, and other forms of terminology associated with psychological treatments, such as mental health difficulties and behavioural difficulties. In line with *Lancet Psychiatry* terminology and for consistency, the term mental disorder is used in this Commission.

A core role for psychological treatments in the future requires a research agenda

The burden of mental disorders is enormous, and yet pharmacological and psychological treatments scarcely reduce the disease burden. Since most patients prefer psychological treatments over pharmacological treatments,[4] increased research efforts are required to develop psychological treatments to a point at which they will have a substantial effect upon the mental disease burden worldwide. To realise the development of psychological treatments, a research agenda is needed that can guide the field for the coming years. For example, a 2014 commentary[2] on improving psychology treatments stated: "By the end of 2015, representatives of the leading clinical and neuroscience bodies should meet to hammer out the ten most pressing research questions for psychological treatments. This list should be disseminated to granting agencies, scientists, clinicians, and the public internationally. Mental-health charities can help by urging national funding bodies to reconsider the proportion of investments in mental health relative to other diseases."[2]

Mental disorders are widespread and costly

Every year almost one in five people worldwide develops a mental disorder,[5] and more than 750 000 people die by suicide.[6] In 2010, mental and substance-use disorders

accounted for 183·9 million disability-adjusted life-years,[7] with most of the disease burden caused by depressive disorders, anxiety disorders, and substance-use disorders. These numbers are likely to be an underestimation since these calculations assume that mental disorders are not associated with excess mortality, except suicide. However, people with mental disorders have a considerably higher risk of dying earlier than those without mental disorders.[8]

Apart from the personal suffering of affected patients and their families, mental disorders pose enormous economic challenges to communities and societies, in terms of production losses and health and social care expenditures.[9–11] The global cost of mental health conditions in 2010 has been estimated at US$2·5 trillion, and these costs are expected to grow to $6·0 trillion by 2030.[12] For this reason, conceptualisations of mental health need to expand beyond the notions of disease or infirmity to functionally related outcomes or, more broadly speaking, the ability to adapt and self-manage.[13]

Treatments make a small contribution to the reduction of the disease burden

Several evidence-based biological and psychological treatments are available for a range of mental disorders. However, these treatments are estimated to be able to reduce the disease burden by only about 40%, and only under optimal conditions and when all patients with a mental illness receive evidence-based treatment.[14] Globally, coverage (ie, the proportion of people who receive a consultation for a mental disorder) is typically much lower than 100%, with coverage well below 50% for some disorders (eg, eating disorders) in most regions,[15] and for some disorders (eg, alcohol-related disorders) coverage is below 10%.[16] The 2014 Adult Psychiatric Morbidity Survey,[17] noted a welcome increase in the number of people with common mental disorders who are receiving treatment. This increase has been largely attributed to the use of psychotropic medication.[17] Unfortunately, most patients who are treated for mental disorders do not receive evidence-based treatments, but instead receive a wide array of treatments including interventions that are not evidence based.[18]

Patient preference for psychological treatment options alongside restricted availability

In the USA, psychotherapy has assumed a less prominent role in mental health care than have treatments with medication.[19] For example, in the USA, antidepressant use almost doubled between 1996 and 2005, from 13 to 27 million individuals, whereas the percentage of people among antidepressant users who underwent psychotherapy declined from 31·50% to 19·87%.[20] In an office-based clinical practice, between 1999 and 2010, on average 8·6% of visits made by adults with depression included the prescription of a second-generation antipsychotic drug,[21] and the frequency of use doubled from 4·6% in 1999–2000 to 12·5% in 2009–10. By contrast, most patients seem to prefer psychotherapy over medication. A meta-analysis[4] of patients with a range of mental disorders (eg, depression, anxiety, insomnia, bipolar disorder, schizophrenia, substance-related disorders, eating disorders, and personality disorders) estimated that approximately 75% of patients prefer psychotherapy as their treatment as opposed to medication. However, some patients do prefer pharmacological treatment, whereas others might have no preference. In this Commission we do not seek to reinforce what we believe to be a misplaced dichotomy between biological and psychological approaches (see Part 3), instead we seek a research agenda that is open to multiple perspectives, does not neglect one perspective at the expense of another, considers links between both perspectives, is informed by patient preferences, and ultimately leads to the greatest clinical effect.

Although most patients prefer psychotherapy to medications,[4] the availability of such treatment is a major problem in many countries.[22] This paucity of availability is attributable to a range of factors, including financial constraints or the scarcity of trained psychotherapists who can deliver the evidence-based treatments. Therefore, psychotherapy is mostly delivered in high-income countries to those who can afford it and know how to find a therapist. In low-income and middle-income countries, psychological treatments are scarce—although notable exceptions exist (see Part 2).[23]

Several approaches are being developed to increase access to psychological services, such as the Increasing Access to Psychological Treatment (IAPT) programme in the UK, in which low-intensity psychotherapies are made available on a large scale and high-intensity therapies are available for those who do not respond to low-intensity therapies.[24] Internet-based interventions (see Part 5) can help in making psychotherapies available to those who need them since these interventions can be offered relatively inexpensively and with a low threshold for access. Another important development to make therapies more accessible is to use so-called lay health counsellors (see Part 2).

Psychological treatment research in tomorrow's science

Improved psychological treatments are needed to help reduce the burden of mental disease worldwide. The landscape of psychological treatment research is ready for innovation, offering exciting and auspicious opportunities for research in the mental health sciences. Insights from different fields of science might allow us to "stand on the shoulders"[25] of existing evidence-based psychological treatments and see further to improve psychological interventions. Greater collaborative endeavours between clinical and basic researchers of many disciplines will help in this regard.[2]

In this Commission, we discuss opportunities to focus future research efforts to improve worldwide mental health. Suitable areas of inquiry for future research

include: understanding the mechanisms that underlie psychological treatments, increasing worldwide access to treatments, developing cross-modal treatment approaches, and enhancing preventive and developmental approaches. To address each of these themes, new tools will be needed, which will be provided by new technologies, improved trial methodologies, and improved training in interdisciplinary mental health sciences, to name but a few sources. In this Commission, we discuss how the goals of people developing and delivering psychological treatments should be to embrace challenging areas, such as the inherent complexities of mental disorders and issues such as suicide prevention. The array of challenges ahead to which a psychological perspective can contribute will require fresh innovation.

Research into these areas will require ideas to be tested, and rejected or developed in line with scientific methods and challenges of mental health of the time (as opposed to therapeutic habit and allegiance to a specific manner of clinical training, or science focused inwardly on itself rather than on genuine application); therefore, attitudes within mental health science will need to change. To illustrate, we make an analogy with a British contemporary art initiative that is engaged with Trafalgar Square's empty plinth in London, UK. Statues are on three of the four plinths in the corners of Trafalgar Square and the fourth plinth stood empty for over a century (figure 1). Now, the so-called Fourth Plinth Programme[26] invites world-class artists to make "astonishing" new artworks for the centre of the capital city. Commissions create a rolling programme of temporary artworks rather than settling permanently on one figure or idea. The resultant sculptures tend to be shown for a year, although sometimes only for a few months, and sometimes the plinth is empty for a period of time; however, the momentum of the programme and scrutiny over the choice of statues continues. Some artworks stand the test of time, whereas some might not. Associated initiatives encourage projects and creative thinking around past and present artworks displayed on the fourth plinth. However, the best use of the fourth plinth remains a subject of debate and discussion in the public, media, and art world.

Figure 1: **The fourth plinth in Trafalgar Square, London, UK**

Like the Fourth Plinth Programme, psychological treatments research needs innovation, rotation of ideas, and robust critical debate as a clear part of advancing research. Although the objects of inquiry might change, the principles of seeking to improve research efforts towards improved mental health will persist. Instead of being prescriptive regarding the future of psychological treatments research, this Commission sets out various suggestions and principles to guide research that should apply across different mental disorders and transdiagnostic processes, approaches, countries, and, indeed, to the new and future generations of mental health researchers. These principles might change over time and how best to strengthen psychological treatments should be a subject of research, debate, and discussion, involving the fields of both psychological treatments and mental health science, and many fields beyond these.

When considering the traditional delivery method of psychological treatments, the changes that can come about from two people talking with each other for a matter of hours during therapy sessions are fascinating, sometimes remediating years of mental distress. Although clearly the presence of another person can be helpful, evidence-based psychological treatments involve far more than just skills that boost therapeutic alliances. Therapeutic effects are now known to be achievable without a therapist being physically present (eg, via internet therapy, see Part 5) and some psychological techniques can be effective when delivered by lay workers with modest training (see Part 2). Moreover, neuroscience continues to reveal how efficiently the mind can work under various parameters (eg, in modulating memory) by a range of techniques that may or may not require another person to be present. The emotional, behavioural, and social changes rendered through therapy pose mechanistic questions for mental health science—eg, how do effective psychological treatments work? The identification of specific targets for mechanistic questions might be facilitated not only by quantitative methods but also by qualitative methods—eg, detailed narratives of individuals' experiences as they undergo psychological treatments. Once potential targets have been identified in this way, they could be subjected to experimental investigation to establish causality for therapeutic change.

We now focus and elaborate on the ten key themes that we see as instrumental to consider in the development of an agenda to progress mental health treatment research. These themes, which were decided as part of a consultation meeting in December, 2015 (panel 1), are not exhaustive and many more are to be welcomed for future scrutiny.

Part 1: How do existing treatments work? Making the case for mechanisms of psychological treatments

Introduction

Although some psychological treatments are effective, little is known about the processes through which therapeutic change occurs. As Alan Kazdin stated in his 2007 review,[27] many evidence-based therapies are available but little understanding exists of the mechansims of change or precisely how they work. Understanding mechanisms of action is essential to derive and hone treatment strategies to directly target the mechanisms, remove irrelevant strategies, and develop novel approaches that are more expeditious and effective than current treatments. Knowledge of mechanisms also allows improved precision in matching psychological treatments to the needs of individuals to improve outcomes compared with current methods.

Research into the mechanisms of treatments offers an exciting opportunity for psychological treatment research. However, most studies in psychopathology have simply described differences between groups of individuals with and without a diagnosis and identified a mechanism of action by use of these differences—an approach that cannot identify causal mechanisms. To move the field toward understanding causality, research on mechanisms should be optimised by framing research within the context of clinical treatment to understand how existing treatments work, and derive new and improved treatments.

What is a mechanism of psychological treatment?

Mechanisms of psychological treatment are defined as "the steps or processes through which therapy (or some independent variable) actually unfolds and produces the change. Mechanisms explain how the intervention translates into events that lead to the outcome."[27] A mechanism is an explanatory construct and not simply an intervening variable that explains the statistical association between an intervention and an outcome. For example, the finding that changes in a patient's perceived self-efficacy and outcome expectancy statistically mediates the subsequent changes in anxiety and functioning[28] does not explain how the changes in self-efficacy and outcome expectancy lead to those outcomes. The underlying changes responsible for symptom improvement could involve multiple processes, including, but not limited to, neural systems, other physiological systems, cognitions, emotions, and behaviours.

The processes through which psychological treatments produce change often overlap with or complement mechanisms that are responsible for the onset or, in particular, the maintenance of psychopathology (hereafter referred to as mechanisms of psychopathology). The US National Institute of Mental Health (NIMH) Research Domain Criteria (RDoC) initiative is directing the search for mechanisms of psychopathology away from the constraints of categorical diagnostic criteria and towards dimensions of observable behaviour and neurobiological measures.[29] The RDoC initiative aims to "elaborate a set of psychological constructs linked to behavioral dimensions for which strong evidence exists for circuits to implement these functions, and relate the extremes of functioning along these dimensions to specified symptoms (i.e., impairment)."[30] Essentially, the RDoC framework aims to identify biopsychological explanations or so-called process constructs for clinical events; these same process constructs could explain change in clinical events throughout treatment. The provisional list of RDoC explanatory constructs includes negative valence systems, positive valence systems, cognitive systems, systems for social processes, and arousal or modulatory systems, with each construct comprising more specific subconstructs.[30] The constructs are assessed with measures that represent at least seven levels called units of analysis: genes, molecules, cells, neural circuits, physiology, behaviour, and self-report. Identifying a mechanism using one unit of analysis does not exclude mechanisms identified using other units of analysis.

Mechanisms of psychopathology vary from being predominantly distal (eg, effects of early life adversity upon inflammatory markers for depression that might not become apparent until many years later[31]) to predominantly proximal (eg, ongoing biases in autobiographical memory for depression;[32] see Roiser's 2015 article[33] for a discussion of these ideas). Mechanisms of psychopathology also vary from being predominantly fixed (eg, within genes, albeit with variations in expression) to predominantly malleable (eg, negative interpretation bias for ambiguous stimuli). Psychological treatments generally target the predominantly proximal and malleable mechanisms of psychopathology—eg, attention bias

Panel 1: **Methodology and approach used in preparing this Commission**

- This Commission arose from an initial consultation meeting in December, 2015, in which researchers from a variety of backgrounds with interests or expertise in psychological treatments research met to discuss challenges in the field, and to lay out possibilities for a future research agenda for advancing the science of psychological treatments
- The group's common interest was captured by Kazdin's call to arms to "reboot psychotherapy research and practice to reduce the burden of mental illness"[22]
- Attendees' backgrounds in terms of subject disciplines included clinical psychology, psychiatry, neuroscience, experimental psychology, and pharmacology
- The language of the meeting was English, and attendees were from the UK, Europe, and the USA; in this Commission we have only cited papers that have been published in English
- The Commission expresses the authors' collective views about some of the key areas in which we see scope for improvements in the field; our goal was not to provide an exhaustive literature review, or a systematic review of specific topics; rather, we have cited sources that are relevant to the issues that we have discussed in the context of each of the ten themes
- We note that many important topic areas and perspectives continue to develop, and that this Commission is a start for necessary and continued discussion

modification training for anxious individuals who have selective attention bias towards threat-relevant stimuli.[34] Alternatively, psychological treatments can target factors that differ from mechanisms of psychopathology but compensate for them—eg, compensatory cognitive training for psychosis.[35] Although less commonly targeted, distal mechanisms can be particularly good targets for prevention efforts. Notably, not all treatment mechanisms are directly tied to mechanisms that are responsible for the onset or maintenance of psychopathology; some treatments work through independent processes—eg, applied behavioural analysis techniques for individuals with autism.[36]

What is the state of the field?

Pivotal evidence-based psychological treatments have evolved by specifically targeting identified mechanisms of psychopathology, one example of which is the treatment of panic disorder. Through a series of experimental investigations and animal modelling, interoceptive conditioning (ie, acquired fear of visceral or other internally generated stimuli due to pairing with an aversive outcome, such as pairing an elevated heart rate with the possibility of a heart attack) and catastrophic misappraisal (ie, misinterpretations of interoceptive stimuli as harmful or threatening) were recognised as mechanisms underlying the fear of bodily sensations that characterises panic disorder.[37–39] Psychological treatments for panic disorder were developed to target specific mechanisms in the form of interoceptive exposure[40] (ie, repeated exposure to interoceptive stimuli in the absence of aversive outcomes) and cognitive restructuring[41] (ie, reasoning skills to replace catastrophic interpretations with evidence-based interpretations). This type of treatment has been shown to be particularly effective for panic disorder, and more effective than non-targeted supportive psychotherapy (Hedges' g 0·35, 95% CI 0·04–0·65).[42] Similarly, the conceptualisation of instrumental reinforcement of compulsions led to a treatment for obsessive compulsive disorder known as exposure and response prevention.[43] In this conceptualisation, the distress-reducing effects of compulsive washing in response to obsessive thoughts of being contaminated act to reinforce and therefore increase compulsive washing with each subsequent obsessive thought. The treatment combines exposure of the individual to reminders of the obsessive thoughts (eg, a dirty piece of clothing) or the thought itself (eg, the thought of being covered in germs) with the prevention of washing. This approach is very effective for patients with obsessive compulsive disorder, and more so than non-targeted psychological control conditions, such as relaxation training (1·29, 0·76–1·81).[44] Another example is behavioural activation therapy, which targets deficits in positive reinforcement as a contributing factor for depression.[45] This approach aims to increase access to positively rewarding stimuli and achieve actions that are value driven and overcome task-related avoidance.[46] In a meta-analysis of behavioural activation treatments for depression,[47] this form of treatment was found to be highly effective compared with comparison-control interventions, which included wait-list and non-targeted psychological control conditions (0·87, 0·60–1·15 when collapsed across control conditions).

Overall, the development of psychological treatments via a mechanistic approach has resulted in more precise, efficient, and effective treatments than those that do not target specific mechanisms. However, the largest effect sizes come from comparisons with non-treatment or wait-list control conditions, with the wait-list control conditions potentially inflating the effect sizes;[48] some of the findings of meta-analyses mentioned earlier included wait-list control conditions.[47] The observation that comparisons of mechanistic treatment approaches with usual care typically yield lower effect sizes than comparisons with non-treatment or wait-list controls[49] could be an indication of the importance of common factors that are relevant to all psychotherapies—eg, goal consensus, therapeutic alliance, empathy, expectations, and therapist effects.[50] Notably, common factors do not obviate the importance of mechanistic research but rather imply the value of taking common factors into account when assessing the mechanisms of specifically targeted therapeutic approaches.

However, despite purported treatment mechanisms, little evidence exists on the precise mechanisms through which psychological treatments actually work. Although mechanistic developments in neuroscience have sparked interest in the psychopathology community, most studies to date have not investigated mechanisms of treatment. Even the study of mediation is often hindered by insufficiently rigorous methodology.[27] For example, although good evidence supports the efficacy of interoceptive exposure and cognitive restructuring for panic disorder, and that the extinction of the fear of interoceptive cues and reduction in catastrophic appraisals occur as a result of treatment, little direct evidence exists that the treatments work through the extinction of the conditional fear of interoceptive cues or reduction of catastrophic appraisals—a claim that requires that the changes in the purported mechanisms explain the subsequent changes in the symptoms. Similarly, although behavioural activation for depression might lead to changes in reward processing, no evidence is apparent that the treatment works through changing neural and behavioural sensitivity to reward.

To make matters worse, the focus of psychological research has slowly shifted away from mechanistically informed approaches toward modifying or adapting existing manualised psychological treatments, sometimes superficially, for different populations and settings. This approach of modification most commonly applies to cognitive and behavioural therapies. Although this shift in focus has been valuable for the advancement of treatment implementation in different settings, it has resulted in a

regrettable divorce from the foundations of mechanistically informed psychological treatments that in turn has impeded the investigation of their mechanisms of action.

Why is it important to understand the mechanisms of psychological treatments?

Without an understanding of the mechanisms of psychological treatments, pathways to intervention development and refinement remain restricted. With a knowledge of how change occurs as a result of treatment, therapeutic strategies can be developed that are more direct, precise, and effective.[51] Also, those therapeutic strategies that do not affect the crucial processes can be removed, making treatments more efficient and effective.[51] Moreover, by refuting a claimed mechanism, research attention can be redirected toward investigating alternative mechanisms and the development of novel treatments that are effective and efficient (panel 2).

Understanding the mechanisms of psychological treatment might uncover moderators of treatment outcome, and thereby lead to improvements in the precision of matching treatments to the needs of individuals.[51] For example, initial interest in training for attention bias modification for anxious individuals waned as a result of mixed findings and small effect sizes.[52] Subsequent research has provided some indication that the effects of training attention bias are larger for individuals with a greater attention bias at baseline,[34] and for those with low-expressing forms of the serotonin-transporter-linked polymorphic region (5-HTTLPR) of the serotonin-transporter gene (*SLC6A4*),[53] than for those with high-expressing forms of 5-HTTLPR. As another example, extinction-based exposure therapy to trauma cues for individuals with post-traumatic stress disorder (PTSD) have been suggested to function in part by enhancing prefrontal cortex inhibitory regulation over the responses of the amygdala.[54] Neuroscientists have identified that some individuals with PTSD fit into subtypes, with the majority showing hyperactivation of the amygdala and hypoactivation of the prefrontal cortex when exposed to trauma reminders, and about 30% showing the reverse pattern of hypoactivation of the amygdala and hyperactivation of the prefrontal cortex.[55] If exposure therapy can be established to work at least partially through enhancing the prefrontal cortex regulation of the amygdala, then exposure therapy might be more effective for the former set of individuals with PTSD than the latter. These examples show ways in which the field of psychological treatment could progress. Conclusive findings will depend upon replication of these results within substantially larger samples.

Not only is the identification of such mechanistic moderators valuable for precision in matching treatment to individuals, but it also improves the elucidation of psychological treatment mechanisms.[51] To follow the previous example of individuals with PTSD, by studying the entire sample (ie, those showing amygdala hyperactivation and those showing amygdala hypoactivation) the extent to which change in amygdala activation serves as a treatment mechanism is likely to be nullified. By recognising baseline differences between individuals, differential mediational pathways could be uncovered—eg, the possibility of amygdala deactivation for those who initially present with hyperactivation, and activation for those who initially present with amygdala hypoactivation. These are illustrative examples, but a mechanistic approach to moderation avoids the default approach of trial and error that assumes that a given psychological intervention strategy works through the same mechanisms for everyone. Another example of a speculative mechanistic hypothesis is the theory that behavioural activation for depression,[46] which involves scheduling activities that are rewarding, leads to symptom improvement for some individuals through enhancing approach motivation or initial responsiveness to reward within positive-valence systems, whereas for other individuals the treatment might reduce threat or potential threat within negative-valence systems or even modulate arousal systems through regulating sleep–wake cycles.

Additionally, psychological treatments with a mechanistic focus can be turned into training in everyday habits that pertain to prevention of and recovery from mental ill health—eg, training in mindfulness techniques to reduce affective memory bias and development of, or relapse into, depressive ruminative states.[56] Another example is the delivery of cognitive behavioural therapy (CBT) as an adjunct to usual primary care for individuals who are depressed and have not responded well to medication alone. In one study,[57] short-term focused CBT was associated with significantly reduced depression 3–5 years after treatment compared with usual care alone. Similarly, another study[58] found that cognitive therapy decreased the recurrence of depression over a 10-year interval in patients with remitted depression who had a history of recurrent depression compared with usual treatment. Together, these data suggest that CBT and cognitive therapy provided patients with skills that were embedded into their daily lives and led to sustained long-term improvements.

Panel 2: **Reasons for understanding the mechanisms of psychological treatments**

- To hone treatments to target the processes that are responsible for change more directly and efficiently
- To uncover essential moderators of treatment outcomes and improve precision in treatment matching
- To develop training programmes for the prevention of and recovery from psychopathology
- To eliminate wasteful and inefficient treatments
- To provide evidence for specificity in treatment beyond non-specific factors that are responsible for the so-called dodo-bird effect

Not understanding the mechanisms of psychological treatments could be detrimental—eg, the development of novel and effective treatments could be hindered by the continued modification of the procedural elements of existing treatments without fully understanding the processes that lead to change. We encourage the development of a larger evidence base of critical processes for therapeutic change and, specifically, of which psychological treatments—existing and newly developed—affect which processes. This evidence base should include common and specific factors of psychotherapies.[30] Additionally, knowing the psycho-logical treatments that exert their effects primarily through common non-specific factors rather than through more targeted specific factors, would be informative, as well as whether the common and specific factors are of greater relevance to one mental disorder or individual than another. Such an evidence base would offer the potential to move the field beyond the long-standing debate of whether all psychological treatments are equally effective (ie, the dodo-bird hypothesis[59]) and whether differential treatment effects exist.[60] We have the opportunity to assess whether matching treatments that are mechanistically focused to individual patients with underlying dysregulation leads to superior outcomes when compared with targeting non-specific factors that are common across psychological treatments. Of course, applying personalised treatments that are mechanistically focused and understanding the role of common factors are not the only ways in which psychotherapy can improve outcomes; other factors that warrant consideration include the personal resources and social context of those in need, and the service delivery systems by which treatments are delivered.

Experimental psychopathology

Understanding mechanisms of psychopathology involves substantial explanatory specificity, and hence is driven by theory.[27] The mechanisms are elaborated through plausible and coherent reasoning on the basis of integration with broader scientific knowledge, and at the same time the explanation provided must be specific in how change in the mechanism accounts for change in the outcome.[27] Once theoretical mechanisms have been elaborated, investigators in the field of experimental psychopathology then assess the validity of the mechanism's causal influences upon selected outcomes (panel 3).

Showing that experimental manipulation of a proposed mechanism leads to symptom change is a powerful method for validation. Experimental studies of this kind in human participants can identify key processes that maintain or change aspects of psychopathology. These studies can also elucidate which of the processes' underlying psychopathology can (or cannot) be modified, and can therefore identify appropriate treatment targets. Burgeoning interest in the mechanisms that underlie psychopathology has been fuelled by advances in cognitive science and neuroscience.[51] For example, an increased activation in affective brain networks and a decreased activation in cognitive control and social cognitive networks has been seen in the brains of young people when they listen to criticism from their mothers, and this activation has been identified as a potentially key mechanism in emotional development.[61] These findings could inform strategies aimed at increasing effective parenting to reduce the risk of mental health problems in offspring.

The direct application of identified mechanisms of psychopathology to mechanisms of psychological treatment is well represented in fear learning and exposure therapies for anxiety disorders—eg, pharmacological drugs that facilitate the consolidation of fear-extinction learning (eg, d-cycloserine) have been shown to have beneficial effects in the context of exposure therapy;[62] although some mixed effects have been reported, possibly due to mechanistic moderators.[63] Another study[64] has shown that retrieving memories that are already stored induces a process of reconsolidation. Once retrieved, the memory has to be rewritten into a long-term memory, which requires neurochemical processes (de novo protein synthesis) in the brain. These processes give rise to the fascinating possibility of changing memories post factum during the period of reconsolidation on retrieval. One study[65] found that engaging an individual with a highly visually absorbing computer game after a memory-reminder cue interrupted the reconsolidation of intrusive visual memories induced by experimental trauma. Pharmacological drugs (eg, propranolol) and behavioural techniques (eg, extinction) have been shown to interrupt the reconsolidation process in human beings, albeit with mixed results,[66] restricting boundary conditions and conceptual challenges.[67]

Evidence that disturbances in autobiographical memory can be potential mechanisms of depression has led to novel therapeutic strategies for depression, including memory specificity training and positive memory elaboration.[32] Additional mechanistic research is needed, and particularly in young people for whom innovative psychological treatments are needed that can precisely target narrowly specified mechanisms that

***Panel 3*: Recommendations for identifying potential mechanisms of psychological treatments**

- Develop a model of explanatory specificity
- Experimental investigation of an explanatory construct to establish causal validity
- Human studies to show that manipulation of a proposed construct leads to symptom change (experimental psychopathology)
- Animal studies to allow more precision and elucidation of targets that cannot be studied in human beings
- Reverse-translation models by use of clinical research to inform models that will be tested in animals
- The flow of iterative and reciprocal information between experimental psychopathology studies in human beings and animals

are consistent with developmental models of causality (see Part 4).

Purported mechanisms can be tested in animals with much more precision with regards to measurement and causality than is possible in human beings. Animal studies are invaluable for identifying basic processes and mechanisms that are not possible to address in human beings because of practical or ethical constraints. Indeed, the first clinical applications of d-cycloserine for exposure therapy and disruption of reconsolidation for fear memories were derived from careful experimentation in animals.[64,68] Animal studies have also elucidated the potential value of disruption of reconsolidation in the treatment of substance abuse or dependence.[69] Ongoing animal work is examining pharmacological drugs that regulate the stress response via inhibition of the renin–angiotensin system (eg, losartan) as another method for enhancing consolidation of extinction learning.[70] Furthermore, advances in understanding the neurobiology of rodent self-grooming could identify potential treatment mechanisms for repetitive behaviours such as compulsions.[71]

In reverse-translation approaches, clinical research informs models to be tested in animals—eg, paradigms for assessing depressive cognitive styles, such as pessimism, that have been validated in human studies have now been reverse translated into paradigms that measure judgment bias in rodents.[72] Similarly, drawing from human-based studies on reward systems, paradigms have been developed to assess decision making in rodents between cues that predict reward versus cues that predict punishment.[73]

Despite these examples of the iterative flow of reciprocal information between experimental studies in human beings and animals, for the most part a huge gap exists between basic and clinical researchers. This gap hinders the development of more refined animal models of psychopathology and treatment and their translation to clinical populations. The reverse and forward translation of advances in basic science and clinical science is essential.

Assessment of mechanisms

Once a mechanism has been identified through careful experimentation, it can be assessed within the context of adequately powered clinical trials. To reach this stage requires measures of the purported mechanisms that are reliable, valid, and sensitive to change, since these measures will become the mediators that are assessed statistically. A major contribution to this effort will be funding to establish a list of candidate mechanisms that explain therapeutic change (based on evidence that the experimental manipulation influences only selected outcomes in animal or human studies) and a list of measures for each candidate mechanism. The RDoC notion of units of analysis provides a helpful framework for choosing measures from multiple modalities.

Kazdin[27] has carefully outlined the steps necessary to establish that a measure is a mediator of a psychological treatment. As an initial step, a strong association must be shown between the psychological treatment and the hypothesised mediator (ie, the mediator changes over the course of treatment), and between the mediator and therapeutic outcome (ie, change in the mediator is related to clinical outcomes). Kazdin lists several methods that allow greater attribution of causality to the mediator—ie, the underlying mechanism. One method is temporal precedence, since mediation cannot be presumed unless changes in the purported mediators occur before, and then predict changes in, the outcomes. Temporal precedence necessitates repeated measurement of mediators and of outcome variables throughout treatment, ideally in every treatment session.

Causality can be attributed to a mechanism more confidently when a single mechanism is specifically associated with a single outcome. Even more convincing than the identification of a single mediator is when the purported mediator of a specific psychological treatment can predict patient outcomes more accurately than a mediator of a different mechanism that has no theoretical association with the treatment. Specificity can also be shown by a stronger mediation via a proposed mediator for a treatment with which it has a theoretical association, compared with a treatment to which it is not theoretically relevant. Evidence for dose–response effects, in which stronger doses of the proposed mediator are associated with greater changes in symptoms than weaker doses, also strengthens the argument for a causal link. The consistency of the associations observed across independent replications is another validator. Although for some mechanistic questions appropriately powered experimental studies of small samples can be informative, validation of the mechanism will require large samples. Collaborative multisite studies will be needed, which will require a strong investment from funders and collaboration among researchers focusing on common goals.

Finally, the field would be advanced by listing the various therapeutic elements that constitute psychological therapies; an effort that has already been initiated.[74] Psychological treatments are mostly packages of different elements, such as cognitive restructuring, self-monitoring, problem solving, relaxation training, or assertiveness training. The more elements that are combined in a psychological treatment, the harder mechanistic specificity is to establish. Improved precision is likely to come from assessing the mechanisms of particular procedural elements rather than combinations of elements (panel 4).[75] Increased collaboration between clinical researchers and basic scientists, combined with new methods and technologies, will help the field of psychopathology to make substantial progress in understanding the mechanisms of change in evidence-based psychological treatments.

Part 2: Where can psychological treatments be deployed? Research to improve mental health worldwide

Introduction

Little or no access to efficacious psychological treatments is not only a major problem for people in low-income and middle-income countries, but is also problematic for many people in high-income countries. Brief, flexible, modular, and efficacious treatments that are derived from mechanistic research could enable the efficient adaptation of such treatments to different cultural contexts. Furthermore, the adaptation of treatments could be of help in the training of lay people who could implement such interventions within a framework of low-intensity treatment using modern techniques on a large scale in low-income, middle-income, and high-income countries. To achieve this goal further work is needed, including: the development of such treatments and adapting them to the local needs, priorities, traditions, and cultural norms of different environments; education and training for lay people to acquire proficiency to deliver such treatments as counsellors in a sustainable way; and the development of delivery models for mental health care with long-term sustainability.

Psychological treatments from an international perspective

As discussed in Part 1, mental disorders constitute a substantial part of the burden of disease worldwide.[7,76] Mental disorders also interact with other serious health conditions—eg, cardiovascular diseases, ischaemic stroke, and HIV—increasing the risk of premature death.[77] Efficacious psychological treatments for a wide range of mental disorders have mainly been developed in North America or Europe, and are typically designed for delivery through one-to-one psychotherapy by highly trained professionals. However, at a global level, 90% of individuals with mental disorders do not receive treatment.[78] Little success will be achieved in decreasing the prevalence and incidence of mental illness without a major shift and expansion in clinical practice and intervention research.[22]

A scarcity of skilled human resources (ie, therapists) and low acceptability of psychological treatments across cultures have been suggested as the two major barriers to increasing access to evidence-based psychological treatments in low-income and middle-income countries.[79] WHO estimated a shortage of 1·18 million mental-health-care workers across 144 low-income and middle-income countries.[80] Other key barriers include prevailing public health priority agendas and inadequate investment in mental health care, stigma associated with accessing mental health care, and challenges in using primary-care settings for implementation of mental health care.[81]

Panel 4: **Recommendations for the assessment of mechanisms of psychological treatments**

- Assess within the context of adequately powered clinical trials
- Develop measures of mechanisms that are reliable, valid, and sensitive to change, and that represent multiple units of analysis (eg, genes, molecules, cells, circuits, physiology, behaviour, cognition, self-report); mechanisms are explanatory constructs, whereas measures are mediators that explain the statistical association between an intervention and an outcome
- Once a mechanism has been identified through experimental work, it can be assessed within clinical trials (see text)
- Establish mediation by showing change in the mediator over the course of treatment, and that change in the mediator precedes and predicts clinical outcomes
  - Temporal precedence (ie, change in the mediator precedes and predicts subsequent change in symptoms); value of repeated measurement
  - Specificity of mediation to a single or restricted number of mediators
  - Specificity of mediation to a theoretically relevant mediator versus an irrelevant mediator for a given treatment, or specificity of a theoretically relevant mediator versus one treatment relative to another treatment to which it does not have theoretical relevance
  - Dose–response relationship between degree of change in mediator and degree of clinical improvement
- Consistency in independent replication
- Assess mediation for elements or specific therapeutic strategies rather than packages of treatment elements

Research to improve worldwide access to psychological treatments

Global access to psychological treatments could become a reality if adequate global and local political support is given and a research agenda is compiled that includes, but is not limited to, the following conditions (panel 5). Psychological treatments that could be scaled up successfully would be brief, flexible, modular, efficacious, and streamlined to remove any unnecessary complexities. Such treatments should be aided by research into mechanisms of action in psychological treatments (see Part 1), and a consideration of the core psychopathology of mental disorders. Large and complicated psychological treatment packages can be delivered only by highly trained professionals and to the minority of people who can afford the high costs that are associated with such treatments. Simplified and clearly defined treatments could be more readily adapted to local needs and delivered by lay mental-health-care workers on a larger scale, and as low-intensity treatments—eg, via the internet. Mechanistically informed treatments could also afford flexibility—eg, shaping treatment to align with local cultural norms and conditions. For example, if one of the major maintaining factors in depression concerns a paucity of positive reinforcement in daily life (see Part 1) then treatment strategies to increase positive reinforcement can be formed in many different ways depending on what is the most relevant, acceptable, and affordable option in the specific context or culture—eg, via various cognitive, behavioural, or psychosocial

techniques. Such treatments could each have flexible forms, but be identical in function.

In low-income and middle-income countries, the development of psychological treatments has typically focused on improving availability and accessibility, and researchers have taken a pragmatic approach to treatment development itself; however, future research efforts should harness scientifically driven developments. Developing psychological treatments on the basis of sound psychological theories and empirical knowledge gained from research on the processes of action in treatment could afford opportunities for cultural adaptation and psychological treatment across international contexts. Research that has tested theories about the mechanisms of action of various exposure therapies for anxiety disorders has provided invaluable knowledge,[82] leading to the enhanced flexibility of exposure therapy, which in turn could be tailored for global adaptation. The findings of research on basic mechanisms will hopefully show the potential for brief and highly efficacious psychological treatments.[2] Future research will need to progress this work into the development of intervention formats and modules that are acceptable and efficacious cross-culturally, and that can be delivered on a wider scale.

The traditional models of one-to-one delivery of psychological treatments by skilled psychotherapists who have had many years of training need to be reconsidered, and new efficient methods of treatment delivery explored.[22,83] Given the small number of highly skilled and trained professionals internationally, a shift towards collaborative models of care delivery has been proposed in which novel strategies, such as task shifting (eg, educating lay people with no previous experience of the mental-health services to become lay counsellors; panel 5), have been successfully used to deliver streamlined treatment of mental disorders with promising results.[79,84,85] Nevertheless, empirical questions remain such as: how best to train people to become lay counsellors in a sustainable way? And what barriers might exist to such sustainability? One solution is the delivery of therapy to a group of patients rather than one-to-one.

Other research questions include: how many training, supervision, and booster sessions will be needed to ensure the high-quality delivery of treatments? Most studies in which potential treatment group leaders have received brief training (1–4 weeks) have shown effective outcomes,[86] but more research is needed in this context. These strategies of task shifting and training the trainer have been pioneering in the global context of mental health, as well as in developed countries. For example, the IAPT programme[24] resembles an advanced form of task shifting, rapidly educating a new category of mental-health professionals called psychological wellbeing practitioners, and the strengths and limitations of the programme can be of use to help in the improvement of future large-scale endeavours. How can technologies be used to train health-care workers on a large scale and maintain the reliability of treatment delivery? Primary-care clinics in the USA have used computerised guides to train inexperienced clinicians to give psychological treatments, albeit on a much smaller scale than IAPT.[87] The outcome and long-term follow-up data from such endeavours will yield many lessons on how to increase access to psychological treatments worldwide.

***Panel 5:* How can access to psychological treatments be increased worldwide?**

- Develop low cost, simple, specific, and effective treatments
- Task shifting: educate people who have not worked within the mental-health services to deliver psychological interventions
- Low-intensity intervention: self-help interventions comprising the most potent components of effective psychological treatments that can be provided through books, CDs or DVDs, the internet, or other media, combined with brief support—usually remote via e-mail or phone—over the course of a few weeks
- Cultural adaptation: rooting the treatment in the sociocultural context (eg, traditions, expectations, cultural norms, symbols) to make sure that it is perceived as intended

Technology is another important tool that can improve the availability of psychological treatments (see Part 5).[83] Providing psychological treatments via the internet or mobile phones, combined with minimal individual support through e-mail or telephone, has shown highly promising results in many studies in high-income countries;[88] however, few studies have tested such interventions in low-income and middle-income countries.[89] Further research is required, particularly since mobile phones are rapidly becoming available worldwide, and the availability of the internet is increasing.[90]

Low-intensity treatments delivered by computerised or mobile-based guided self-help technologies are an ideal early option in a stepped-care model of treatment. National guidelines are starting to propose the use of low-intensity treatments as a first option to improve the availability of efficacious treatments (eg, for bulimia nervosa and binge eating disorder[91]). Countries such as Sweden and Australia have led the way in research on internet-based treatment and the implementation of low-intensity treatments, with examples from eating disorders[92] to parent training[93] (for a meta-analysis of mental and somatic disorders see Anderson et al[94]). Work such as this provides models and lessons that can be used or developed to improve access to care worldwide—eg, the internet could offer enhanced possibilities for long-term follow-up after a standard course of psychological treatment has ended and the implementation of booster sessions.

Contextual factors have an essential role in any efforts to increase access to psychological treatments and are a topic for future implementation research. The involvement of all stakeholders is a key factor in scaling up services to ensure support and to facilitate sustainability.[95] Initiatives to improve mental health in low-income and middle-income countries need to be rooted in the local society to assure sustainability, and to inform ways to maximise and achieve this goal. Methods

to improve societal involvement could include engaging the local government, considering local legislations and traditions, involving patient organisations, and creating conditions for continued education and mutual exchange. One area that needs further research is the effort to help people who are refugees from war and persecution;[96] for these individuals, not only is the development of treatments essential, but particular contextual factors require investigation—eg, moving populations, multiple trauma experiences.

The stigma related to mental health problems is another barrier to improved access to treatment that requires further research. Understanding and addressing the association between religious or cultural beliefs and attitudes towards mental health is a crucial factor. The potential of media, such as radio and television, to combat the stigma related to mental health problems and seeking treatment for mental health problems warrants investigation. As an example, stigma is clearly associated with talking openly about family planning among people living in poor communities in some low-income and middle-income countries. The successful use of a well designed television series to improve family planning and to reduce fertility rates in Mexico is a good example of the effective application of such strategies to reduce stigma.[97] The Headspace initiative in Australia provides a model that could be adapted to different cultural contexts with the goal of decreasing the stigma of mental illness and facilitating access to treatment.

For the **Headspace initiative** see https://www.headspace.org.au

The economic aspects of international efforts to improve mental health should also be subject to more rigorous research. Evidence from the UK[98] suggests that psychological treatment approaches—eg, early intervention for psychosis, conduct disorder, and suicide prevention—can have a cost-effectiveness ratio higher than 10 (ie, for every £1 invested in such an intervention, there will be more than £10 of benefit). Future research designs should include cost-effectiveness analyses regarding the broader provision of psychological treatments in resource-limited settings, both in developed and developing countries.

Panel 6: **Example directions for future research to improve access to psychological treatments worldwide**

- Build brief, flexible, modular, and efficacious treatments that are streamlined on the basis of knowledge from research on mechanisms of action in psychological treatments
- With knowledge of the mechanisms of action of psychological treatments, derive treatments aligned with the local needs, priorities, traditions, and cultural factors, which will be specific to the environment in which the treatment will be given
- Investigate how much education and training is needed for people without or with little previous experience of work within mental health care to acquire proficiency to give basic psychological treatments as lay counsellors in a sustainable way
- Investigate how models of delivery of psychological treatments can be scaled up in a sustainable way
- Investigate the use of media such as television, radio, and the internet to combat the stigma related to mental disorders

Research collaboration and exchange between cultures

The best way to enable the improvement in psychological treatments would be by an international mutual exchange of knowledge, experience, and expertise between disciplines (panel 6). Opportunities for students and professionals—both scientific and clinical—from different parts of the world to visit one another and learn about conditions for, and challenges in, improving access to psychological treatments in different contexts could prove to be a key factor in creating the enthusiasm and lasting collaborations needed to develop sustainable interventions (see Part 7). Such exchanges could also facilitate cross-cultural comparisons that might contribute to understanding and more efficient prevention and treatment of mental disorders.

Work needs to continue towards increasing global access to psychological treatments, both for individuals in low-income and middle-income countries and those in high-income countries. Research into psychological treatments will allow the psychiatric community to continue to develop and assess the efficacy of brief and flexible interventions, which could be focussed on precise mechanisms of action, that could in turn be adapted to meet the needs of individuals in different cultural contexts. Training lay people to deliver such interventions, and scaling treatments for delivery in a manner that is sustainable in the long-term, are two key directions for future work.

Part 3: With what? The potential for synergistic treatment effects—using and developing cross-modal treatment approaches

Introduction

Both pharmacological and psychological interventions are commonly recommended as first-line treatments in psychiatry and the potential for enhancing treatment action through combination approaches has started to attract research interest. However, the optimal method for treatment combination is far from clear and requires dedicated research in preclinical studies, experimental medicine models, and randomised controlled trials. We advocate that such an approach should consider the potential for synergy between key mechanisms of action across different treatment modalities and consider these different treatment methods within the same research framework. The potential for negative effects from treatment combinations should be included in future research programmes.

Creating synergy and avoiding harm with combination treatments

An individual with a mental disorder or comorbid mental disorders is likely to receive a combination of different treatment approaches as part of his or her care, often

including psychological therapies, different types of medication, and social interventions (panel 7). However, clinical guidelines include little about combination treatments and the vast majority of research focuses on single treatments, often with the presence of another treatment as an exclusion criterion to participation in randomised controlled trials; although, some meta-analyses have been completed of existing studies on combination treatments.[101,102] The generalisation of research based on single treatments to the typical clinical reality of combination treatments is not always valid in practice. Therefore, exciting basic and clinical science questions arise about what happens when a psychological treatment is combined with other therapeutic approaches.

Empirical studies suggest that combination treatment might have small benefits over single treatments—eg, when a psychological treatment, such as CBT, and a pharmacological treatment, such as a selective serotonin-reuptake inhibitor (SSRI), are combined in the acute treatment of emotional disorders, including depression.[103] However, the longevity of effects after treatment discontinuation could actually be reduced in some cases compared with each single treatment alone. For example, in the treatment of anxiety disorders, post-treatment relapse has been reported to be higher in patients who also received benzodiazepine or antidepressant treatment during CBT than in those who received CBT alone or in combination with a placebo.[100,104] Findings such as these emphasise the importance of capturing clinical effects after treatment ends and during the acute response phase, and also of focusing on potential mechanisms that could underlie these differential outcomes (panel 7).

Mostly, combination treatments in the clinic are driven pragmatically—eg, an individual might receive two effective treatments, often with each from a different practitioner, such as a clinical psychologist and a psychiatrist. This sort of approach contrasts with attempts to combine treatments on the basis of a mechanistic understanding or model. The hope is that scientifically informed combination treatments have the potential to create synergy and to support a better therapeutic response than either treatment offered alone. This scientifically informed approach could be of use to potentiate the mechanisms that are hypothesised to support a therapeutic effect or to overcome the limitations or barriers to a particular mechanism applied on its own (see Part 1). Interventions that are given together with psychological treatments could include the addition of drugs, neuromodulation, social, nutritional, or other forms of psychological intervention such as computerised training (eg, cognitive bias modification).

Boosting psychological interventions by use of contemporary cognitive neuroscience research

Developments in neuroscience and experimental psychology[82] have been used by researchers who are focused on boosting the effects and retention of psychological treatments. Understanding the molecular basis of memory processes provides targets that might be manipulated to facilitate learning and the extinction and reconsolidation of memories, which are key components of many psychological treatments for a number of mental disorders.[64,105]

Augmentation of existing psychological treatments

A growing area of interest is the use of drugs that target the glutamatergic system (eg, d-cycloserine) to facilitate underlying processes of extinction and retention during exposure therapy for anxiety disorders such as agoraphobia, social anxiety, and PTSD.[63] However, identifying the factors that might moderate this benefit is challenging, and a 2015 Cochrane review[99] found no evidence that d-cycloserine versus placebo conferred any advantage overall when combined with CBT in the treatment of anxiety disorders. Techniques that directly stimulate the brain (eg, transcranial magnetic stimulation) applied over the medial prefrontal cortex have been reported to modulate conditioned fear learning and extinction in healthy volunteers.[106] Hopefully, add-on treatments that affect the underlying mechanisms of learning and memory might speed up overall treatment effects, reduce the number of treatment sessions needed, or even prolong treatment effects. However, better understanding is needed of the best methods to facilitate learning in an area about which much is still unknown. For example, the optimal parameters to support learning pharmacologically or through neuromodulatory devices are elusive and require dedicated strategic focus to support preclinical work in healthy volunteers and animal models (see Part 1).[63]

A focus on mechanistically derived combinations also requires an understanding of and the ability to predict the effects of a psychological treatment alone and in combination with other treatments—eg, enhancing learning by pharmacological means in an exposure

***Panel 7*: What is a combination treatment?**

Combination treatment

The application of two or more types of intervention that have been specifically assessed for efficacy in combination.

In this Commission, the combination of psychological treatments is referred to with other types of interventions across modalities, including the addition of drugs, neuromodulation, social, nutritional, or other forms of psychological intervention such as computerised training.

Synergistic *vs* harmful combination treatments

Some treatments might work well together and have greater efficacy than either applied on its own—eg, the use of a drug to improve learning has been hypothesised to enhance retention of the benefits of CBT;[63] however, no tested drugs exist that relaibly do this.[99]

By contrast, some treatments could impair efficacy in combination—eg, patients who receive benzodiazepines during psychological treatment can show reduced long-term benefits of CBT after drug discontinuation.[100]

CBT=cognitive behavioural therapy.

treatment that has failed, or in which extinction has not occurred, would be expected to have counterproductive effects, strengthening poor outcomes. These complexities underscore the necessity and potential effects of understanding the mechanisms of treatments in isolation and in combination.

The need for better preclinical models

These observations of the potential outcomes of combination treatment highlight the crucial role of preclinical and experimental medicine models in understanding both the key processes and mechanisms that are important for combination treatments and assessing early signals of efficacy for future clinical testing. Animal models are commonly of use in the pharmaceutical industry to screen novel drugs, but are rarely of use in a combination approach—ie, by testing the effect of a drug together with a psychological intervention. This single-treatment approach could lead to the early rejection of a drug that might have weak effects on its own, but which could be clinically useful in an adjunctive role with psychological treatments. Strategic focus and funding are needed for mechanistically informed approaches to treatment combination in animal and human models. The back translation of findings from the clinic to these models needs to be enhanced, and increased interest is needed in using combination models to assess novel treatments, including as part of drug development within the pharmaceutical industry. Research in this area needs to incorporate measures that can assess and predict when and for whom combination treatment will be helpful. Regulatory support for this approach from the US Food and Drug Association and the European Medicines Agency, linked to approval and licensing of drugs, will be required to allow pharmaceutical companies to develop and test these kinds of combined treatments, both to facilitate potentially beneficial combinations and to reduce potentially harmful ones.

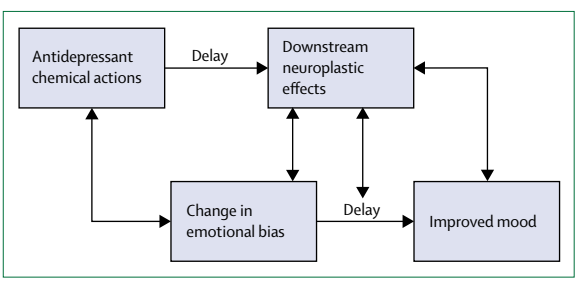


Figure 2: **Combining antidepressant drugs and psychological interventions to speed up the therapeutic effects**
Antidepressant drugs are hypothesised to work via early changes in negative affective bias—ie, by reducing the influence of this key maintaining factor in depression.[108] This theory raises the possibility that psychological treatments could be used in combination with chemical actions to boost the effect of antidepressants on negative affective bias, avoid delays in action, and facilitate the translation of effects on bias into clinical action—ie, improved mood. Reproduced from Harmer et al,[108] with permission from Elsevier.

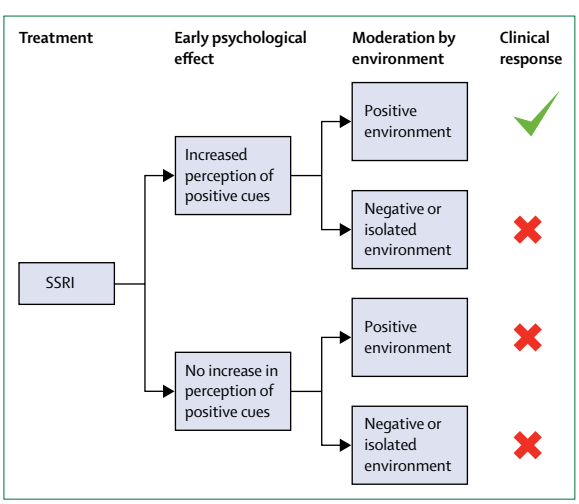


Figure 3: **Effect of patient environment on the clinical efficacy of antidepressant drugs**
Increased perception of positive cues has been associated with delayed clinical response with SSRI treatment, but this effect is moderated by environmental and social factors. Therefore, increased positive bias is only associated with improvements in depression in the context of a relatively supportive or positive environment. In the absence of changes in emotional bias, the patient's environment has little effect.[111] SSRI=selective serotonin-reuptake inhibitor.

Unifying approaches and measures across treatment research

Treatment combination across different treatment modalities can be restricted by barriers between researchers, clinicians, and funders. These barriers include different frameworks, languages, focus, and outcome measures, making it difficult to see natural synergy between the fields. However, exploring treatments using a common framework could help to overcome these barriers and lead to novel hypotheses that could not be predicted by a single approach alone. For example, studies have used measures across scientific fields to understand treatment effects, such as neuroimaging to understand and predict therapeutic response to psychological treatments,[107] and psychological outcome measures to explore the effects of drug treatment.[108]

As an example, efforts to understand the mechanism of an antidepressant drug usually focus on the molecular, cellular, or chemical interactions, but evidence is increasing that antidepressants affect core psychological processes that are important in depression before therapeutic effects are observed, which could help explain their delayed clinical actions in depression (figure 2).[108] Antidepressants increase the relative processing of positive versus negative information early in treatment, which could be important in the recovery process from depression since the patient has more positive feedback and reinforcement, countering the negative biases that are hypothesised to play a key role in maintaining the disorder.[109,110]

A key barrier to the successful translation of these effects into clinical benefit is the need for interaction with the environment. If a patient is socially isolated or in a socially detrimental environment, then increased

positive bias and processing would be expected to have only a small effect. Shiroma and colleagues[111] reported that increased positive bias, induced by treatment with antidepressant drugs, interacted with interpersonal support in the patients' environment to predict the therapeutic response (figure 3). This kind of interdisciplinary approach to treatment has the potential to generate new hypotheses concerning combination treatment that would not have been predicted from either approach alone. Using this example, the combination of early phase treatment with an antidepressant drug in combination with a psychological intervention is predicted to increase the patient's interaction with the environment (eg, behavioural activation), and could remove a barrier to successful treatment with an antidepressant drug (figure 2).[108]

To facilitate interdisciplinary combination approaches to treatment, increased communication and translation are key. Greater collaboration and joint meetings, the use of similar assessments and measures, and joint funding initiatives will help support this aim to improve combination treatments in the future. These improvements will require organisations, funding bodies, and researchers to work together, but the results will no doubt be exciting. An example of this collaborative approach to treatment occurred following a joint symposium and was presented at two very different meetings; the British Association for Psychopharmacology and the British Association for Behavioural and Cognitive Psychotherapies. The joint symposium, supported by the charity MQ: Transforming Mental Health, focused on the divide between psychological and biological treatment development and stimulated approaches to start to bridge the gap and align research strategies between psychopharmacology and psychotherapy.[112] Researchers in the field need to build on this exciting initiative, call researchers across all mental-health fields, and get strategic funding to strengthen this promising endeavour.

Testing the efficacy of combination treatments

Developing and assessing the efficacy of combination treatment also creates complexities in trial design and methodology (see Part 6). Treatment trials that compare active treatment with control treatment often require large sample sizes to have sufficient statistical power to isolate true effects from demand or placebo effects. Exploring interaction effects in comparison with individual treatments can require even larger sample sizes, depending on the study design. In particular, the effects of two treatments will often be assessed in isolation, as well as in combination, leading to a factorial design with four groups. Mechanistic studies also need to consider possible state dependency of learning—ie, that memory will be enhanced if tested in the same state versus a different state, including internal states produced by a drug.[113] The field of combination treatments will therefore benefit from a variety of approaches and from testing the effects of treatment at different time points and under multiple conditions.

Experimental medicine can be used to test hypotheses in smaller controlled studies and using surrogate markers of treatment success. This approach has revealed key effects of both pharmacological[114] and psychological[115] treatments that are used for anxiety disorders on the same underlying cognitive processes, and it has been used to explore the effects of combined treatment. For example, the effects of pairing computerised training for cognitive bias modification with brain stimulation of the dorsolateral prefrontal cortex were assessed using reactivity to a stressor as a proxy marker of efficacy in healthy volunteers.[116] The effects of cognitive bias modification and SSRI treatment alone and in combination have been explored by use of the same outcome measure, along with effects on negative memory bias. The results of this study showed that the combined effects could be worse than either applied in isolation in healthy volunteers.[117] Early changes in these measures are associated with delayed therapeutic benefit in patients[111] and can therefore be of use to guide initial proof-of-principle studies for treatment combinations and to reject those that have little therapeutic promise. Combinations that appear to be successful with these surrogate markers can be put forward for the next stage of clinical assessment, typically in a randomised controlled trial. This approach might be supported by big-data approaches in which the data are collected under more naturalistic conditions (eg, large-scale analysis of medical records or prescribing patterns; figure 4). Promising treatment combinations and timing of treatment combinations might be isolated by pattern analysis from large datasets. To facilitate this analysis, assessment and treatment elements must be standardised (see Part 8). The triangulation of experimental medicine, randomised controlled trials, and big-data analysis will be necessary to develop, assess, and understand combination approaches of the future.

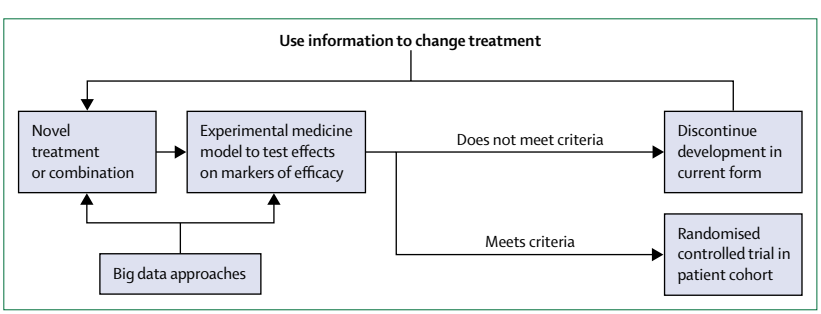


Figure 4: **Experimental medicine models for earlier assessment of efficacy of novel treatments and combinations**

Surrogate markers within experimental medicine models can be of use to screen new treatment combinations in small groups of patients or volunteers. This information is used to refine decision making for subsequent application in and design of randomised controlled trials. If insufficient evidence of efficacy is seen in the model, this information can be used to change treatment focus, the dose or duration, or the treatment target. If pre-set criteria are met, the efficacy of the treatment combination can be assessed using randomised trial designs. Approaches with big data can be useful to highlight particularly promising treatments or combinations and provide additional evidence of efficacy from naturalistic data-capture methods.

***Panel 8:* Potential future research directions in combination treatment**

- Development and validation of preclinical animal and human models for proof-of-principle studies and mechanistic focus in combination-treatment research
- Elucidating the optimal parameters for enhanced learning with drug-treatment approaches and the consideration of individual differences in this response
- Encouraging pharmaceutical companies to develop and assess novel drugs for a combinative role with psychological interventions; cultivate an understanding of this approach within the regulatory community
- Clinical studies informed by proof-of-principle work to test the efficacy of treatments alone and in combination across mental disorders
- Consideration of the potentially harmful effects of combination treatment for treatments that work well in isolation, including a focus on attribution bias and long-term outcomes
- Research the views and acceptability of combined treatments in mental disorders and the importance of patient preference and views about treatment for their clinical symptoms
- Patient preference needs to be considered in formal research programmes that attempt to bridge the psychological–pharmacological divide; the views, acceptance, and opinions of the individual receiving treatment can influence its effects
- Preclinical research using animal or human models is needed to understand key mechanisms and the effects of novel interventions before translation to clinical research programmes
- Back translation can be used to determine the success of translational research since success depends in large part on the validity of the experimental model that is used to mimic the disorder in the laboratory; back translation describes how evidence from clinical research and experience is used to drive, test, and refine the development and validation of animal and human preclinical models
- Experimental medicine and experimental psychopathology: investigators use models, typically human models in laboratory settings, to explore key mechanisms and processes that are hypothesised to be important for treatment action in psychiatry and psychology; these models can be of use for screening novel treatments and refining their application before full clinical testing

Breaking down barriers: from patient perspectives to research of the future

Finally, patient preference should be considered when assessing the effects of combination treatment. Individuals often express a preference for either psychological or pharmacological treatment, so the option of a combination of treatments might be a difficult choice for some. This view that a dichotomy exists between a psychological or biological view of mental disorders is challenged by evidence that psychological and biological treatments tap into the same core processes and represent different methods, rather than different concepts.[108] Challenging these assumptions and creating synergy at multiple levels (including among the public, clinicians, and scientists) will be a crucial step towards the optimal development of treatments. The ethical implications of combination treatments and their development should be incorporated within research strategies for these areas. Additionally, the attribution of treatment effects needs to be considered from the patient's perspective—eg, if any benefits from combined treatments are attributed to the medication, then the long-term advantage of CBT can be lessened.[118] Studies to characterise attribution bias in combined treatment approaches and consideration of the strategies that might be effective in reducing these effects are key priorities for future work (panel 8).

In summary, research to date that has tested the efficacy of combination treatments has shown great promise for the clinical utility of combining psychological and pharmacological approaches. However, many unanswered questions remain that need to be addressed regarding the optimal method for treatment combination in preclinical studies, experimental medicine models, and randomised controlled trials.

Part 4: When in life? Psychological science, prevention, and early intervention—getting the approach right from the start

Introduction

Opportunities for prevention and early intervention into mental health problems exist throughout a patient's lifespan; however, the early years of life are perhaps the best opportunity to set an individual on a path to good mental health. This process requires both population-based change and the accurate identification of those at risk, and for both approaches effective and safe interventions are needed. Many approaches have little or no scientific underpinning, and so the rigorous and sustained application of approaches that are based on psychological science to this area of practice is crucial and offers enormous promise. The focus of this section is primarily on children and young people.

Prevention and early intervention

The prevention of mental disorders is one of the main challenges for the future of mental health care because of their high burden of disease for individuals and societies, the relatively small effect of treatments to date, and the enormous societal costs of mental disorders once they have emerged.[119] The imperative to reduce risk factors across the population and to intervene at the earliest point when symptoms or precursors of mental distress occur makes sense on a human, societal, and economic level.[120,121] Psychological science can inform and underpin the development of early preventive interventions, even if the risk factors are social in origin.

The early years of life, from conception through to childhood and adolescence, are a good opportunity to set an individual on a path to good mental health. Most mental disorders have their origin or onset before the age of 18 years.[122] The greater plasticity of the brain during childhood and the nature of the emotional and behavioural responses of a child mean that the potential to intervene successfully and powerfully could be greater than at any other point in life. Nowadays, the potential role in early life for psychological approaches is greater than that of pharmacological and other physical interventions; however, many interventions remain under-researched, such as nutritional approaches. For psychological interventions to make progress into the

effective prevention of mental disorders, some key requirements and scientific and clinical challenges have to be met.[2]

Requirements and challenges for prevention and early intervention

Preventive approaches in childhood and adolescence (panel 9) require the identification of risk factors or at-risk groups (unless an intervention is going to be delivered to the whole population).[120] Key risk factors in early life include exposure to severe adversities, such as maltreatment, disturbed parenting, parental substance misuse, exposure to domestic and other violence, and loss events—eg, serious parental illness or death of a parent.[126] However, further research is needed into these and additional risk factors, as well as into the interactive effect of risk factors.

For change to occur, effective and acceptable interventions should be available. These interventions might target modifiable risk factors or use other theoretical approaches to affect change, including tackling key psychological mechanisms. However, many early interventions do not have sufficient evidence to be considered as effective. Developing and testing early interventions that might reduce the risk of psychological illness is a fundamental and largely unmet challenge.

Current research limitations regarding early interventions

Any kind of early intervention is often implicitly assumed to be better than no intervention, but this assumption is not correct. Almost any intervention that can do or change something has the potential to cause harm if applied in the wrong circumstances, as discussed by Carter and colleagues[127] regarding eating disorders. The possibility for harm is often overlooked and is probably one of the key blind spots in the field of prevention of psychological problems, particularly when translated into policy. Crucially, clincians and researchers need to acknowledge that not all interventions are the same; even those interventions that overlap in appearance or content can have different effects.[128]

A paucity of evidence on the effectiveness of psychological treatments exists in many areas of child and adolescent mental health practice, particularly for very young children. However, this area does hold promise since the differences in effectiveness for different treatments can be seen where high-quality evidence exists.[60,129] A related consideration is that an intervention might not have the same treatment effect in every setting or with all individuals equally (eg, the apparently contradictory findings for the Family Nurse Partnership intervention[130]). Disentangling these challenging problems is made more difficult if the components of a psychological intervention are not clearly specified or publicly available, perhaps because of some commercial or other protective reason.

***Panel 9:* Psychological treatments: what are preventive and early interventions?**

Prevention is often defined as those interventions that are done before people meet formal criteria for a disorder.[123]

Three types are described:

- Universal prevention, which is aimed at the general population or parts of the general population regardless of whether they have a higher than average risk of developing a disorder (eg, school programmes or mass media campaigns).
- Selective prevention, which is aimed at high-risk groups who have not yet developed a mental disorder (eg, the Nurse Family Partnership programme developed in the USA that initially aimed to prevent later adverse psychosocial outcomes for women and their children in socioeconomically deprived areas).[124]
- Indicated prevention, which is aimed at individuals who have some symptoms of a mental disorder but do not meet diagnostic criteria (eg, the intervention developed by Rapee[125] for parents of preschool-aged children who are at risk of anxiety disorders).

A further challenge is the paucity of understanding of the mechanisms by which an intervention occurs in many preventive and early interventions. As set out in Part 1, an understanding of the mechanism of action is crucial to the development of new and more effective methods of successful treatment. However, the mechanisms of action are likely to be more changeable in early life than at other points in life, complicating efforts to understand them in a preventive and developmental context. For example, different mechanisms could operate at different points in childhood, and each of these mechanisms could be different from those operating in adulthood, even for the same condition or problem that is presented (see Part 8). Although few well studied examples of this divergence between childhood and adult mechanisms seem to exist, studies are emerging—eg, Ewing and colleagues[131] found that children at risk of anxiety disorders do not have the specific cognitive biases for emotional stimuli that are seen in adults at risk of anixety disorders. For patients in early childhood, clinicians and researchers will need to go beyond the individualised mechanisms suggested in the RDoC explanatory constructs (see Part 1). Instead, other mechanisms existing in the social world of young children might open crucial pathways to help change precursors of psychopathology—eg, via the early relationships or attachments that children form to their parents or carers. Parental sensitivity has been shown to be a key mechanism of change (eg, in the context of attachment),[132] although the detailed processes which might then lead to the development of psychopathology remain to be elucidated.

Making interventions stick—persistence of effects

Another challenge for preventive and early intervention approaches, which is shared with many other forms of psychological intervention, is how to make interventions stick—ie, not only how to make the effects of psychological treatment last beyond the end of the treatment, but also how to make them generalise to other areas of functioning. Few psychological interventions have convincing evidence of sustained benefit.

Panel 10: **Examples of promising preventive and early intervention approaches**

Example 1: video feedback to promote positive parenting

During infancy, brief and focussed interventions, such as video feedback to promote positive parenting,[135] can improve parental sensitivity and the child's attachment relationship with their primary carer or parent; this technique draws on both attachment theory and social learning theory; some evidence of effects on child behaviour exist for this intervention, which are largely lacking for other video feedback parent-focused approaches to date.

Example 2: parental interventions for childhood anxiety

An intervention for parents of children aged 3-5 years who have an increased risk of anxiety disorders (identified by having high levels of behavioural inhibition) has been shown to reduce the risk of subsequent anxiety disorders within the child; this intervention was brief (six sessions), and used an educational approach with some behavioural components focussed on exposure; effects from the treatment were still seen 11 years later, although only convincingly in girls, and were shown to be cost-effective using Australian criteria for cost-effectiveness.[125]

Example 3: parenting programmes for child behavioural problems

Among school-age children (aged 3–7 years), consistent evidence has shown the benefit of parenting groups based on social learning theory, such as Scott and colleagues' Parenting Programmes to improve child behaviour;[136] longer-lasting benefits have been shown, and economic modelling studies point to societal, financial, and individual health gains.[137]

Panel 11: **Research questions in prevention and early interventions**

- When are the optimal times to intervene to prevent mental disorders?
- Who are the key at-risk groups that will most effectively respond to early or preventive treatment?
- What are the potential harmful effects of specific early-intervention approaches?
- How do we increase the so-called stickiness of treatment effects? How do we make them last beyond the end of treatment?
- How can we deliver interventions on the scale needed (including internationally) to reach at-risk children and young people?
- How can insights from mechanisms of change help prevent or delay disorders and reduce the recurrence of episodes?
- How can insights about prevention be applied across the human lifespan?

Developments are needed in psychological science to inform how to take psychological interventions outside of the therapy room—which could make interventions more widely available and acceptable, and make the effects of interventions more generaliseable to everyday life functioning. Technologies could help in this regard (see Part 5)—eg, by use of gaming and other technologies to prevent or treat early signs of depression.[133] A further approach is to take interventions into schools.[134] To date, both of these approaches have utilised primarily cognitive behavioural interventions, although other approaches, such as interpersonal therapy, also show promise for the treatment of depression in children and young people.

Positive examples for the future

Panel 10 contains three examples of intervention types for young children and their parents that have shown that preventive and early interventions are possible from very early in life, and that longer-lasting benefits are possible. All three interventions are derived from scientifically rigorous and sustained approaches to intervention development and are informed by theory. Other preventive or early interventions do exist, with varying levels of research evidence to support them, for a range of psychological and psychiatric conditions.

Prevention of mental disorders in adults

In the past two decades, randomised controlled trials have shown that preventing or at least delaying the onset of mental disorders is possible in adolescents and young adults, especially depression and psychotic disorders. Psychosocial preventive interventions, typically based on psychological treatments such as CBT or interpersonal psychotherapy, have been tested in at-risk populations and in people with subthreshold symptoms of depression or psychosis. Meta-analyses[138,139] confirm that these interventions effectively reduce the incidence of new cases of depressive disorders by about 20–25%, and prevent or delay the onset of about 50% of psychotic disorders in those at high risk for developing a psychotic disorder.[140,141] Preventing the onset of mental disorders is one of the most promising areas in which research on psychological interventions can help to reduce the disease burden of mental disorders.

The challenges ahead

Clearly, more research is needed to expand the repertoire of approaches and the range of mental disorders that can be treated. These approaches need to be theory driven and rigorously trialled (see Part 1 and Part 6).

Particular attention should be given to ensuring that interventions can produce effects with lasting benefits for children and adolescents, and substantial efforts need to be made to develop or adapt interventions so that they can be of use across a range of settings and accessible internationally (see Part 2).[142] Although preventive and early intervention approaches for mental disorders potentially have huge health benefits, they face particular challenges in terms of showing reliable efficacy and being applied consistently and thoughtfully in everyday practice in health care. The examples considered in this section provide optimism for future developments, but health-care professionals and researchers need to look carefully at the limits of effectiveness, and at the potential to cause harm (eg, potential negative effects of screening and classifying high-risk groups, offering unnecessary treatment to young people with only temporary distress or symptoms, or harmful side-effects of individual psychological treatments; panel 11). Knowledge of these benefits and harms should be pooled from patients of all ages. Although a lot of work still has to be done before effective methods of prevention for mental disorders are widely available, the rigorous and sustained application of psychological-science approaches to these areas of practice offers enormous promise.

Part 5: Technology—can we transform the availability and efficacy of psychological treatment through new technologies?

Introduction

Internet-based psychological treatments have been applied across a broad range of mental disorders. The rise of eHealth and mHealth approaches that use information technology (eg, the internet, virtual reality, serious gaming) and mobile and wireless applications (eg, text messaging, apps) marks a new era for psychological assessment and treatments. Technological innovations offer considerable possibilities to innovate psychological treatments, adjust them to daily life and the people using them, and improve access to treatment. Such knowledge could be of use to better understand how therapies work, make them easier to use, and enable more people to benefit from psychological treatments. Developments in technology-based treatments should be theory driven and properly assessed.

Internet-based psychological treatments

Most research into psychological treatments has been done with somewhat traditional internet interventions. In these interventions, patients work through self-help materials on a computer, learning how to apply a psychological treatment to themselves with the help of a coach or psychologist.[143] Such self-help materials have often been very close in content to face-to-face psychological therapy (eg, CBT). Accordingly, the materials are as if a hard-copy paper manual has been converted into a computerised form, sometimes with simple additional content such as video clips. Direct comparisons between face-to-face interventions and guided internet interventions suggest that no major differences are apparent in efficacy between the two treatment formats.[94] The efficacy of internet-based therapies (see appendix) has been shown for a broad range of mental disorders, including depression,[144] anxiety disorders,[145] sleep problems,[146] bulimia,[147] and alcohol problems.[148]

Internet interventions have many advantages, including saving time for therapists, reducing waiting lists, allowing patients to work at their own pace, removing the need to schedule appointments with a therapist, saving travelling time, reducing the stigma of going to a therapist, and facilitating psychological help for individuals who are hard of hearing.[149] Furthermore, internet interventions might reach patients who cannot be reached with more traditional forms of treatment (eg, because of distance or stigma). Interventions can be quite easily adapted to specific patient groups, with a wide range of attractive audiovisual information with voices giving instructions via a character of whichever gender or age, with whichever accent or language, or perhaps game format, the patient prefers. Internet interventions are probably more cost-effective than face-to-face treatments, but further economic research is needed to verify this.

From a research perspective, internet interventions have many advantages. One major advantage is that recruiting patients for randomised controlled trials of internet interventions is much easier and more cost-effective and efficient than doing trials of traditional face-to-face psychotherapies (see Part 6). Research into these interventions should stimulate further development of personalised treatments for mental disorders by allowing large-scale trials that are powered to examine complex questions (see Part 8) or test for weaker effects (eg, prevention trials).

However, internet interventions have limitations. The quality of interventions that are offered through the internet is not clear, and despite portals for evidence-based internet therapies, such as Beacon, the possibility that low-quality therapies are being offered remains a threat. Beacon is a webservice through which a panel of health experts categorise, review, and rate websites and mobile applications for internet-based psychological treatments. It is part of a suite of self-help programmes that have been developed and delivered by the National Institute for Mental Health Research at the Australian National University, although it is unfortunately not being updated. Drop-out rates are higher in internet-based interventions than in face-to-face therapies,[150] and it is unknown whether the condition of these patients gets worse as a result of the intervention, or in general, since they cannot be followed-up. Internet interventions might affect the therapeutic alliance between therapists and patients, but most evidence suggests that internet-based therapies are at least equivalent to face-to-face therapies in terms of therapeutic alliance.[151] Little research has focused on the long-term effects of internet interventions; however, the same is true for face-to-face psychological treatments. Furthermore, we acknowledge that internet interventions might have unknown disadvantages, such as misunderstandings due to reduced communication channels in unguided interventions and the potentially confusing depiction of content as graphs and schemes. Additionally, data security and privacy should be guarded for any intervention that is offered through the internet.

For **Beacon** website see https://beacon.anu.edu.au

See **Online** for appendix

Finally, despite increasing access, the internet is not yet accessible to many potential users around the world, and dissemination will depend on the attitudes of possible users and health-care professionals. However, even in low-income and middle-income countries, access to the internet and mobile phones is expanding (see Part 2), although creative solutions (eg, regarding literacy) might need to be taken into consideration where applicable.

Other technological opportunities

Interventions can increasingly be offered through smart phones and watches, Google glasses, virtual-reality headsets, and other kinds of innovative devices. Many of these devices have the advantage that they can be worn by the patient and collect information during daily life (ecological momentary assessment;[152] see Part 8). The

Panel 12: **Potential directions for future research with new technologies for psychological treatments**

- Treatment and theory development: health behaviour theory can be of use to inform technological treatment innovation across all areas of psychological treatments
- Treatment evaluation: trials to assess the effectiveness of new products such as apps
- Learning: maximising and innovating learning methods during psychological treatment by fresh means—eg, skills learning, habit change—such as via serious gaming
- Devices: the incorporation of new technologies—eg, avatars , smart watches, and other devices—into existing psychological treatments to facilitate delivery and improve outcomes
- Harnessing new technologies to advance methods of examining causal mechanisms, refine treatments, and derive treatment approaches that are mechanistically driven
- Health monitoring: enable large-scale data mining and data interpretation to predict the onset and course of mental disorders
- Personalisation of technology-based interventions
- Technologically aided preventive treatment approaches adapted across all age ranges and globally

information collected might considerably improve prediction models for individual patients and thus potentially improve and increase the effect sizes of existing treatments. Computerised adaptive-testing techniques assess symptoms online with greater sensitivity and specificity from fewer items than traditional forms of outcome monitoring—ie, pen and paper questionnaires.[153] Several virtual-reality treatments have been developed, mainly for anxiety disorders. Patients are not confronted with the real stimuli that provoke their anxiety but with their virtual counterparts using real-time computer graphics, body-tracking devices, and other sensory input devices.[154] This form of treatment has shown some effectiveness;[155] however, many of the trials have been small and of suboptimal quality. Many studies have shown that telephone-supported therapies are effective in the treatment of common mental disorders.[156]

The range of mental-health applications (ie, apps) available is rapidly growing, offering a range of psychological interventions;[157] however, most apps are not based on health behaviour theory and little evidence supports their effectiveness.[158] Future researchers should develop theory-driven interventions and assess their effectiveness, since only a few interventions have been tested in randomised controlled trials.[159,160] Specific adaptations to the design of a randomised trial might be needed because of rapid technological developments.[161] Widely available and untested products pose a risk to the public. Although the field of technology-based interventions is still young, and efforts to progress treatment development have started, international approaches are needed to develop regulated approaches and procedures.

The format of new technologies could allow new treatment techniques to be developed that are not part of existing face-to-face psychological treatments, offering novel information processing options (eg, virtual-reality exposure, and possibly interpretation of bias training). Serious gaming, such as the SPARX program, also opens opportunities for interdisciplinary research and new methods of treatment delivery.[133] Serious games refer to games with a purpose other than providing entertainment, which in this case is the delivery of a psychological treatment using game principles. SPARX is an interactive fantasy game designed to give CBT to adolescents seeking help for depression.

At some point, the automated support of these new technologies might replace the therapist altogether (ie, therapist-free therapy[162]), and lead to improved, personalised treatments (see Part 8). New technologies can also be of use in predicting the development and outcome of mental disorders. For example, mobile phone apps are available to monitor associations between psychological risk and suicidal ideation,[163] and evidence exists that the use of specific phrases and personal pronouns can, for example, predict an individual's depression status from their blog posts (see Part 9), although we acknowledge that such monitoring could raise ethical concerns.[163] Because of the huge quantities of data can be collected through mobile phones and other devices that can be connected with existing databases, data-mining techniques could be helpful to predict the onset and course of mental disorders. This data accumulation could aid the development of innovative psychological interventions that could be integrated into new technologies that become part of the daily lives of patients. However, to increase the likelihood of success, new technology and data accumulation alone will not suffice. A sound theoretical framework should be incorporated to drive hypothesis alongside clinical knowledge.

Finally, eHealth and mHealth approaches that use information technology and mobile and wireless applications are examples of ways that technology has been harnessed to innovate psychological treatments, their availability, and their assessment. Technology-based treatments need to improve with advances in psychological theory and understanding of mechanisms of change. Future technological innovations offer considerable possibilities to innovate psychological treatments (panel 12), including adjusting treatments to patients' daily lives and using the information gained to better understand how therapies work, improve the treatments, and improve the technology's ease of use, so that people across all age ranges and worldwide can benefit from psychological treatments.

Part 6: Trials to assess psychological therapies

Introduction

Several key issues in the design and conduct of clinical trials to assess psychological therapies must be addressed to develop therapies that are evidence based. These issues offer several opportunities for improvement and

some specific challenges, given the complexities of both the therapies being assessed and the populations who are receiving them. The challenges include more accurate reporting of clinical trials, eg, specification of therapy protocols and inclusion and exclusion criteria, choice of outcome measures, measurement of adverse effects, and prevention of bias in trial design and analysis. The opportunities include the increasing role of service users and carers in all aspects of trial design and conduct, the development of methodologies for achieving a consensus regarding research questions and outcome measures, the development of new methods for analysis of mediators and mechanisms, and innovations in the design of clinical trials (eg, adaptive trial designs and mixed methods approaches that incorporate nested qualitative studies).

These challenges and opportunities will be considered in this section of the Commission in the context of a feasibility study (the COMPARE trial, ISRCTN06022197) and the potential for a subsequent trial to assess CBT for people with psychosis. This subsequent trial would be a direct comparison of CBT, an antipsychotic medication, and as a combined treatment, which is a research recommendation in the UK National Institute for Health and Care Excellence guideline for the treatment of psychosis in children and young people (for additional information see appendix).[164]

The need to improve clinical trial methodology

Clinical trials are the cornerstone of evidence-based approaches to decisions about access to health care, but in the field of mental health such trials often have substantial methodological shortcomings that result in low-quality evidence. Many psychotherapy trials are not registered in an international database before recruitment starts,[165] therefore other researchers cannot be sure whether the outcomes that are reported were those originally intended, and raises the possibility of selective reporting of outcomes (ie, focusing on those results that were statistically significant), or that negative trials remain unpublished. A systematic review[166] found that many psychotherapy trials did not attempt to maintain blinding (ie, masking) in the people rating the treatment outcomes increasing the likelihood of bias. Additionally, treatment protocols were broad and not based on a specific model, which makes assessment of fidelity and replication problematic. These limitations could be overcome by ensuring linkage between experts in trial design and methodology and statisticians and innovators in psychological therapy development. Accredited clinical trials units, with their extensive experience of trial design and conduct, could coordinate with academic methodologists who are at the forefront of developments in trial statistics and methodologies.[167] In the past decade, substantial improvements in psychological treatment trials have been made, with more studies adopting clinical trial registration and pre-specification of primary outcomes, including application of Consolidated Standards of Reporting Trials (CONSORT) criteria (appendix). Such procedures are increasingly required by leading journals and ethical review boards. However, to apply these procedures to psychotherapy trials particular adaptations of both trial design and reporting guidelines will need to be developed—eg, around issues such as double-blinded studies, a trial design that cannot be maintained with a therapist-delivered psychological treatment. However, double-blinded studies can also be problematic for pharmacological treatments, since aspects of the treatments can become apparent despite the investigators' best intentions—eg, the rapid and dramatic weight gain and parkinsonian side-effects found with both first-generation and second-generation antipsychotics. Another possibility is that subjective cognitive effects[168] unmask participants.

Additionally, the potential negative effects of psychotherapy are increasingly being recognised, and unwanted effects and serious adverse events need to be documented and reported to ethics committees as part of safety monitoring. Historically, psychological therapy trials have been poor at both monitoring hypothesised side-effects and deterioration, and reporting serious adverse events.[169] Negative effects and adverse events that require documenting range from the worsening of existing symptoms, to issues such as novel symptoms, poor therapeutic relationship, and perceived coercion.[170] Such adverse events are possible in both traditional psychotherapy and internet-based interventions.[171] A procedural model and checklist are available for clinicians and researchers,[172] and the detection and management of adverse events in treatment trials is considered a sign of good practice. Formalised measures of possible harms (ie, side-effects) caused by trials should be the rule, rather than the exception, in psychotherapy research.[169]

To ensure that psychological therapy trials are credible, the minimum standards expected in other fields should be met (eg, those standards in pharmaceutical trials). However, psychotherapy researchers have an opportunity to develop their own standards, which could ensure superior trial design, conduct, and reporting, which other fields could aspire to meet.

A set of reporting standards specifically tailored to psychological therapy trials are being developed as an extension of the original CONSORT guidelines.[173] These reporting standards include recommendations to improve internal and external validity, address measurement issues (psychological therapy trials often have many measures, of which many assess latent constructs), improve reporting of recruitment processes and representativeness of participant groups, and increase contextual information—eg, factors that helped or hindered the interventions. Additionally, research on general trial methodology (eg, on how to deal with the issue of masking participants) will be an important area of future inquiry.

Conflicts of interest

Management of a clinical trial by the developer of a psychological therapy could be considered equivalent in terms of bias to a pharmaceutical company managing a drug trial, and investigator allegiance effects have been observed in psychological therapy trials.[174,175] The focus of investigations into this bias has been more on allegiance to a given type of psychotherapy than on financial interests. Steps can be taken to reduce bias, including the declaration of interests (ie, personal financial interests such as training fees, book royalties, and non-financial interests), registration of protocols, prespecification of plans for statistical analysis, and involvement of independent methodologists in the trial design and data analysis. Trial steering committees and data-monitoring committees with independent clinical, statistical, and service-user representation also increase study confidence and minimise bias. These committees can provide constructive criticism and protect the safety of participants and scientific integrity of the trial. Expertise in all relevant approaches is important for trials that compare two or more therapies—eg, the team for the COMPARE trial includes researchers with expertise in both CBT and antipsychotic medication.

Inclusion and exclusion criteria

The selection and justification of inclusion and exclusion criteria are crucial to good trial design. The criteria should be specific enough to allow the identification of suitable participants and replication of a trial, but broad enough to reflect real-world clinical settings and permit generalisability according to the purpose of the trial. Historically, many psychological therapy trials require a single diagnostic category or symptom as an entry criterion, not allowing those with several or at least specific comorbidities (eg, other mental disorders, physical health issues, drug or alcohol use). These exclusion criteria are difficult to justify when the clinical reality of a mental health difficulties is complex and comorbidities are the norm (see Part 8). Trials in the past two decades that have assessed CBT for psychosis have typically been good in terms of generalisability, allowing for inclusion of participants who meet broad criteria (which is also true for trials of psychological therapy for depression[176]). Even trials that have focused on mechanisms of change—eg, whether reducing worry processes results in a reduction in paranoid thinking—have allowed participants with comorbidities.[177] However, in these situations compromises might have been made between clinical pragmatism (ie, having broad entry criteria) and the ability to scrutinise specific mechanisms within the trial. Trials that attempt to address transdiagnostic processes by targeting a specific mechanism (eg, modification of attention biases or extended perseverative processing) or problem (eg, sleep difficulties or irritability) across diagnostic groups offer potential advantages in terms of recruitment, generalisability, and implementation in mental health care (for further discussion of these issues see Part 8).

Improved integration of research trials within clinical settings would facilitate the generalisability of results to the real world. One goal is for every individual who attends a hospital clinic because of a mental health problem, or engages with a community mental health team, or attends an appointment in primary care, to be offered participation in psychological therapy research (if willing and able to provide consent). For example, for interventions for which genuine uncertainties in treatment exist (eg, what dose of CBT for psychosis is required), all willing participants could be randomised into groups with different treatment durations.

Choice of control condition

Appropriate control conditions for psychological therapy trials are a matter of considerable debate—eg many argue that so-called treatment as usual is not appropriate since such conditions can be highly variable and at times include access to the treatment that is being provided in the experimental group. The use of an active control condition is often recommended, which reduces confounds such as non-specific factors (eg, attention, warmth, human relationships); however, inclusion of an active control condition could oversimplify the issue of therapeutic relationship—itself a topic of research and debate about its importance. The provision of an alternative therapy can raise other confounds, such as the so-called match between therapist and participant, and the ability of a therapist to switch between, and adhere to, different treatment protocols even though they probably have greater skill and allegiance to one protocol over another. Ways to deal with such issues include having multiple therapists who can provide each active condition, perhaps across trial sites, so that different trial sites can have different expertise but can provide all therapies (eg, a trial of CBT for psychosis compared with befriending).[178] Furthermore, mental health problems might differ in their response to psychological placebos—eg, the effects of non-directive supportive therapy are similar to CBT and interpersonal psychotherapy for depression,[179] although CBT is superior for patients with psychosis.[180]

Experts in clinical psychology trials, such as Alan Kazdin, have formulated models to guide the type of trial needed to address the type of question asked. In part, design solutions will depend on the specific research question. For example, if the pragmatic question is whether an intervention works better than the current provision, then a two-arm trial design would allow the comparison of the new intervention with a specified and defined treatment as usual that follows best practice—eg, CBT plus monthly engagement and monitoring of the participant's daily difficulties compared with monthly engagement and monitoring alone.[181] If the question is whether one form of psychotherapy is better than another, then a direct comparison might be required.

However, if the question is why a treatment works, or whether a specific element is necessary, then the comparator treatment should be a therapy that controls for specified factors (eg, human contact) but in which the active ingredient has been removed. Findings from meta-analyses suggest that wait-list controls should be avoided, since they can lead to inflated effects sizes for the experimental treatment, possibly because people abandon their efforts to solve their mental health problems or recover independently because they are waiting for therapy.[48]

Outcome measures

Most trial methodologists would recommend a single primary outcome and a single prespecified timepoint at which this main outcome should be measured (eg, total symptoms at final follow-up assessment). This method can sit uncomfortably with basic aspects of psychological assessment—eg, the need for multiple assessments of a construct for validity, and multiple timepoints for reliability, as well as tracking the time course of the response. Having more than one primary outcome is justified in some situations (eg, in psychosis studies clinicians prefer psychiatric symptoms whereas service users tend to prefer social outcomes).[182] However, multiple primary outcomes require larger sample sizes. Additionally, the use of data obtained at multiple timepoints can give the most accurate estimate of treatment effects over the full follow-up duration. This process can be done by specifying an analysis involving all available data for a particular measure, which might be preferable to anchoring judgements regarding efficacy to a single assessment timepoint.

The most important outcome can be a subject of debate. Clinicians often prefer clinical outcomes (eg, psychiatric symptoms) whereas service users might prefer social outcomes (eg, recovery, social functioning, and quality of life).[182] Consensus regarding outcome measures for a specific condition would enable individual participant data meta-analyses,[183–185] which could hopefully provide information about the moderators and mediators of the treatment response. Integration with and adoption of routinely collected service user outcome data would also facilitate understanding of mediators and moderators. As part of a UK initiative that aims to establish agreement about sets of core outcomes for particular health conditions (Core Outcome Measures in Effectiveness Trials [COMET]), work is underway to establish consensus about a set of core outcomes for assessments of interventions for people with psychosis.[186] Regarding reporting outcome measures, it is unclear whether having a detailed interviewer-administered rating scale, which could provide rich data and be more engaging for participants than self-administered rating, is preferable to a self-report measure, which could be more reliable since inter-rater reliability is not needed across sites and staff and it avoids rater bias. A combination of both approaches could be a reasonable solution that maximises the benefits of both, so long as they are clearly prespecified as dual primary outcomes. If a trial with dual primary outcomes shows consistency across these outcomes, then the confidence in the findings would be increased.

Another important consideration when selecting outcomes is the time required to complete all assessments. Psychological therapy trials often include numerous secondary outcome measures, which might be of substantial interest. However, a large assessment burden on participants is more likely to impair retention in the trial, subsequently resulting in missing outcome data and reducing the internal validity of the trial. Limiting the number of outcome measures is likely to minimise attrition, but it restricts opportunities for understanding the processes of change. Similarly, agreement on the frequency of assessments and the length of follow-up would facilitate the pooling of data and the capacity for comparisons across trials. A compromise usually needs to be made between collecting meaningful data that will permit identification of what approaches work for whom across a broad range of outcomes and that facilitate mediation and moderation analyses, and not jeopardising participant retention. The involvement of service users who would be eligible for trial participation in the design of the trial, and ensuring pilot and feasibility work has been done, are both likely to be useful strategies in achieving a balance between these factors. Another possibility for minimising the assessment burden and maximising ecological validity and multiple measurements of outcomes is by use of experience sampling methods or ecological momentary assessment data as outcomes. This approach would allow reporting of symptoms, emotions, and indicators of functioning (eg, use of time in daily life—how many hours are spent engaged in constructive activity such as employment, education, parenting, housework, and leisue) as primary outcomes in clinical trials (see Part 8).

In addition to the measurement of wanted effects, such as improvements in symptoms or quality of life, measuring unwanted effects and reporting serious adverse events to ethics committees are important to safety monitoring. Historically, trials of psychological therapies have been poor at both monitoring hypothesised side effects and deterioration and reporting serious adverse events.[169] Several trials[181,187] of CBT for psychosis have attempted to measure adverse effects via qualitative and quantitative approaches. Some critics have suggested an association between CBT for psychosis and increasing stigma, encouraging deterioration or destabilisation, leading to serious adverse events such as admissions to hospital. However, these trials[181,187] also showed the opposite effect when compared with control conditions. This result is surprising when the inbuilt detection bias inherent in the design and implementation of these studies is taken into account (ie, therapists might have

For **COMET** website see http://www.comet-initiative.org/about/overview

weekly contact with a participant, whereas raters might only have contact at baseline, end of treatment, and follow-up, which clearly reduces the likelihood of detection of serious adverse events).

Public and patient involvement

Public and patient involvement is another area that can help to improve how psychological therapy trials are run.[188,189] People with mental disorders can provide unique insights into clinical trials, including identification of the most important and relevant research questions and thus outcome measures. For example, a definitive trial comparing CBT with antipsychotic medication would need to decide whether the most important question is one of superiority (ie, is combination treatment superior to monotherapy), equivalence (which would enable choice), or non-inferiority (in which case choice might depend on adverse-effect profiles). The assessment of acceptability of psychological therapies and the exploration of potential adverse effects can be informed by embedded qualitative interviews and analyses that can be led by service users (eg, the COMPARE trial is incorporating such a study). Finally, the involvement of service users as staff and, ideally, coapplicants, and investigators, should ensure meaningful participation in all phases of the design of the trial, running the trial, and reporting (eg, COMPARE has two service users as co-investigators and two as grant holders).

Public and patient involvement can be via consultancy groups (which is the case for the COMPARE trial), via priority setting partnerships that identify and prioritise the top ten unanswered questions (the James Lind Alliance facilitate the development of such partnerships in the UK), which has been done for the treatment uncertainties related to a diagnosis of schizophrenia,[190] or by the use of Delphi methods to establish consensus on topics with experts with experience (the COMPARE trial is also informed by Delphi studies of people with psychosis for both defining recovery[191] and identifying treatment priorities and preferences[192]).

For the **James Lind Alliance** website see http://www.jla.nihr.ac.uk/

Mechanisms and mediators of change

Trial design should also attempt to facilitate the identification of potential mechanisms, mediators of change (see Part 1), and moderators of treatment effects to inform on how a treatment works, what components are necessary and sufficient, and what treatments work for whom. The identification of mechanisms could be built into all clinical trials, which would also allow pooling of data, although this pooling would require consensus among researchers about the instruments that should be included in the trials. When a specific research question involves testing a mechanism, the trial must have sufficient statistical power for the mechanistic hypotheses and any between-group predictions.

The identification of mediators and moderators requires considerable thought at the planning stage to ensure that the appropriate factors are measured at the appropriate timepoints. The development of new statistical methods for the analysis of mediation and moderation should help with the accurate identification of mechanisms of change and mediators of treatment outcome. Traditional approaches to mediation analysis[193] assume the absence of confounding due to an unmeasured variable being responsible for changes in both the mechanism and outcome. These approaches are problematic because the assumptions made are unrealistic in many instances, especially given the complexity of potential influences on mental health. Subsequent developments that might be better suited to mediation analysis include attempting to measure and adjust for all important confounders,[194] or attempting to adjust effectively for unmeasured confounders (hidden confounding) by use of instrumental variable-based methods analysed on the basis of principal stratification.[195] Examples that are relevant to CBT for psychosis include the finding that participants with a psychosocial causal explanation of their difficulties could be more likely to engage with and benefit from CBT than those with a biological explanation,[196] and that participants with a good therapeutic alliance with their therapist are likely to benefit from a high number of CBT sessions, whereas participants with a poor alliance might be more likely to be harmed as the number of sessions increase.[197]

Innovation in trial design and methodology

The wider context of an individual trial should be considered. The reliability and validity of the findings from meta-analyses that are used to inform policy, guidelines, and service recommendations are largely dependent upon the quality of the trials that are included and the suitability of the selection criteria (ie, whether the included trials were designed to answer equivalent questions). Designing high-quality trials with a long-term perspective provides an opportunity to improve such meta-analyses. Collaboration between research groups, investigators, and methodologists with regard to future pooling of data could be facilitated by establishing collective research groups that would be recognised by group authorship, which would incentivise such involvement and cooperation.

Sometimes, alternative approaches to the traditional two-arm randomised controlled trial are needed, such as multiarm multistage trials.[198]

New methodologies, including adaptive designs, preference trials, and sequential multiple assignment randomised trials (SMARTs), will permit better generalisability to routine practice and more ethical and efficient trial conduct than traditional approaches. For example, a SMART that permits investigators to re-randomise patients who do not respond to CBT or antipsychotic medication after a relatively short period of time into the other monotherapy group or the combination group would confer future clinical

advantages—eg, arriving at a suitable treatment for an individual faster than with traditional trial designs. A preference trial would maximise recruitment in a field in which both service users and clinicians can have strong treatment preferences and opinions about psychological therapy or medication that could jeopardise recruitment, generalisability, or adherence to allocation in a standard randomised controlled trial. An adaptive design with a planned and prespecified interim analysis could permit the early abandonment of a treatment group that proved to be inferior. The cohort multiple randomised controlled trial design[199] allows several randomised controlled trials to be done simultaneously within a large patient cohort. For each randomised trial, all people who are eligible in the cohort are identified, then some are randomly selected to be offered the experimental intervention. The outcomes in the randomly selected participants are compared with the outcomes in those who were eligible but not selected (ie, receiving standard care or treatment as usual). Such designs could overcome recruitment difficulties and increase statistical power, efficiency, representativeness of samples, and comparability between trials, as well as increasing knowledge about the natural course of mental disorders and the likelihood of collecting data on long-term outcomes. This approach would be ideally suited to mental disorders that are seen within specialist teams (eg, eating disorders or first-episode psychosis), especially when the teams are linked within a national or international network and routinely monitor outcomes in a standardised way.

Improvements in the detection of patients who can be classified as responders and non-responders could be achieved by the selection of appropriate measures, incorporation of experience sampling or momentary assessment in the early phases of a trial (see Part 5), use of improved inclusion and exclusion criteria, and the development of statistical methods for mediation, moderation, and consideration of individual response trajectories rather than aggregate effects.

Notably, researchers should recognise that identifying successful interventions is not just about randomised trials, and clinical trials should complement other types of research questions and evidence. For example, randomised trials need to include embedded qualitative studies to obtain rich data alongside quantitative outcomes to inform understanding of active treatment processes and generate new hypotheses that can be tested empirically. The COMPARE trial involves interviewing participants about their experiences of both CBT and medication, focusing on acceptability, credibility, and wanted and unwanted effects (these interviews are designed, completed, and analysed by researchers with lived experience of psychosis). The results of these interviews have the potential to inform the design of a definitive trial related to the selection and recruitment of participants, inclusion and exclusion criteria, outcome measures, and treatment protocols.

If all of the above improvements can be achieved, the ability of researchers to identify and answer the most important questions will improve, trials will be run with greater reliability and validity, and confidence in and acceptance of the findings of these trials will increase (panel 13). Meaningful involvement of service users and carers will allow the identification of appropriate research questions and methods, ensure the relevance of outcomes (including adverse effects), and improve the retention of participants. Additionally, creation of large-scale datasets will enhance the credibility of the results of clinical trials, either by consensus regarding design considerations and measures that enable pooling of data, developments in individual participant data meta-analyses, or by use of routinely collected service data. Psychological treatment trials should also benefit from advances in trials in other areas of medicine.

Part 7: Training—can we cultivate a vision for interdisciplinary training across mental health sciences to improve psychological treatments?

Introduction

In this section, we discuss why the field of mental health science should endeavour to improve links between

Panel 13: **Directions and priorities for future research in clinical trials of psychological treatments**

- Establish a consensus among stakeholders (ie, the innovators and developers of psychological treatments, service users, and methodologists) regarding outcome measures, appropriate scheduling of assessments, and the length of follow-ups
- Routinely build into the design of clinical trials the ability to analyse for mechanisms of treatment
- Engage with commissioners and providers of psychological services to maximise the likelihood that such services can facilitate the routine collection of data to contribute to the evidence base and include clinical trials as part of service delivery when uncertainty exists
- Ensure quality trial design and valid, reliable analysis of data by routine and early engagement with clinical trials units, registration for all trials (including production of prespecified statistical analysis plans), and ensure that data analysis adheres to plans and is done by independent specialists in trial statistics
- Involve service users in all aspects of trial design and conduct, from decisions regarding research questions and methods, through to involvement in trial management and governance, research administration, and interpretation and dissemination of findings
- Carefully match comparators to the specific research questions that trials are seeking to answer
- Measure unwanted and wanted effects and arrive at a consensus about how to measure and report adverse effects
- Increase the use of innovative trial designs that maximise value for money, value for participant input, and reflect clinical practice; such designs include adaptive trials, multiple trials within cohorts, SMARTs, and preference trials; different designs will be suited to different research questions and clinical contexts
- Encourage career paths for those focused on advancing methods in the methodology of psychological treatment trial design, statistics, and other areas that will aid in future research

clinical psychology, psychiatry, and basic research training, and make some proposals about how this aim might be achieved. We review some early successes in innovation in psychological treatments in which basic researchers and clinicians have worked together, and discuss the reasons that such productive interaction has decreased in the past several decades. We offer some recommendations to bridge the gap between clinical practice and basic research into psychological interventions.

Historical shifts in interdisciplinary training

In 1949, in Colorado, the American Psychological Association held the Boulder Conference on Graduate Education in Clinical Psychology to agree on a standard model for clinical psychology training in the USA. Heavily influenced by the ideas of David Shakow, the conference adopted a scientist–practitioner training framework that encouraged clinical psychologists to use scientific research to inform their practice.[200] This proposal facilitated the development of effective new psychological interventions, which was catalysed by clinicians who did basic research, and basic researchers who understood the principles of psychological treatments (see appendix). This confluence of expertise resulted in crucial insights into the mechanisms of onset, maintenance, and treatment of symptoms of mental disorders, and, in some cases, completely revolutionised the psychological treatments available.

By taking a scientist–practitioner approach, training in psychological treatment becomes far more than just learning how to deliver a treatment described in a manual. Understanding the principles on which a treatment was derived can help the practitioner to deliver the treatment well and adapt the treatment to a given situation or patient. An example of a situation in which basic training was important was the development of various types of exposure therapy (incorporating response prevention) for anxiety disorders, including phobias, PTSD, and obsessive compulsive disorder. This treatment was initially derived from research on fear extinction in rodents, which showed a reduction in Pavlovian responses to negatively conditioned stimuli when the aversive outcome was omitted (see Part 1).[201,202] Notably, the focus on response prevention—ie, encouraging patients with anxiety not to engage in their usual coping strategies when confronted with an anxiety-provoking stimulus (eg, avoidance for phobias, rituals for obsessive compulsive disorder)—came from the insight that these behaviours can maintain the conditioned association through preventing extinction.[203] This approach might seem counterintuitive to the patient because, acutely, the prevention of coping behaviours increases their anxiety in the short term, but leads to a reduction in anxiety in the long term. Since this approach can also be counterintuitive from the perspective of some other therapeutic approaches, understanding the principles behind exposure techniques is important. Another example of practitioners benefiting from understanding the underlying science via their training is in the context of depression—namely, the influential learned helplessness model,[204] and its later modifications associated with hopelessness.[205] The learned helplessness model originated from the finding that animals that were exposed to inescapable aversive stimuli subsequently failed to escape when they had the option to do so.[206] Learned helplessness theory has made notable contributions to the understanding of risk factors for depression, especially associated with the roles of attributional style and perceived controllability.[207] Moreover, this theory has inspired numerous animal models that remain the mainstay of testing procedures for new antidepressant drugs in preclinical research, and translational research in this field has yielded valuable insights into the basic cognitive and brain changes that underlie depressive symptoms and their response to treatment.[208]

Over the past several decades, the links between basic research, clinical psychology, and psychiatry have become weaker, the reasons for which could be numerous. One simple fact is that because of the rapid expansion of psychology, basic researchers and practitioners rarely work in the same building. This distance reduces opportunities for interaction and the sharing of ideas between researchers and practitioners. Another important issue is that basic researchers and clinical psychologists often do not read the same journals, or even attend the same conferences, meaning that opportunities for interaction are few.[2]

Renewing the links between basic research and psychological treatments

Clinicians providing psychological treatments need training in basic research

In most countries, little teaching of contemporary basic research (eg, experimental psychology, neuroscience, genetics, physiology, pharmacology, data science, social science, economics) is incorporated into the clinical syllabuses of clinical psychology or psychiatry, or of allied professional training in the treatment of mental disorders. Canada and the USA are notable exceptions, since many clinical psychologists in these countries complete a doctoral training programme lasting at least 5 years, which includes substantial teaching in basic research together with an extensive research-based thesis and clinical training. The basic science content of training courses for psychiatry trainees in the USA has been emphasised,[209] although professionals within the field recognise that further training in basic science would be desirable.[210,211] Other than these examples, the basic research content included in clinical psychology programmes is small, even at the doctoral level (eg, PsyD in Canada and the USA, which is completed by approximately half of all qualified clinical psychologists in these countries; DClinPsy in the UK). In other countries, in which a master's degree is the standard educational

qualification required to become a clinical psychologist (including most of the European Union, Australia, New Zealand, and South Africa), very little basic research is in the curriculum.

This paucity of basic research content in clinical psychology programmes raises a serious concern about the training of clinical mental health researchers of the future and the risk that they will not be equipped with the tools to understand, critically assess, and use basic research that might be relevant to the development of new treatments or preventive strategies. Psychological interventions might become stuck in the past—relying on outdated models that are not supported by contemporary research or theory. This disconnect between basic researchers and clinical psychologists hinders innovation, and slows the emergence of effective and truly novel psychological treatments. Unless clinical psychologists and psychiatrists have the skills to assess research on both risk factors (eg, genetic and socioeconomic influences) and proximal mechanisms (eg, cognitive and neural processing of information), improving preventive strategies and treatments will be difficult.

Basic researchers need training in clinical conditions and psychological treatments

Although most basic researchers are enthusiastic that their research might contribute to improved treatments for mental disorders, they tend to have only a vague idea of what standard psychological interventions entail, since clinical practice is not generally taught even in undergraduate psychology degrees. Specifically, many basic researchers have little knowledge of the evidence base that supports standard psychological treatments, and have little opportunity to interact with clinical psychologists, see therapy in action, or find out what the common techniques comprise. Indeed, in our experience, the view that psychological treatments are primarily given in the context of an antiempirical psychoanalytical couch tradition, and that they are not derived from solid scientific theory or supported by robust evidence from clinical trials, is worryingly prevalent among basic researchers.[2] To formulate relevant research questions, basic researchers who are interested in contributing to the development of psychological treatments need to understand what the symptoms of mental disorders are (and are not), what the most common evidence-based psychological interventions entail and how theoretical models guided their development, and what the key questions are that need to be solved in the future.

The future of interdisciplinary training

Training clinicians in basic research

How can we ensure that the next generation of research leaders, both clinical and basic, are able to bridge the growing divide between their fields? One priority is to provide extra opportunities for academic training to trainees and qualified practitioners, and to attract those with a strong aptitude for research. In the UK, although competition for places on professional doctoral courses in clinical psychology is intense, and they recruit students who are highly academically able, very few graduates subsequently have a career in clinical research. Funding opportunities for the academic training of qualified clinical psychologists are highly competitive. That said, some major UK research funding bodies, such as the National Institute for Health Research (NIHR) and the Medical Research Council, offer academic training pathways for clinicians. These training pathways offer clinically qualified, non-medical health-care professionals the chance to undertake a PhD, while covering a clinical-level salary, tuition, travel, training costs, and research consumables. This training provides a valuable springboard for a career in clinical research, but there is scope for uptake by more clinical psychologists than at present, in part because they might not be aware of these opportunities or have sufficient support or research experience to develop a strong application. Another way of improving academic training in clinical psychology would be to create longer training programmes specifically for those trainees with a strong aptitude for research. These courses could be similar to the North American PhD model, providing students with sufficient time to complete an extensive research project and teaching relevant scientific material alongside clinical skills. The Psychological Clinical Science Accreditation System model that has been developed in the USA, which emphasises the science of clinical psychology in training and internships, would also be an effective way of increasing opportunities for research training. A similar training model is offered at The University of New South Wales (UNSW), Sydney, Australia, in which students are enrolled in a clinical training programme and a PhD programme concurrently, and they are awarded both degrees at the conclusion of their course (eg, Master of Psychology [Clinical] and PhD).

Training pathways also need to be developed for mental health researchers that cultivate an interdisciplinary approach both between clinical psychology and psychiatry, and between disciplines of clinical mental health and a variety of relevant basic research. One possible way to achieve this interdisciplinary approach would be to encourage clinical psychologists to undertake internships or placements in basic-research settings across a range of relevant disciplines, from economics and social science, to neuroscience and genetics. Psychiatrists in the UK already have such an opportunity through the NIHR Academic Clinical Fellowships scheme, but no equivalent programmes seem to be available for clinical psychologists, in either the UK or other European countries. Multiskilled clinical academics, trained in an interdisciplinary environment, would have the advantage of being able to speak the languages of both clinical and basic research. They would also be best placed to develop the metaprofessional skills needed to do truly interdisciplinary

For more on the **UK National Institute for Health Research fellowship for research** see https://www.nihr.ac.uk/funding-and-support/funding-for-training-and-career-development/training-programmes/nihr-hee-ica-programme/nihr-hee-ica-programme-cdrf.htm

For more on the **UK Medical Research Council Clinical Research Training Fellowship** see http://www.mrc.ac.uk/skills-careers/fellowships/clinical-fellowships/clinical-research-training-fellowship-crtf/

translational research, and to use the knowledge derived from basic research to drive innovation in the development of psychological treatments.

Training basic researchers in psychological interventions

Basic researchers with an interest in understanding and contributing to the development of new psychological treatments need to be provided with the opportunities to do so. In the same way that a first-year neuroscience PhD student might learn about the principles and practice of neuroimaging analysis, and therefore be able to assess neuroimaging evidence more effectively because they understand the potential pitfalls (even though they might never use the technique), basic researchers need a route through which they can learn about what psychological treatments are used in practice and how they are hypothesised to work. This knowledge would provide a new generation of researchers who understand the basic principles underlying psychological interventions and could bring a fresh perspective on driving innovation. Even sitting in the same lectures and tutorials as clinical trainees would increase the opportunities for meaningful interaction, and encourage clinical and non-clinical students to value input from each other when developing collaborations. Although neuroscience and cognitive or experimental psychology students are obvious candidates for such an approach, students with backgrounds in a whole range of disciplines—from social science and economics, to computer science and mathematics, and molecular biology and genetics—might have an interest in psychological interventions and could contribute important ideas.

A culture change is needed to accept more crossover

To address these problems that are hindering interdisciplinary interaction several obstacles will need to be overcome, which will require bold changes in thinking within the health-care system. These obstacles exist for both clinical accreditation and funding. A huge number of mental health practitioners have research talents that are being underutilised, and perverse disincentives often discourage clinicians from entering academia, including a possible reduction in salary and a perception that research will not help in their career progression. Additionally, the procedures for obtaining funding for a research doctorate are not widely understood among trainees, and the opportunities to gain the research experience that would contribute to a competitive application are sporadic and invariably depend on locally available supervisors; therefore, the trainees with the most research potential might be overlooked. Furthermore, unlike for clinical training (at least in the UK), an absence of national recruitment is apparent for research training in clinical psychology.

These obstacles could be addressed through longer, targeted clinical academic programmes (like the PhD programme in North America) that include a substantial research component in the professional doctorate, alongside standard clinical training, and national recruitment to attract trainees with the greatest research potential. More substantial research projects than are completed nowadays in most clinical psychology courses would also help to address the concern that learning about techniques could be forgotten if they are not put into practice. Many European training programmes for clinical psychology successfully blend clinical training with basic research; however, the courses are at a master's level, and so do not have the requirement of a doctoral-level thesis, and therefore trainees do not receive the same quality of research training as those in the North American PhD model. For example, in the past decade a pioneering model for training clinical psychologists has been adopted by the Karolinska Institutet in Stockholm, Sweden. In this model, teaching is based within the Division of Psychology in the Department of Clinical Neuroscience, and within a medical university. This design has resulted in the students being exposed to both psychology and neuroscience, and encouraged awareness of the rich links between clinical psychology, neuroscience, psychiatry, and physical medicine. Almost all of the instructors are involved in research, and the majority have at least 50% of their time devoted to research. Although only a master's level qualification is required to become a clinical psychologist in Sweden, Karolinska students are poised as members of the new scientist–practitioner generation. The development of similar programmes elsewhere would be a positive step toward interdisciplinary training, as would an examination of the outcomes of different international models. To our knowledge, such an investigation has not been done to date, but would be extremely valuable.

Models of shared research supervision

Another major factor that restricts access to interdisciplinary training is that those trainees who do enter research training are often supervised only by clinicians, rather than by basic researchers. As discussed, this separation between clinical training and basic research affects both fields with very few opportunities available for trainees in basic research who are keen to understand psychological treatments, to find out what they entail, and the diverse approaches that they adopt. Such exposure to ideas, and understanding of how psychological interventions are actually administered, is an important first step for basic researchers to start to formulate valuable research questions. Therefore, allowing basic researchers to have an active part in the supervision of research projects of clinical psychology trainees would be desirable when possible, and vice versa. Encouraging joint doctoral supervision (whether for research or clinical students) between principal investigators within basic research and clinical

psychology would be a simple and valuable step in the right direction in this regard. Returning to the Australian example, at UNSW Sydney, students who are studying for a combined clinical and PhD degree often do their PhD research under the supervision of a basic researcher (eg, behavioural neuroscientists) and test questions with clear clinical relevance (eg, on topics such as fear extinction, and drug addiction), alongside their clinical training programme. Such a model of supervision facilitates a broad training experience and a unique opportunity for mentorship from both clinical supervisors and basic researchers.

Mixing and mingling—the role of conferences

Finally, even among those clinical psychologists who do enter academia, few forums exist for exchanging ideas with researchers from other disciplines, since the journals they read and the conferences they attend are typically discipline specific (with some notable exceptions—eg, the MQ: Transforming Mental Health annual science meeting; the meeting on neuroscientific research into psychological treatments arranged by the European College of Neuropsychopharmacology;[212] and the annual meeting of the German Association for Psychiatry, Psychotherapy and Psychosomatics). Some clinical psychologists and neuroscience researchers have started to work together to produce new ideas for intervention. A good example is the adoption of ideas from the literature on the neuroscience of reconsolidation—the modification of old memories during their reactivation—in the formulation of new treatment approaches for PTSD.[213] Several studies have tested the possibility that reactivated memories could be disrupted through pharmacological intervention with propranolol,[214,215] with some preliminary indications of positive effects. Other studies[65,216,217] have tested whether the reconsolidation of established memories can be disrupted by use of simple psychological interventions based on cognitive science, with promising results. Engagement with a simple visuospatial task (the computer game Tetris) following memory reactivation was shown to substantially reduce subsequent intrusive memories of experimental trauma.[65] Although this line of research requires considerable further work to show robust clinical efficacy (see Part 6),[216,217] it is an intriguing example of the type of interdisciplinary innovation between basic and clinical research that holds promise for improved treatments in the future. Other good examples of interdisciplinary innovations have been found in the development of new psychological interventions for anhedonia (panel 14).

In the 1950s and 1960s, the development of new psychological interventions transformed the treatment of mental disorders, with the creation of effective treatments on the basis of novel, empirically testable models. Inspired by ideas that were drawn from cognitive psychology and behavioural neuroscience, interventions that were developed through collaborations between previous generations of basic researchers and clinicians have become the treatments of choice. Despite these successes, improvements in treatments are still needed since patient responses to psychological interventions are highly variable. However, in the past few decades the productive interaction between those who deliver psychological interventions and basic researchers has waned. The gap between these

For the **German Association for Psychiatry, Psychotherapy and Psychosomatics** website see https://www.dgppn.de/

***Panel 14:* Could understanding reward processing in the brain help in the development of new treatments for anhedonia?**

Over the past decade, interest has been renewed in a core symptom of depression, anhedonia, which is the loss of interest or pleasure in previously enjoyable activities; anhedonia is also an important component of many other mental disorders, including schizophrenia and addiction, as well as a prominent symptom in neurological disorders, such as Parkinson's disease.

In depression, anhedonia is associated with a more severe course of illness and poorer response to standard antidepressant drugs[218] and psychological treatments[15] than depression without anhedonia; clinicians appreciate that this symptom is an area in which treatments are inadequate.

Given that anhedonia is intrinsically related to an absence of motivation and hedonic response, researchers have proposed that this symptom could arise because of a disruption of the brain's reward circuits,[219] which have been characterised in extensive detail by neuroscience research over the past 30 years.

This idea is not new; in the 1970s Jeffrey Gray first proposed that symptoms of depression might be explained by changes in a behavioural activation system and a behavioural inhibition system,[220] although most researchers focused on the behavioural inhibition system and its association with neuroticism.

An important conceptual advance in this theory has been the notion that the reward system (the behavioural activation system) comprises several relevant cognitive processes: hedonic response to reward delivery, valuation of rewards, reward learning, propensity to exert effort, and decision making; these components at least partially dissociate, and are linked with activation in different brain circuits and neurochemical systems.[221]

This knowledge from neuroscience research has been exploited by clinical psychologists seeking to develop treatments specifically targeted at anhedonia—eg, positive affect treatment;[221] this treatment builds on behavioural activation therapy and positive event scheduling, which are both effective treatments for depression[222] that were originally motivated by ideas derived from behaviourism,[45] and that are known to increase responsivity in the brain's reward system.[223]

Drawing on the finding that reward processing comprises a diverse set of processes, the aim of positive affective treatment is to increase engagement in, attention to, and anticipation of enjoyable activities.[17]

From a complementary angle, another novel approach based on cognitive science (ie, the processes of mental imagery and interpretation bias) has been via positive imagery training; in trials with individuals with depression, post-hoc analyses show early indication of an effect on anhedonia;[224,225] this type of focussed approach could be developed into the wider package of positive affective training.

Although these novel interventions require further assessment, specifically in groups of individuals with anhedonia and depression, the research so far provides examples of how scientific discoveries are of use to fuel development of innovative psychological interventions.

Panel 15: **Example directions for the future of training and links between clinical and basic science**

- Opportunities for integrated clinical and academic training in psychology, through extended programmes that are targeted at those clinicians with the greatest research potential
- Training for basic researchers in psychological treatments, including hands-on experience of techniques and interactions with clinicians, so that they can formulate research questions that are relevant to psychological interventions
- An expectation of interdisciplinary research for psychological treatment researchers, including cosupervision of the research component of professional qualifications by clinical and non-clinical principal investigators
- The provision of seminars on the next steps, focused on academic training as a standard part of programmes for clinical training in mental disorders
- Improved dissemination of research internship and doctoral funding opportunities for clinical psychologists, such as that provided by the Society for a Science of Clinical Psychology
- Training programmes in which trainees in clinical psychology, psychiatry, and basic research can learn alongside each other
- High-level interdisciplinary meetings between basic researchers, clinical psychologists, psychiatrists, and others, including forums in which practitioners can propose questions that they think are important to basic scientists; with tangible outcomes such as papers, grant applications, and implementation work
- Use of the continuing professional development framework to enhance the understanding of basic science among psychological treatment practitioners

For **Society for a Science of Clinical Psychology** website see http://www.sscpweb.org/

disciplines impedes innovation in the development of new psychological treatments, both because basic researchers do not understand what psychological interventions entail, and because clinicians are not familiar with relevant advances. In this section, we have outlined a number of proposals for how to bridge this gap; these proposals should promote a much more extensive interdisciplinary interaction and dialogue than exists nowadays (panel 15).

Part 8: Whom should we treat, for what, and with what? Embracing the complexity of mental disorders from personalised models to universal approaches

Introduction

Most theoretical models and evidence-based psychological treatments have typically been designed for specific, categorically defined mental disorders—eg, major depressive disorder, social phobia, or PTSD. Leading clinical guidelines recommend specific treatments for each mental disorder, usually categorically defined by symptomatology.[226,227] However, mental disorders are more complex than these guidelines take into account, and are characterised by huge varieties between individuals with a given disorder.Heterogeneity in symptomatology across mental disorders is very common,[228] and many individuals have more than one mental disorder.[229,230] Additionally, many individuals have subsyndromal symptoms of other disorders, and could have symptoms that shift between disorders over time. Mental health researchers—and those in psychological treatment research specifically—need to embrace the complexity of mental disorders to make progress in reducing the burden of these disabling conditions. The complexity of mental disorders is a challenge for research and clinical practice. Treatment solutions to deal with this complexity include both highly individualised (ie, personalised) approaches, and so-called universal or transdiagnostic approaches that target common mechanisms. More studies are needed to examine whether these approaches improve the effectiveness of treatments for mental disorders.

Why are mental disorders so complex?

Unlike most areas of medicine, mental disorders are defined predominantly by their symptoms. A paucity of knowledge about the causes of mental disorders contributes to this approach. Symptoms are often considered as manifestations of an underlying latent factor (eg, sad mood and loss of interest are caused by an underlying major depressive disorder). However, these symptoms might not only serve as an output from so-called underlying processes, but could also mutually reinforce one another, as presumed by the network approach.[231] For example, in depression, insomnia might lead to concentration problems, which in turn might cause sadness and loss of pleasure, which in turn might lead to fatigue, feelings of guilt, and suicidal ideation, resulting in the full syndrome of major depressive disorder. Thus, whether these symptoms are indeed manifestations of an underlying factor is still uncertain.[231]

Mental disorders are dimensional, and yet most mental health researchers use a categorical model to study the effects of treatments. The Diagnostic and Statistical Manual of Mental Disorders-5 (DSM–5)[227] is a categorical nosology for classification, to identify, for instance, a depressive episode, and to study the effects of a disorder-specific treatment for depression, such as behavioural activation. In the past few years, initiatives have been taken—eg, the RDoC initiative[29]—to stimulate research on the dimensions of observable behaviour and neurobiological measures of mental disorders, instead of categorical diagnostic criteria (see Part 1).

An additional complicating factor is the differences between individuals and the specific characteristics of their psychopathology. Studies using network analyses have given new insights into the variation of psychopathology between patients.[228,232] These studies show that, although for some people—eg, those with a strongly connected network of symptomatology—the transition from feeling healthy to being fully depressed can be abrupt (categorical), for others—eg, individuals with a weakly connected network of symptoms—external stressors (such as not being able to pay rent) could lead to an increase in symptomatology; although these symptoms gradually decrease after the stressor is gone.[233] These differences in psychopathology could be explained

by a dimensional model of psychopathology—ie, that the individuals with strongly connected networks might be those with increased neuroticism. However, whether these differences between individuals can be explained by an underlying dimensional mechanism or categorical disorder remains unclear.

Mental disorders are complex to study because of the interplay between an individual's emotions, cognitions, physiology, and other factors, as well as how they interact with the environment, which can change over time or as a consequence of having a mental disorder (for the differentiation of mechanisms responsible for onset *vs* mechanisms that are responsible for maintenance of psychopathology see Part 1). For instance, for individuals with depression, major life events (eg, the death of a loved one) are consistent risk factors for the onset of the first episode, whereas for those who have had one or two previous depressive episodes, comparatively less stressful events (eg, getting a minor traffic ticket) are sufficient to trigger a subsequent depressive episode.[234] Huge differences have been found between individuals in how their emotions fluctuate—an important part of many mental disorders—and huge differences over time.[235]

Furthermore, at least 45% of people with mental disorders have more than one disorder (for definitions see appendix), while over half of people with a mental disorder have subsyndromal symptoms of other mental conditions.[229] The lifetime comorbidity of common mental disorders (ie, anxiety disorders with major depressive disorder) can be as high as 73%.[230] The Global Burden of Disease Study[236] estimated that comorbidities for mental disorders for 188 countries between 1990 and 2016 had risen substantially. Comorbid disorders are consistently associated with a greater demand for professional help, poorer prognosis, greater interference with everyday life, and a higher incidence of suicide than disorders without comorbidities.[237,238] An improved understanding of comorbid mental disorders is crucial to give insight into their causes and to improve psychological treatments for all mental disorders and other conditions.

Heterogeneity and comorbidity have been studied in some fields of mental health to explain the causes of mental disorders, including comorbid disorders.[239,240] Dimensional models have been proposed to explain the cause of comorbid disorders; most suggest shared factors for the concurrent disorders (eg, neuroticism),[241] and some add specific factors that differentiate among mental disorders.[242] For instance, the dimensional tri-level hierarchical model of anxiety and depression includes the following levels: a shared higher level factor for anxiety and depression (ie, general distress); two additional factors that are at an intermediate level in terms of specificity for anxiety and depression (ie, anxious misery; fears that explain covariation in positive affect, anhedonia, and sad mood; social fears and fears to explain covariation in social fears; and fears of specific stimuli and interoceptive sensations, and agoraphobic fears); and five further specific unique factors for depression and anxiety disorders (ie, depression, fears of specific stimuli, anxious arousal, social fears, and interoceptive or agoraphobic fears; figure 5).[243]

Alternatively, a network approach can be of use to explain comorbidities through spreading symptom activations. Comorbidities are hypothesised to result from direct associations between the symptoms of multiple disorders—ie, a symptom of one diagnostic category (eg, major depressive disorder) can evoke other symptoms that in turn evoke symptoms of another diagnostic category (eg, anxiety about several events, chronic anxiety or worry).[231] Thus, a comorbidity might be the result of shared symptoms across mental disorders, so-called bridge symptoms.

Figure 6 is an example of a dynamic network of symptoms of major depressive disorder that mutually reinforce other symptoms of the disorder and comorbid symptoms of generalised anxiety disorder.[228,231] For example, disturbed sleeping, which is a symptom of depression, could lead to fatigue, concentration problems, and irritability or agitation (bridge symptoms), as well as other specific generalised anxiety disorder symptomatology. The bridge symptoms are criteria of major depressive disorder and generalised anxiety disorder.[231,244] Additionally, between different individuals comorbidities can develop in different ways, resulting in many different paths to the comorbidity depending on the individual and their environment. However, the network approach does not explain why some individuals are more prone to developing comorbidities (ie, having more symptoms) than others.

Both the network model and the dimensional (hierarchical) model could contribute to the explanation of mental disorders, including comorbidities. These models emphasise the necessity of translating findings from group studies to specific individuals struggling with mental health problems. The role of symptoms, individual differences in symptoms and emotions, and potential underlying mechanisms as maintenance factors in mental disorders, are key elements that require further study.

Personalised models of mental disorders

Although some disorder-specific treatments have positive effects on comorbid disorders in addition to the specific presenting disorder (eg, CBT for specific anxiety disorders also reduces depressive symptomatology),[245] improvements in treatment outcomes are still needed for people with mental disorders, including those with comorbid mental disorders.

Research should embrace the complexity of mental disorders to make progress in psychological treatment research (panel 16). One way forward is to study both interindividual and intraindividual differences. An experience sampling method or ecological momentary assessment can be of use to develop personalised models of psychopathology.[246] The experience sampling method

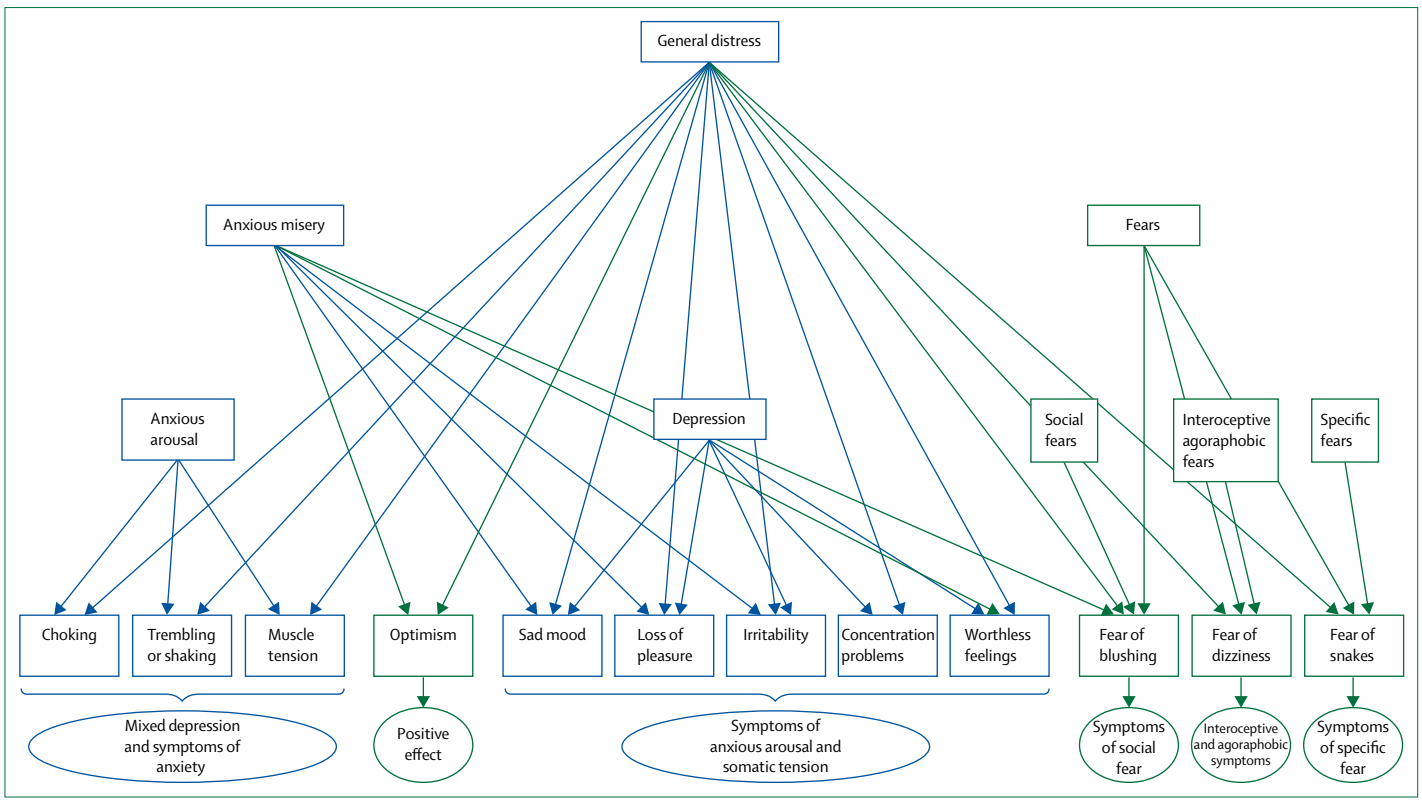


Figure 5: **Tri-level hierarchical model of the comorbidities associated with major depressive disorder and generalised anxiety disorder**
Blue and green boxes and lines show how factors and symptoms are associated with the major comorbidities. Adapted from Prenoveau et al,[243] with permission from Elsevier.

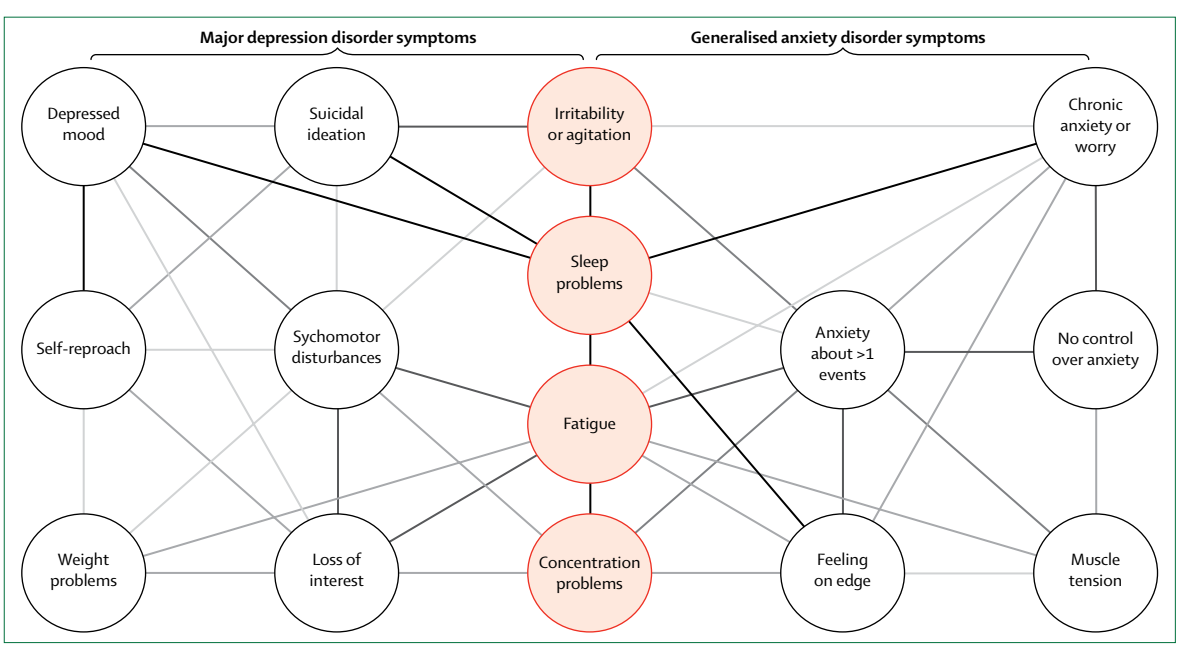


Figure 6: **Hypothetical dynamic network of the symptoms of major depressive disorder that mutually reinforce other symptoms of the disorder and comorbid generalised anxiety disorder symptoms**
Circles contain symptoms and the lines show the causal relationship between those symptoms. Darker lines indicate a stronger relationship between the symptoms. Red circles are bridge symptoms of major depressive disorder and generalised anxiety disorder. Adapted from findings in Borsboom et al[231] and Cramer et al.[228]

is a collection of research methods by which a service user reports on symptoms, affect, behaviour, and cognitions close to when they occurred in the service user's daily life—eg via an application on a mobile phone (see Part 5). Given that the experience sampling method can gather extensive data for each individual, individualised analyses can generate personalised models on the dynamics of each patient's network of psychopathology. Therefore, for instance, the centrality (or the strength) of a specific symptom or mechanism for a specific person can be defined—eg, a loss of interest might be a central symptom for one person with major depressive disorder, whereas the central symptom for another person with the same disorder could be sad mood.[246] This experience sampling method would offer new insights into mental disorders and personalised models of psychopathology. Systematic reviews have emphasised the value of the experience sampling method for assessing symptom fluctuations and interactions over time in anxiety disorders,[247] depressive disorders,[248] and substance use.[249] Studying the transient processes of emotions, cognitions, symptoms, and stress (and other relevant factors) in daily life can be done in prospective and experimental studies—eg, in a randomised controlled trial (see Part 6). In one study,[250] alongside a randomised trial of the effectiveness of three relapse-prevention treatments for depression, an ecological momentary assessment study was incorporated for a subset of patients who had remitted from recurrent depression. This momentary assessment study assessed the participants' emotions, cognitions, symptoms, and imagery-based processing ten times a day, 3 days a week, for 8 weeks, using the *Imagine your mood* application on a mobile phone.[250] Given these ecological momentary assessment studies involve self-reporting questionnaires, addition of physiological and behavioural measures might be useful for such investigations.

Personalised treatment approaches

Research on personalised models might disentangle the complexity of mental disorders, including comorbidities, and enable the optimisation of psychological treatments (appendix). The goal of the personalised medicine approach is to optimise the patient's response to treatment on the basis of their unique characteristics (ranging from genetic and neurobiological factors to symptoms) and underlying mechanisms (appendix). Ecological momentary assessment might improve insight into specific diagnoses[251,252] and offer valuable information that might improve matching treatments to patients. For instance, assessing daily fluctuations in positive and negative emotions by use of an experience sampling method for patients with depression predicts their response to treatment.[253] Assessing an individual's change in emotions (and other processes) over time as they are undergoing therapy might offer valuable empirical information on patterns and mechanisms of change during treatment.

An alternative route to improve the matching of patients to treatment is to use a machine-learning approach to identify the characteristics of an individual on the basis of group studies, which predict the patient's differential responses to existing treatments. An example of this technique is the calculation of a personalised advantage index score,[254] generated by comparing psychological treatments with pharmacological treatments for depression. Future studies should examine whether treatment matching can be improved for individuals with comorbid mental disorders. Similar approaches include clinical-risk scoring,[255] as is used in the field of medicine—eg, treatments for lung cancer are improved by molecular testing for targeted therapies that can overcome resistance to first-generation drugs.[256] Within the field of mental disorders, further studies are needed to examine the relevant variables of these index scores to optimise treatment matching and incorporate, for instance, machine learning.

Additionally, as discussed in Part 1, research on the mechanisms of psychological treatments might reveal crucial moderators of treatment outcomes that lead to better matching of patients to treatment, such as cognitive and biological markers.

Apart from enhancing treatment matching, feedback to the clinician and the patient on daily fluctuations might be of use to adapt treatment and thereby improve the treatment outcomes. Feedback on daily fluctuations via momentary assessment might enable clinicians to adapt interventions immediately—ie, within the session—by

Panel 16: Potential directions for future research regarding the complexities of mental disorders

- Embrace the complexity of mental disorders, including comorbidities, by studying interindividual and intraindividual differences in daily life, and investigate individual dynamics of emotions, cognitions, symptoms, and stress (and other relevant mechanisms) in prospective studies, and in experimental studies, such as randomised controlled trials
- Study models that explain comorbidities in mental disorders and treatment approaches for comorbid disorders
- Investigate whether psychopathological models can be personalised to the extent that treatments can be adjusted, and thereby improve treatment outcomes
- Investigate which patients should be treated, and with what; a disorder-specific treatment, a personalised treatment, or a transdiagnostic or universal treatment, or a combination of these approaches
- Examine the effects of transdiagnostic or universal treatments for several mental disorders, including the comorbid conditions, in comparison with evidence-based disorder-specific treatments

giving real-time feedback on progress to the clinician and the patient.[234] A randomised controlled trial[257] of 102 patients with depression showed that the efficacy of pharmacological treatment could be enhanced by the addition of feedback to the clinician and patient on the personalised patterns of positive affect via an experience sampling method. The collection of data from ecological momentary assessments, with comparable assessments within clinical settings on a patient's patterns of daily fluctuation of change over time while undergoing treatment, would be of great value in a large population with mental disorders (including outcomes after treatment; see Part 6). Mobile devices and applications could increasingly be of use for personalised and immediate interventions. In the future, researchers could make empirical data available to clinicians and patients, which could help them to work together on improving treatment outcomes. Close collaboration will be needed with computer scientists and mathematicians, drawing on advances in these fields (eg, areas of complexity, dynamical systems, and handling big data). Further research is needed on the dynamics of symptom outcomes, rather than just static assessments—eg, time-series analysis of data on mood in patients with bipolar disorder.[258] For now, studies are needed to examine whether personalised treatments are indeed more effective than traditional treatments. A crucial question is, can psychopathological models be personalised to the extent that treatments can be adjusted for the individual, and thereby improve outcomes (see Part 6)?

One size fits all or a universal approach?

Most traditional disorder-specific psychological treatments contain a package of several interventions that target underlying mechanisms of psychopathology (see Part 1). Another, approach is to consider common features between mental disorders via a so-called universal approach (appendix; panel 9)—eg, adverse life events are consistent predictors for the onset of most mental disorders.[259] A risk factor—eg, stress sensitisation—might prove to be a valuable target for treatment, since changing sensitisation might also influence other symptoms in the network, such as rumination or sleeping problems.[260] Alternatively, changing stress sensitisation might reduce a latent factor (eg, neuroticism) and thereby reduce symptomatology. Research efforts could be focussed on trying to identify universal underlying mechanisms across numerous mental disorders, and targeting these mechanisms by universal interventions (panel 16; see Part 4). This transdiagnostic approach has begun to give very promising results—eg, in the treatment of eating disorders.[261,262]

Another example of a transdiagnostic approach to psychological treatment is Barlow's unified protocol for the treatment of emotional disorders.[263] This approach targets transdiagnostic mechanisms that are hypothesised to be responsible for the development and maintenance of psychopathology broadly, rather than addressing disorder-specific mechanisms or symptomatology (especially studied in patients with a principal anxiety disorder). A more personalised approach is taken as part of this protocol than in most treatment protocols, including an assessment of how each patient's dysfunction is associated with the underlying mechanisms of their disorder. The patient's personal profile can then be used by a clinician to select additional interventions that are specific to the mechanisms underlying their symptomatology.[264] Further studies are needed that examine whether these unified approaches are indeed more effective than traditional disorder-specific treatments.

Finally, despite the apparent contrast between a personalised and a universal approach, we suggest that future research agendas embrace the complexity of mental disorders, including comorbidities, and consider both ends of the treatment spectrum—ie, examine approaches that could offer universal treatment and, if necessary, add disorder-specific interventions alongside personalised treatment solutions (panel 16). Solutions to the problems regarding the complexity of mental disorders need to consider both highly individualised approaches and universal or transdiagnostic approaches to target common mechanisms.

Part 9: Target: suicidal behaviour—protecting lives

Introduction

In this section, we discuss how many of the principles outlined earlier in the Commission could be applied to the development, assessment, and implementation of treatments to reduce suicidal behaviour. Although the causes of suicide and suicidal behaviour are complex, they are psychological at their core, since an individual who attempts suicide makes a decision to end their life. In the past 25 years, substantial advances have been made in understanding who is most at risk of death by suicide and what factors increase this risk in some individuals but not in others. Moving forward, the growing evidence base for psychological treatments can be built on to reduce the risk of suicidal behaviour. However, despite these advances, key gaps are apparent in the understanding of suicidal behaviour that require urgent attention. Addressing these gaps is an excellent opportunity to develop more effective treatments that can be replicated, are more precise than treatments to date, and can reach those who are most vulnerable irrespective of who they are or where they live.

Suicide and suicide attempts are the most tragic outcomes that result from an inability to effectively treat those with mental disorders. Suicide is a major public health concern, with at least 804 000 people dying by suicide globally each year.[265] Since suicidal behaviour is a transdiagnostic occurrence that is

associated with many mental disorders, we believe that it is an ideal test case of how the methods that have been discussed elsewhere in this Commission can be applied to a specific problem.

In addition to the personal tragedy associated with death by suicide, the economic cost of suicide is huge. For example, in countries in the European Union, the average lifetime cost associated with a suicide is approximately €2 million.[266] Although the science of suicide research is still relatively new compared with other mental health sciences, in the past few decades several welcome advances in the understanding, treatment, and prevention of suicidal behaviour have been made.[267] These advances include a better understanding of the common risk factors for suicidal behaviour,[268–271] evidence that some psychological treatments reduce suicidal ideation and behaviour,[272–279] and growing evidence that public health interventions are associated with reductions in suicide.[278,279] In this section, we discuss the advances that relate to psychological treatments for suicidal behaviour in more detail and identify a number of urgent calls to action (panel 17). We focus on psychological treatments, but clinicians and researchers should keep in mind how the principles outlined in this Commission can be applied to the primary prevention of suicide.

Although suicide most often occurs in the context of mental disorders,[280,281] the need to move beyond diagnostic categories to explain and treat suicidal behaviour is widely recognised,[282] as is the central role of psychological factors in the cause and course of suicidal behaviour.[271] Arguably, suicide is the cause of death that is most closely associated with psychological factors, given that an individual makes a decision to end their own life.[271] Despite advances in the knowledge of the risk factors associated with suicidal behaviour, the ability to predict who is most likely to die by suicide is poor because no markers of suicide risk are sufficiently specific—eg, although depression is the mental disorder most associated with suicide risk, less than 5% of people with depression die by suicide.[271,283]

New psychological models of suicide have been developed that have identified more proximal and specific markers for risk of suicide than previous models.[284–290] In addition to the theoretical importance of identifying proximal markers of the final common pathway to suicidal behaviour, proximal markers are crucially important for clinical practice and should be treatment targets. Specifically, constructs that are among the key predictors of suicide attempts include feelings of defeat, entrapment, not belonging, and being a burden, as well as future thinking, goal adjustment, reasons for living, and fearlessness of death;[271,286–288,291,292] therefore, these constructs should be targeted in psychological treatments and suicide prevention activities. To date, insufficient focus has been on these suicide-specific psychological proximal markers. Moreover, little is known about which factors are responsible for the observed effectiveness of approaches to suicide prevention (see Part 1). Trials of psychological treatments for suicidal behaviour should routinely assess theoretically derived mechanisms (both psychological and biological) that could explain the treatment effect. A concerted focus on potential biomarkers—eg, salivary cortisol or the serotonin metabolite 5-hydroxyindoleacetic acid—is also required, ideally tested in combination with other factors.[293,294]

Evidence for psychological treatments and their effect on suicidal ideation and behaviour

Psychological treatments reduce suicidal ideation and the frequency of suicide attempts,[272,274,295] although little evidence is available that such treatments have a marked effect on subsequent incidences of death by suicide.[296] Indeed, in 86 (50%) of the 172 WHO member states, between 2000 and 2012, the incidence of death by suicide either remained approximately the same, or increased by more than 10%.[265] Most people who die by suicide are not in contact with clinical services in the 12 months before death, so until the reach of psychological treatments can be expanded beyond those already in contact with clinical

***Panel 17*: Calls to action for research into psychological treatments for suicidal behaviour**

- More large-scale psychological treatment trials (including psychotherapeutic and brief-contact interventions) targeting suicidal ideation and behaviour are urgently required
- Establish whether psychological treatments work for different sociodemographic populations (eg, men *vs* women, adolescents *vs* older adults, individuals from different ethnic backgrounds) and in different settings (eg, primary or secondary care *vs* acute settings), patient groups (eg, treatment as an inpatient *vs* as an outpatient) and countries (eg, low-income and middle-income countries *vs* high-income countries)
- Rigorous investigation of those individuals at imminent risk of suicide
- Replication of psychological treatments by independent research groups
- Agree on common measures of core outcomes (ie, suicidal ideation and behaviour) and complete multicentre treatment studies and harness so-called big-data techniques to establish whether psychological treatments can prevent suicide
- Assess potential mechanisms derived from psychological theories that are hypothesised to account for treatment effects in all trials (risk and protective mechanisms) and moderators of the effects
- Use techniques derived from experimental psychopathology to establish whether hypothesised mechanisms account for changes in symptoms or wellbeing (see Part 1)
- Establish the active ingredients of psychological treatments, including the role of therapeutic alliance
- All psychological and social treatments trials (irrespective of whether suicidal ideation or behaviour is the target) should routinely include a measure of suicidal ideation or behaviour (even as a secondary outcome) that could be harvested in big-data analyses
- Ascertain the barriers to seeking treatment—particularly for males
- Investigate the extent to which new technologies might be of use to engage difficult to reach populations (eg, men, adolescents)
- Those with lived experience of suicidal behaviour (eg, bereaved by suicide or with personal experience) should be involved in all stages of psychological treatment research

services, these services are unlikely to have a direct effect on national suicide rates. Given the complexity of the risk factors for suicide, multilevel interventions offer the most promise.[279,297]

Nonetheless, meta-analyses show that CBT is effective in reducing suicidal behaviour in adults, although not in adolescents.[275] A systematic review and meta-analysis[275] of psychosocial interventions following self-harm in adults concluded that CBT "seems to be effective in patients after self-harm", and specific studies (which require replication) provide support for dialectical behaviour therapy (for individuals with borderline personality disorder),[298] psychodynamic interpersonal therapy,[299] and mentalisation-based therapy.[300] Efforts have also been made to establish whether the Collaborative Assessment and Management of Suicidal ideation and behaviour (CAMS) is feasible and clinically effective.[301] The Attempted Suicide Short Intervention Program (ASSIP), a brief intervention consisting of integrated therapy and personalised letters, showed encouraging findings in patients who have attempted suicide.[302]

A meta-analysis[303] of therapeutic interventions for attempted suicide and self-harm in adolescents found that therapeutic interventions are effective in reducing self-harm when it is treated as a global category that includes suicidal and non-suicidal self-harm, but that the effects are weaker when suicidal and non-suicidal behaviour are examined separately. This weaker effect when separately analysing suicidal and non-suicidal behaviour is consistent with the findings of a Cochrane review of interventions for children and adolescents who self-harm.[277] The authors of the review found only 11 trials, most of which were single trials, from which they concluded that therapeutic assessment, mentalisation, and dialectical behaviour therapy "warrant further evaluation"(see also Part 4).[277] Treatments that target depression are not effective in reducing suicidal thoughts or suicide attempts.[304] A marked heterogeneity is notable across treatment studies in the field, and many studies have small sample sizes and evidence of publication bias is clear since no published studies seem to report negative findings.[275] Replication of the existing treatments by independent groups is needed, as is the development of evidence-based assessment measures that are clinically useful in the field of treatment research for suicidal behaviour (see Part 6).

The development, assessment, and implementation of psychological treatments for suicidal behaviour should be prioritised. Moreover, the extent to which psychological treatments are effective for different sociodemographic populations needs to be established (eg, men *vs* women, adolescents *vs* older adults, individuals from different ethnic backgrounds), as well as in different health-care settings (eg, primary or secondary care *vs* acute settings) and patient groups (eg, psychiatric inpatients *vs* outpatients; see Part 8). The sex-specific research is especially important, because more men die by suicide than women in all countries worldwide,[265] but many more women participate in treatment trials for suicidal behaviour.[276] Additionally, the optimal time to give treatment interventions to reduce the risk of future suicidal behaviour among those who have attempted suicide is still unclear.

Psychological treatments are not a panacea. For those psychological treatments that are effective, the overall effect sizes are small.[276,305,306] Also, for many reasons, including access and suitability, psychological treatments reach only a minority of people who die by suicide or who are suicidal. Given the inequality gradient for suicide (ie, people from lower socioeconomic backgrounds are substantially more likely to die by suicide than people in a higher socioeconomic situation[307]), the structural inequalities (eg, poverty) that contribute to the excess in suicide mortality among those from low socioeconomic backgrounds needs to be challenged.

Most suicides occur in low-income and middle-income countries,[265] so the extent to which treatments that are developed in high-income countries are generalisable worldwide needs careful consideration (see Part 2). When developing and assessing treatment trials, consideration should be given to whether a tailored or modular approach is desirable and feasible, whether the treatment is based on principles or manualised (eg, a person-centered approach or an approach with a specified session plan), and whether the interventions account for different risk profiles and inequalities (see Part 8). Furthermore, as noted in Part 1, efforts need to be re-focussed to ensure that when treatment successes occur, the mechanisms responsible for them are understood (eg, does prevention of suicide depend on changes in reward sensitivity?). An appreciation of mechanisms will help explain why treatments that are expected to be effective are not.

Challenges and opportunities for research

Panel 17 highlights the key challenges and opportunities for treatment research for suicidal behaviour in the next decade and beyond. Since individuals who are at imminent risk of death by suicide are usually excluded from treatment trials, little is known about which treatments might be effective in this patient group. Similarly, most people who are suicidal do not receive treatment;[308] therefore, an understanding is needed of the barriers to seeking help and accessing treatment. One reason some people in distress are reluctant to seek psychological or psychiatric treatment could be for fear of stigma. Organisations such as Headspace in Australia (see Part 2) offer promising stepped-care treatment models that aim to remove the stigma of mental disorders, are set in the community, and provide people with a way to seek help for relatives and friends. Another challenge is that patients with suicidal behaviour or ideation are difficult to keep in treatment;[309] an understanding of the factors associated with

disengagement is needed, so that the treatment given can be maximised when patients are in health-care settings—eg, innovative brief-contact interventions have been shown to offer some promise in acute settings.[273,310–312] Maximised treatment approaches should be considered as adjuncts to existing treatments and could be effective in reducing the likelihood that individuals act on their suicidal thoughts.[310,311] Although some public health interventions for suicide prevention have a multilevel approach and explored synergies through a combination of interventions,[297,313] few examples exist in which interventions for suicide prevention have explored combining different psychological treatments (see Part 3). Given the heterogeneity of individuals who attempt suicide or die by suicide, exploring the efficacy of treatment combinations is likely to be a rewarding approach. However, potential iatrogenic effects should be monitored in such studies, as well as in monotreatment studies (see Part 6). The potential for harm during psychological treatments research has been highlighted in the Royal Australian and New Zealand College of Psychiatrists guidelines for deliberate self-harm.[305]

To facilitate the pooling of findings across treatment studies, we urge researchers of suicidal behaviour and ideation to agree on a common set of core outcome measures (see Part 6). In the USA, some movement has been made in this regard;[273] however, an international consensus would be ideal. Agreement on such a set of measures would be aided by the gathering of an international, interdisciplinary working group. We also call for all psychological treatment trials to include a measures of suicidal ideation and behaviour as an outcome measure. Although suicidal behaviour occurs transdiagnostically, the differential prevalence of suicidal ideation and behaviour across psychiatric categories needs to be considered to understand why, for example, individuals with bipolar disorder are at particularly high risk of suicide.[314] Research into psychological treatments needs to embrace the assessment of potential mechanisms to account for treatment efficacy, and establish the active ingredients of effective treatments for suicidal ideation and behaviour (see Part 1).

The extent to which new technologies could be useful to engage so-called difficult to reach populations (eg, men, adolescents) needs to be investigated.[315,316] For example, could gaming technology be harnessed to engage young people in seeking help and treatment? Mobile applications offer opportunities to monitor suicidal ideation and mood in real time and have the potential to enhance the ability to identify (and intervene) when individuals are at their most vulnerable; however, these applications should be developed with the same rigor as traditional methods of psychological treatment (see Part 5). Arguably, the field of suicide prevention has not given sufficient consideration to the cultural influences and pressures on men, women, and adolescents (eg, depictions of masculinity). Given the high incidence of death by suicide among male individuals, the improved integration of such factors into the understanding of suicide risk and suicide prevention efforts is crucial.[317–319]

Those with lived experience of suicidal behaviour (eg, individuals bereaved by suicide, and those with personal experience) should be involved in all stages of treatment development.[320] Since little is known about what protects vulnerable people from engaging in suicidal behaviour, research into potential buffering factors should be central to the development of treatment protocols (see Part 4).

Finally, multidisciplinary collaboration is key to the success of developing, assessing, and implementing psychological treatments to prevent suicide. Since suicide is an end product of the interplay between psychological, social, biological, clinical, and cultural factors, an interdisciplinary approach should be the norm in psychological treatment research (see Part 7). However, since an individual who attempts suicide makes a decision to end their life, in the context of a range of different risk factors, psychology needs to be at the centre of future developments in the field.

Now is an exciting time to be working in research for psychological treatments for suicidal behaviour, since the theoretical and empirical foundations are available for promising treatments. However, in the next decade and beyond, innovative thinking and practice will be needed to ensure that the promise of research into psychological treatments is realised and leads to a reduction in suicidal ideation and suicide attempts.

Part 10: Active innovation and scrutiny of future psychological treatments research

Inspecting ideas and making space for future ideas

Psychological treatments are highly effective for many patients, but a large proportion of patients either do not respond to existing therapies, or the therapies are inaccessible to them. New ideas are needed, and they should be critically inspected, with the progression and rejection of ideas via rigorous and high-quality research.

In the Introduction, we used the metaphor of the fourth plinth in London's Trafalgar Square. The plinth is a metaphor to make contemporary ideas visible and to give them critical consideration. Although some pieces that are displayed on the plinth will be preserved for posterity, others might not be. Some psychological treatments or research ideas should not stand on the plinth forever, whereas some might stand the test of time. Ideas for the plinth need to be generated, inspected, and replaced over time, all within the context of a science-driven framework. Psychological treatment is a relatively young field compared with some medical treatment fields, and the notion of innovation and turnover are crucial parts of its future.

How might this innovation work for psychological treatments? The wide range of potential topics would need to be considered, as well as how these topics could

be selected, where they would be used, how they could achieve visibility, in addition to the need for a repeated cycle of this endeavour, the ultimate aim of which would be to improve the discussion and debate of the pertinent issues to make a difference for mental health. Topics could include both longstanding challenges and novel ideas such as new findings that would benefit from constructive and rapid scrutiny (eg, therapeutic approaches that emerge from the findings of preclinical studies, new ideas from sister disciplines, and new technology and ethical issues). Exciting new directions that emerge in these and other contexts should be clearly formulated, considered, and reflected upon. The ideas would need to undergo rigorous debate within and beyond the field of mental health science, and empirical assessment in the context of scientifically sound studies—eg, well controlled randomised trials.

Open and constructive debate needs to be encouraged, without new ideas being too swiftly quashed by tradition and vested interests in maintaining a status quo. However, new ideas and trends in thinking must be scrutinised before being accepted into clinical practice. One problem for the field of psychological therapy is the need to promote the use of evidence-based treatments by practitioners, who might prefer to ignore the evidence and use techniques for which they have a personal preference. For example, exposure is a treatment technique that is theoretically driven with an excellent evidence base and there is a strong scientific understanding of the mechanisms that underlie its effectiveness;[82] however, in practice, a substantial proportion of therapists do not use this technique.[321] This reluctance and sparse uptake of empirically supported interventions, or aspects of them, among practitioners is an issue that needs to be understood and rectified.

The plinth metaphor also provides a way to question older ideas that are now taken for granted, but that would benefit from further examination. Many broad issues that affect the whole field of psychological treatment require discussion (eg, the diagnostic systems, the quantity of academic publications *vs* their capacity to affect patients, and funding issues specific to psychological treatments), in addition to many issues that are relevant to science generally—from reproducibility to open data. Psychological science is a young discipline compared with many other fields, and emphasis on the progression of psychological treatments over the past century could be beneficial to stimulate innovation. Parallels exist between some of our suggestions in this Commission and the Science in Transition initiative in the Netherlands, which calls for several key reforms in science with the goal of achieving reproducible outcomes.[322,323]

For the **Science in Transition** website see http://scienceintransition.nl/

How can topics be selected? In the art world, the Fourth Plinth Initiative is an open competition to artists and is subject to a review panel. For research into psychological treatments, an equivalent competition or selection process could be held, with specific calls for people to raise challenging ideas that can catalyse progress. This process would generate topics outside what can be imagined now, and potentially create a way to capture the concerns and questions of younger generations in the field (eg, why is neuroscience not being used in treatment research more?), or those of researchers with several decades of experience (eg, why have effect sizes for psychological treatments not improved?).

Such debates and discussions could be included in a dedicated session at conferences and cross-disciplinary meetings, in a specific type of journal article, and in electronic media and areas and settings that allow debate and scrutiny. The metaphor could be adapted to fit a range of outlets, and journal editors and conference organisers could be encouraged to provide space for it. To bring attention to the resulting ideas, an annual prize could be awarded for topics that have attracted attention and made constructive progress.

The empty plinth metaphor highlights the need for repetition in the process of innovation, so that novel ideas for psychological treatment would constantly be generated, tested, and disseminated. This iterative process would not only encourage innovation, but would also enable differentiation of the new treatments and ideas that can stand the test of time, and allow long-held assumptions to be questioned to bring about progress. Essentially, these processes all occur throughout the scientific process, but—as we have discussed throughout this Commission—because of the huge scale of mental disorders globally, progress needs to speed up within psychological treatments research. Borrowing an idea from the arts gives a metaphor for one way (among many needed) to start achieving this goal.

Mental disorders and psychological treatments provide crucial and demanding targets for research enquiry. Creative but realistic solutions require communication and meaningful multidisciplinary collaborations among researchers and funding agencies, and some so-called blue skies thinking from outside the field. Additional researchers from across all disciplines are needed within the psychological treatment field, since a vast range of important questions remain that need to be addressed. This need within the field poses a great opportunity for many early career scientists to make landmark contributions, and other researchers should be encouraged into the field.

Debatably, research has stagnated in some areas of psychological treatment. Outcomes for many mental disorders (ie, depression, obsessive compulsive disorder, schizophrenia, and bipolar disorders) have not improved since the original interventions were developed, and might even be declining.[324] Understandably, an emphasis has been put on increasing access to psychological treatments,[24] given the large unmet need and changing models of service delivery.[22,83,325,326] However, an equally

strong need exists for the development of innovative new psychological treatments for the large proportion of people who do not engage with or respond to existing interventions, or who relapse after a seemingly successful course of treatment. The proportion of people who are in one of these categories varies by disorder, age group, and research study, but it can be considered to be at least 50%.[327,328] We also see a pressing need for multiple solutions, given the scale of the challenge ahead. A range of approaches could be valuable in this endeavour, including the dissemination of evidence-based therapies and increasing the accessibility of evidence-based psychotherapies. Therefore, although we see the need for a multipronged approach to tackling mental disorders worldwide, we argue that the development of new psychological treatments is one of the most promising approaches, especially given the scale of the problem of mental disorders from a public health perspective.

What factors might encourage stagnation or innovation? Branding, communication, and funding

One obstacle to innovation in the field of psychological treatment research is branding of psychological interventions, with the accompanying restrictions due to intellectual property issues. Such branding prevents the dissemination and implementation of psychological therapies, and stifles innovation by implying ownership.[329] A sustainable, not-for-profit model for the development of psychological interventions is an alternative and potentially better model than the branding model. Some research groups are under increasing pressure from so-called knowledge transfer departments at universities to brand their work for uniqueness—this pressure needs to be resisted. Instead, departments and research groups should be in favour of developments in psychological therapies that are more open, that highlight shared common components, and that are described to an extent that they can benefit from examination by the wider psychological treatment community. The issue of sharing knowledge is clearly complex because of concerns regarding incent-ivising investment in psychological treatments from a range of sources, and the need for quality control within some interventions. The development of citizen science has the potential to counteract branding and provide a fertile ground for innovation.

Noticeably, as discussed in Part 7 about training, the majority of psychological treatment researchers stick to what they know. Such adherence is rewarded by strong CVs, grant funding, and in-depth knowledge of a specific field. However, this approach can also lead to insularity. Input from fields such as neuroscience, maths, pharmacology, and more diverse disciplines, such as medical geography,[330] could help clinical researchers and practitioners think differently. Jointly reviewing advances in areas such as cognitive and social science to identify which innovations might be relevant to improving psychological therapies is entirely feasible. Such an approach has substantial potential to facilitate the introduction of new, scientifically sound ideas into psychological treatment. Innovation can benefit from creativity, including taking ideas from one area and seeing if they apply to another for treatment benefits.

Improvements are needed in communication between service users, clinicians, and across the health services. Mental and physical health-care services are typically entirely separate services, with minimal overlap despite their close relationship in terms of pathology, service use, and cost to the health services around the world.[331] Improving communication between providers of these two health-care services via shared training, resources, or even co-location would be a fundamental step toward innovation, with scope to give substantial benefits to the entire health-care system. Drawing on multiple areas of expertise will be important—particularly, obtaining input from patients and carers, which is a topic that is receiving increasing attention,[188] but which requires further consideration.

The issues of innovation and improvement cannot be dissociated from those of dissemination and implementation. Innovations that stay localised will benefit some patients but the effect will be minimal (see Part 2). Furthermore, the time taken for a treatment to get from bench to bedside will continue to be unacceptably high (estimated to be 17 years, although some argue the development of psychological treatments is quicker than pharmacological treatments[3,332]) unless dissemination and implementation are part of the development plan from the outset. Communication between stakeholders is essential to ensuring the effects of innovations are as anticipated. Only through the development of meaningful networks can genuine collaborations be built—eg, joint training, conferences, and funding. Such joint networks need to be funded appropriately for the stage of development, with basic researchers and clinicians having a bidirectional conversation, initially by email but then face to face in a relaxed atmosphere with time to think creatively, discuss constructively, and develop testable hypotheses.

The role of funders in promoting or stifling innovation cannot be overemphasised. The NIMH's influence on funding has been profound, and inclusion of a category on the RDoC entitled "Other"—so that researchers are not restricted to only studying established research domains—encourages new ideas.[333] Although researchers understand that funding agencies tend to want to avoid risks, the funding of high-risk studies is fundamental to the development of new treatments. Agency support to fund proof-of-concept studies in psychological therapies could be especially important to the field. The extent of funding for international research into mental disorders, and psychological treatments in particular, is far too low;[334,335] increased funding is essential for progress and to take risks in new areas.

Globally, within large funding organisations, mental health is often included with other diseases or with, for example, neuroscience. Representation within these funding organisations of people with experience in mental health research can be minimal and people with genuine expertise in mental health are needed within the decision-making bodies of the major funding organisations. Clearer representation of expertise in psychological treatments would also be of benefit. A review of the international funding organisations that address mental health would be useful, including the extent to which psychological treatment research is accommodated. Some charities fund research, which is of course welcomed, but unfortunately many smaller charities often do not have the capacity for a rigorous research review process. The quality and effect of studies that do not benefit from peer review and scientific support is often suboptimal. Funding models whereby smaller charities that support mental health research are themselves supported by larger charities, with regard to their commissioning and execution of research, are likely to improve both the quality of research and the value for money of the research projects. The creation of a framework for peer review for mental health in general, and psychological treatment in particular, or even a possible outsourcing model for such processes, might help many organisations with funding initiatives in this area.

How can the effectiveness of efforts toward new treatments be assessed?

Broadly, our aim in undertaking this Commission was to identify the scope of advancing research efforts to improve mental health globally via improvements in the effectiveness and the global reach of psychological treatments. We have outlined an agenda of some of the areas in which we see real scope to improve treatment research and treatment delivery to enable more effective interventions and greater accessibility of such treatments to individuals with mental disorders than have been available to date. Treatment protocols that effectively treat and prevent the onset of mental disorders will have a key role as one of the many contributions that are needed to relieve the substantial worldwide burden of mental disorders.

The ability to assess in a tangible and meaningful way whether the goal of improving treatments for mental disorders has been achieved remains a challenge for the field. The initial indicator of success is within the outcomes of the treatment trial—ie, whether the effect sizes indicate improved efficacy of novel and refined psychological interventions. In the longer term, meta-analyses will outline whether new treatment approaches have improved effectiveness, and thus, in turn, contribute to reducing the prevalence and the burden of mental disorders. In the future, the findings of epidemiological studies that show changes in the prevalence of mental disorders over time will reveal the success of scalable treatment and prevention approaches. We acknowledge, however, that measurement in this field can be complicated and ambitious—eg, changes in the diagnostic classification systems over time complicate comparisons. We therefore see a need for research on how to define and quantify the burden of mental disorders globally and over time. We see scope for progress to be made, not only by examining changes in prevalence, but also by investigating improvements in the functional effect of mental disorders—from impairments in social and occupational functioning, through to quality of life. Such a suggestion aligns with our acknowledgment of the value of expanding ideas of mental disorders beyond the notions of disease and infirmity, to outcomes with broad functional relevance (eg, an individual's capacity to adapt, and self-manage; see Introduction).

Innovation to create new treatments—what ideas can be put on the plinth in the first round?

Increasing access to effective psychological treatments is a priority, but investment in innovations that will energise the research field of psychological treatment and improve therapeutic outcomes is equally important.[22,83] Many books and journal articles have been dedicated to the issue of innovation, and even an entire journal is devoted to this topic (*Healthcare: The Journal of Delivery Science and Innovation*), which commenced in June 2013. Innovation is clearly a challenging area and what is presented as an innovation can often be seen as old wine in a new bottle. Innovation needs to be put into a historical context, so that existing ideas are not repackaged with enthusiasm as an innovation.[336] Engagement is needed in the critical inspection, progression, and rejection of ideas through research. One approach is to change the nature of the questions being asked; here we begin with two examples.

What matters to patients?

Most clinical research has tended to focus on single diagnoses, despite many patients having multiple coexisting disorders.[230] Clinicians have guidelines for the treatment of specific diagnoses but almost no data to guide them with regard to evidence-based decision making for patients who have common co-occurring disorders—eg, anxiety and depression. Patients' difficulties can alternatively be considered in terms of the problem they are having rather than in diagnostic terms—eg, loneliness or betrayal.[337] Linking with social psychology and having a problem-based approach to the development of psychological treatments, instead of a disorder-based approach, is likely to lead to new ways of thinking about and addressing mental disorders, which was partly the intention of the RDoC initiative.[333] Such approaches could increase engage-ment in and the acceptability of therapies, but challenges would still

exist for agreeing operationalised definitions of the problem, and ensuring that such difficulties were affecting people's lives in ways they value and that could be viewed within a psychological framework.

What matters to researchers?

Many things matter to researchers, but most scientists are curious about what does not work, as well as what does. Data that do not obey the expected rules are essential to scientific progress. For psychological treatments research, defining non-responders, identifying which people relapse, and those who do not engage in treatment, are all necessary and crucial steps.[328] A thorough and focused analysis of the characteristics of those individuals who do not respond to existing treatments, and having dedicated funding for such research, are priorities that would have a positive effect and bring generalisable benefits to existing and new treatments. Additionally, in areas in which no existing treatments work adequately, the generation of novel treatments is essential.

What next?

We see mental health as a substantial global challenge, but at the same time we recognise that nowadays we are faced with an array of pressing priorities that demand global attention and action, including, but in no way limited to: climate change, international conflicts, famine, and the displacement of millions of people from their home countries. Notwithstanding that many such substantial problems exist in the world, in the domain of mental health, we call for increased research efforts to advance psychological treatments, so that more effective interventions will serve as an essential part of our set of approaches that are needed to make an impact upon the burden of mental disorders worldwide and improve lives.

We acknowledge that our call for developments in psychological treatments for mental disorders is but one endeavour in the context other similar timely initiatives. For example, Wykes and colleagues[338] have laid out six key priorities for a mental health research agenda for Europe and worldwide. Mental health is increasingly being recognised as an area that needs to move forward on a global scale. Furthermore, psychological interventions can be applied not only to mental disorders, but have been increasingly of use across a range of areas—eg, in changing health behaviour, managing the psychological aspects and effects of physical health problems (ie, pain management and somatic concerns, psycho-oncology), and instituting organisational change.

Clinicians, researchers, patients, carers, funders, commissioners, managers, policy planners, change experts, and the general public all have a part to play in innovating psychological therapies, and a focus on any one of the ideas presented in this Commission has the potential to bring about substantial and much-needed improvements. More ideas will be needed than just those included here. This Commission is not a specific roadmap, all relevant areas of research and mental health science need to be considered to gain traction in this endeavour. Innovations arising from thoughtful effort have genuine potential to transform the science and practice of psychological therapies, and the lives of all of those who are affected by mental disorders.

Contributors

All authors contributed equally.

Declaration of interests

EAH reports her primary affiliation is the Division of Psychology, Department for Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden. She is an Honorary Professor of Clinical Psychology at the University of Oxford, Oxford, UK, in the Department of Psychiatry, and holds an honorary position at the Medical Research Council's (MRC's) Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, UK, and receives no renumeration for these roles. She serves on the board of the charity MQ: Transforming Mental Health and receives no remuneration for this role. She is on the board of overseers for the charity Children and War Foundation, Oslo, Norway, and receives no remuneration for this role. She is an associate editor for *Behaviour Research and Therapy*, for which she receives an honorarium, and was a founding associate editor for *Clinical Psychological Science*. She is on the editorial board of *Psychological Science* and *Cognitive Behaviour Therapy* and receives no remuneration for these roles. She has presented keynote addresses at conferences, such as European Association for Behavioural and Cognitive Therapies (EABCT), and received an honorarium for this. She has presented clinical training workshops, some of which include a fee. She receives royalties from her co-authored book *Oxford Guide to Imagery in Cognitive Therapy*. AG is a Professor of Clinical Psychology at the Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden. He is on the editorial board of *Behaviour Research and Therapy* and receives no remuneration for this role. He has presented clinical training workshops and provided supervision to clinicians at eating disorder treatment units, most of which include a fee, but are not related to the current contribution. He also receives royalties from his co-authored books on eating disorders and body image. CJH has received consultancy fees from P1vital, Lundbeck, and Servier. She has grant incomes from Johnson & Johnson, UCB Pharma, Lundbeck, and Sunovion. She is employed by the University Department of Psychiatry at Oxford University, Oxford, UK, and holds an associate Professorship at the Department of Psychology, University of Oslo, Oslo, Norway. She is an associate editor for *Psychological Medicine*, for which she receives an honorarium. PGR reports grants from the UK National Institute for Health Research (NIHR) outside the submitted work. He is employed by Imperial College London, London, UK, and Central and North West London Foundation National Health Service (NHS) Trust. He has been involved in the development and adaption of psychological interventions, but receives no payments from these. PC is Head of the Department of Clinical, Neuro and Developmental Psychology at the Vrije Universiteit, Amsterdam, for which he receives his annual salary. He is deputy editor of *Depression and Anxiety*, for which his university receives a fee. He is also leader of the Mental Health Program of the EMGO Institute for Health and Care Research at the Vrije Universiteit and Vrije Universiteit Medical Center in Amsterdam, Honorary President of the Canadian Psychological Association (until June, 2017), coordinator of the Global Consortium for Depression Prevention, a member of the Guideline Development Panel for Depressive Disorders of the American Psychological Association, Chair of the Sir Henry Wellcome Building for Mood Disorders Research Expert Advisory Board in Exeter, UK, adviser for several national and international research projects, and a member of the editorial board of several scientific journals; he receives no remuneration for any of these positions. He does receive expense allowances for his membership of the Board of Directors of the Fonds Psychische Gezondheid and Korrelatie, and for being Chair of the Parliamentary Assistance Coordination Office Committee of the Raad voor Civiel-militaire Zorg en Onderzoek of the Dutch Ministry of Defence. He also receives royalties for books he has authored or co-authored and

for occasional workshops and invited lectures. APM's primary affiliation is the School of Psychological Sciences, University of Manchester, Manchester, UK. He is also Director of the Psychosis Research Unit, Greater Manchester West NHS Trust, Manchester, UK. He serves on several editorial boards and receives no remuneration for these roles. He has presented keynote addresses at conferences and given clinical training workshops, some of which have included a fee. He receives royalties from several co-authored and edited books. He delivers cognitive behavioural therapy within the UK NHS and has received funding from both the MRC and UK NIHR to do evaluative research into the efficacy of psychological therapies. JPR is a consultant for Cambridge Cognition and Takeda. He is an associate editor for *Neuroimage: Clinical* and receives an honorarium for this role. He is a handling editor at *Royal Society Open Science* and is on the Editorial Board of *Computational Psychiatry*; he receives no renumeration for these roles. He has received the British Psychological Society's Spearman Medal, the British Association for Psychopharmacology's Senior Award, and a Philip Leverhulme Prize in Psychology. He has received renumeration for acting as a panel member for the Research Council of Norway. None of these positions or awards have any direct relation to this Commission. CLHB is Professor of Clinical Psychology at the Department of Psychiatry at the Academic Medical Center at the University of Amsterdam, Amsterdam, Netherlands, and at the University of Groningen, Groningen, Netherlands. She also has a Guest Professorship at the Faculty of Psychology and Pedagogy at Ghent University, Ghent, Belgium. She is co-editor of *PLoS ONE*, and receives no honorarium for this role. She serves on the board of the section for Affective Disorders at Dutch Cognitive Behavioural Association and is a boardmember of the Dutch multidisciplinary guideline for anxiety and depression; she receives no remuneration for this role. She is an adviser for the minister on National Health Care in the Netherlands on issues of care for inclusion in the statutory insured package (Advies Pakket Commissie, ZIN); she receives an honorarium for this role and it has no direct relation to the current contribution. She has received a fellowship at the Netherlands Institute of Advanced Sciences, supported by the Royal Netherlands Academy of Arts and Sciences and this is not directly related to the current contribution. She has given keynote addresses at conferences such as EABCT 2014 and the European Conference of Psychology and received an honorarium. She has given clinical training workshops, some of which included a fee. She receives royalties from her books and co-edited books. RCO'C's primary affiliation is the Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK, where he heads the Mental Health and Wellbeing Research Group. He is also Director of the Suicidal Behaviour Research Laboratory at the University of Glasgow. He is joint chief editor for *Archives of Suicide Research* and associate editor for *Suicide and Life-Threatening Behavior and Behavior Therapy*. He also serves on several editorial boards and is a current Vice President of the International Association for Suicide Prevention and past President of the International Academy of Sucide Research. He receives an honorarium for his role at *Behavior Therapy*, but does not receive a remuneration for any of the other roles. He was a member of the National Institute of Health and Care Excellence's guideline development group for the longer-term management of self-harm and he is a member of Health Education England's Self-Harm & Suicide Prevention Competence Framework Expert Reference Group. He sits on the Scottish Government's Suicide Prevention and Implementation Monitoring Group. He receives royalties from one co-authored and several co-edited books, occasional workshops, and invited addresses. RS served as a Senior Adviser for the MQ: PsyImpact programme from February, 2015, to February, 2016. She was a consultant for Big Health. She also receives royalties for books she has authored or co-authored (American Psychological Association Books, Elsevier Press) and occasional workshops and invited addresses. MLM's primary affiliation is the School of Psychology, UNSW, Sydney, NSW, Australia, where she is a Professor and PLuS Alliance Fellow. She is a consulting editor of the *Journal of Experimental Psychology: Applied* and *Clinical Psychological Science*, and a member of four additional editorial boards; she receives no remuneration for these roles. She has given keynote addresses at conferences and for some of these she has received honoraria. MGC's primary affiliation is the Department of Psychology, University of California, Los Angeles (UCLA), CA, USA, where she is Vice Chair, which provides her annual salary, supplemented by summer funds from grants from the US National Institute of Mental Health or the Defense Advanced Research Projects Agency. She is editor-in-chief of *Behaviour Research and Therapy* and associate editor of *Psychological Bulletin*, for which she receives remuneration. She is director of the UCLA Anxiety and Depression Center; co-director of the UCLA Staglin Family Music for Behavioral and Brain Health; former President of the Association for Behavior and Cognitive Therapy; co-chair of the Human Studies Section of the UCLA Grand Challenge for Depression; member of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) Steering Committee for the American Psychiatric Association; member of the Scientific Advisory Board for the Center of Excellence on Generalization Research at the University of Leuven, Leuven, Belgium; Honorary Member of the Experimental Psychopathology Group (Dutch-Flemish Postgraduate School for Research and Education); and Honorary Fellow of the Department of Psychiatry at Oxford University, Oxford, UK. She receives no remuneration for any of these positions. She received remuneration for her awards as Eleonore Trefftz Guest Professorship (Technical University of Dresden, Dresden Germany) and the International Francqui Professor (Univeristy of Leuven). She also receives royalties for books she has authored or co-authored (American Psychological Association Books, Elsevier Press) and occasional workshops and invited addresses.

Acknowledgments

Additional contributors who also attended The *Lancet Psychiatry* meeting in London, UK, in December, 2015 include: E Barley, N Balmer, S E Blackwell, M Browning, K Carroll, S Cartwright-Hatton, C Creswell, T Dalgleish, M Di Simplicio, S Dix, B Dunn, P Fearon, C Hirsch, J M Hooley, L Iyadurai, S Jones, S Kamboj, A Milton, A Reinecke, U Schmidt, and M Van Ommeren. We also thank C Fairburn and G Goodwin for helpful discussions. We thank R Emsley for his consultancy regarding clinical trial methodology, and we thank L Iyadurai and E L James for help preparing the manuscript. We are grateful for support from MQ: Transforming Mental Health for travel expenses to the Commission meeting held by The *Lancet Psychiatry*. EAH is supported by the Karolinska Institutet and the Lupina Foundation of Toronto. She has received support from the MRC intramural programme (MRC-A060–5PR50) and the NIHR Oxford Biomedical Research Centre Programme. AG is supported by the Swedish Foundation for Humanities and Social Sciences (Riksbankens Jubileumsfond). CJH has research funding from the Wellcome Trust, MRC, and the NIHR Oxford Health Biomedical Research Centre. PGR has funding from the NIHR to develop and evaluate early interventions in randomised controlled trials, and receives support from the Imperial NIHR Biomedical Research Centre. PC has funding from the European Union (FP7 and H2020 programmes), ZonMw (Dutch Health Research Council), and the Protestants Fonds Geestelijke Volksgezondheid (Protestant Foundation for Mental Health). APM has received funding from the NIHR. JPR is funded by the Wellcome Trust. CLHB is supported by the Department of Psychiatry at the Academic Medical Center of the University of Amsterdam, and the Netherlands Institute of Advanced Sciences (2017), which is supported by the Royal Netherlands Academy of Arts and Sciences. RCOC has received funding from the US Department of Defense, NIHR, NHS, NHS Greater Glasgow and Clyde, NHS Health Scotland, MRC, MQ: Transforming Mental Health Research, the Chief Scientist Office, and the Scottish Government. RS's research at Great Ormond Street Hospital, NHS Foundation Trust and Univeristy College London Great Ormond Street Institute of Child Health, London, UK, is supported by the NIHR Great Ormond Street Hospital Biomedical Research Centre. MLM is supported by the Australian National Health and Medical Research Council. She also receives support from the PLuS Alliance Fellows Funding Scheme (UNSW, Sydney, Australia). MGC is funded by the US National Institutes of Mental Health (R01 MH1001171, R01 MH1014531, R34 MH101359, R01 MH102274), the US Defense Advanced Research Projects Agency (R21 MH1010336), and the US National Aeronautics and Space Administration (NNX15AP57G). The views expressed in this Commission are those of the authors and are not necessarily those of any of the funding bodies or agencies.

References

1 James W. The principles of psychology. New York, NY: Henry Holt & Co, 1890.
2 Holmes EA, Craske MG, Graybiel AM. Psychological treatments: a call for mental-health science. *Nature* 2014; **511:** 287–89.
3 Joyce C. Transforming our approach to translational neuroscience: the role and impact of charitable nonprofits in research. *Neuron* 2014; **84:** 526–32.
4 McHugh RK, Whitton SW, Peckham AD, Welge JA, Otto MW. Patient preference for psychological vs pharmacologic treatment of psychiatric disorders: a meta-analytic review. *J Clin Psychiatry* 2013; **74:** 595–602.
5 Steel Z, Marnane C, Iranpour C, et al. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980–2013. *Int J Epidemiol* 2014; **43:** 476–93.
6 Värnik P. Suicide in the world. *Int J Environ Res Public Health* 2012; **9:** 760–71.
7 Whiteford HA, Degenhardt L, Rehm J, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet* 2013; **382:** 1575–86.
8 Cuijpers P, Vogelzangs N, Twisk J, Kleiboer A, Li J, Penninx BW. Comprehensive meta-analysis of excess mortality in depression in the general community versus patients with specific illnesses. *Am J Psychiatry* 2014; **171:** 453–62.
9 Chisholm D, Sweeny K, Sheehan P, et al. Scaling-up treatment of depression and anxiety: a global return on investment analysis. *Lancet Psychiatry* 2016; **3:** 415–24.
10 Hu TW. Perspectives: an international review of the national cost estimates of mental illness, 1990–2003. *J Ment Health Policy Econ* 2006; **9:** 3–13.
11 Gustavsson A, Svensson M, Jacobi F, et al. Cost of disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol* 2011; **21:** 718–79.
12 Bloom DE, Cafiero ET, Jané-Llopis E, et al. The global economic burden of non-communicable diseases. Geneva: World Economic Forum, 2011.
13 Huber M, Knottnerus JA, Green L, et al. How should we define health? *BMJ* 2011; **343:** d4163.
14 Andrews G, Issakidis C, Sanderson K, Corry J, Lapsley H. Utilising survey data to inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders. *Br J Psychiatry* 2004; **184:** 526–33.
15 Baxter AJ, Patton G, Scott KM, Degenhardt L, Whiteford HA. Global epidemiology of mental disorders: what are we missing? *PLoS One* 2013; **8:** e65514.
16 Park-Lee E, Lipari RN, Hedden SL, Kroutil LA, Porter JD. Receipt of services for substance use and mental health issues among adults: results from the 2016 National Survey on Drug Use and Health. Rockville, MD: Substance Abuse and Mental Health Services Administration, 2017. https://www.samhsa.gov/data/sites/default/files/NSDUH-DR-FFR2-2016/NSDUH-DR-FFR2-2016.htm (accessed Jan 5, 2018).
17 Appleby L, Asherson P, Ashtarikiani, et al. Mental health and wellbeing in England. Adult psychiatric morbidity survey 2014. Leeds: NHS Digital, 2016.
18 Wang PS, Lane M, Olfson M, Pincus HA, Wells KB, Kessler RC. Twelve-month use of mental health services in the United States: results from the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005; **62:** 629–40.
19 Olfson M, Marcus SC. National trends in outpatient psychotherapy. *Am J Psychiatry* 2010; **167:** 1456–63.
20 Olfson M, Marcus SC. National patterns in antidepressant medication treatment. *Arch Gen Psychiatry* 2009; **66:** 848–56.
21 Gerhard T, Akincigil A, Correll CU, Foglio NJ, Crystal S, Olfson M. National trends in second-generation antipsychotic augmentation for nonpsychotic depression. *J Clin Psychiatry* 2014; **75:** 490–97.
22 Kazdin AE, Blase SL. Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspect Psychol Sci* 2011; **6:** 21–37.
23 Fairburn CG, Patel V. The global dissemination of psychological treatments: a road map for research and practice. *Am J Psychiatry* 2014; **171:** 495–98.
24 Clark DM. Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. *Int Rev Psychiatry* 2011; **23:** 318–27.
25 Turnbull HW. Correspondence of Isaac Newton, volume 1: 1661–1675. Cambridge: Cambridge University Press, 1959.
26 Greater London Authority. What's on the Fourth Plinth now. https://www.london.gov.uk/what-we-do/arts-and-culture/art-and-design/fourth-plinth (accessed May 4, 2016).
27 Kazdin AE. Mediators and mechanisms of change in psychotherapy. *Annu Rev Clin Psychol* 2007; **3:** 1–27.
28 Brown LA, Wiley JF, Wolitzky-Taylor K, et al. Changes in self-efficacy and outcome expectancy as predictors of anxiety outcomes from the CALM study. *Depress Anxiety* 2014; **31:** 678–89.
29 Sanislow CA. Updating the research domain criteria. *World Psychiatry* 2016; **15:** 222–23.
30 Kozak MJ, Cuthbert BN. The NIMH research domain criteria initiative: background, issues and pragmatics. *Psychophysiology* 2016; **53:** 286–97.
31 Miller GE, Cole SW. Clustering of depression and inflammation in adolescents previously exposed to childhood adversity. *Biol Psychiatry* 2012; **72:** 34–40.
32 Dalgleish T, Werner-Seidler A. Disruptions in autobiographical memory processing in depression and the emergence of memory therapeutics. *Trends Cogn Sci* 2014; **18:** 596–604.
33 Roiser JP. What has neuroscience ever done for us? *Psychologist* 2015; **28:** 284–87.
34 Fox E, Zougkou K, Ashwin C, Cahill S. Investigating the efficacy of attention bias modification in reducing high spider fear: the role of individual differences in initial bias. *J Behav Ther Exp Psychiatry* 2015; **49:** 84–93.
35 Twamley EW, Vella L, Burton CZ, Heaton RK, Jeste DV. Compensatory cognitive training for psychosis: effects in a randomized controlled trial. *J Clin Psychiatry* 2012; **73:** 1212–19.
36 Smith T, Iadarola S. Evidence base update for autism spectrum disorder. *J Clin Child Adolesc Psychol* 2015; **44:** 897–922.
37 Barlow DH, Craske MG. The phenomenology of panic. In: Rachman SJ, Maser JD, eds. Panic: psychological perspectives. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988: 11–36
38 Bouton ME, Mineka S, Barlow DH. A modern learning theory perspective on the etiology of panic disorder. *Psychol Rev* 2001; **108:** 4–32.
39 Clark DM. A cognitive approach to panic. *Behav Res Ther* 1986; **24:** 461–70.
40 Barlow DH, Craske MG. Mastery of your anxiety and panic. Albany, NY: Graywind Publications, 1988.
41 Salkovskis PM, Clark DM, Hackmann A. Treatment of panic attacks using cognitive therapy without exposure or breathing retraining. *Behav Res Ther* 1991; **29:** 161–66.
42 Hofmann SG, Smits JA. Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomized placebo-controlled trials. *J Clin Psychiatry* 2008; **69:** 621–32.
43 Foa EB, Steketee G, Grayson JB, Turner RM, Latimer P. Deliberate exposure and blocking of obsessive-compulsive rituals: immediate and long-term effects. *Behav Ther* 1984; **15:** 450–72.
44 Öst LG, Havnen A, Hansen B, Kvale G. Cognitive behavioral treatments of obsessive-compulsive disorder. A systematic review and meta-analysis of studies published 1993–2014. *Clin Psychol Rev* 2015; **40:** 156–69.
45 Lewinsohn PM. A behavioral approach to depression. In: Friedman RM, Katz MM, eds. The psychology of depression: contemporary theory and research. New York, NY: Wiley, 1974: 157–85.
46 Jacobson NS, Martell CR, Dimidjian S. Behavioral activation treatment for depression: returning to contextual roots. *Clin Psychol Sci Pract* 2001; **8:** 255–70.
47 Cuijpers P, van Straten A, Warmerdam L. Behavioral activation treatments of depression: a meta-analysis. *Clin Psychol Rev* 2007; **27:** 318–26.
48 Mohr DC, Spring B, Freedland KE, et al. The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychother Psychosom* 2009; **78:** 275–84.
49 Cuijpers P, Cristea IA, Karyotaki E, Reijnders M, Huibers MJ. How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. *World Psychiatry* 2016; **15:** 245–58.

50 Wampold BE. How important are the common factors in psychotherapy? An update. *World Psychiatry* 2015; **14:** 270–77.

51 Kazdin AE. Evidence-based psychotherapies I: qualifiers and limitations in what we know. *S Afr J Psychol* 2014; **44:** 381–403.

52 Cristea IA, Kok RN, Cuijpers P. Efficacy of cognitive bias modification interventions in anxiety and depression: meta-analysis. *Br J Psychiatry* 2015; **206:** 7–16.

53 Fox E, Zougkou K, Ridgewell A, Garner K. The serotonin transporter gene alters sensitivity to attention bias modification: evidence for a plasticity gene. *Biol Psychiatry* 2011; **70:** 1049–54.

54 Milad MR, Rosenbaum BL, Simon NM. Neuroscience of fear extinction: implications for assessment and treatment of fear-based and anxiety related disorders. *Behav Res Ther* 2014; **62:** 17–23.

55 Lanius RA, Vermetten E, Loewenstein RJ, et al. Emotion modulation in PTSD: clinical and neurobiological evidence for a dissociative subtype. *Am J Psychiatry* 2010; **167:** 640–47.

56 van Vugt MK, Hitchcock P, Shahar B, Britton W. The effects of mindfulness-based cognitive therapy on affective memory recall dynamics in depression: a mechanistic model of rumination. *Front Hum Neurosci* 2012; **6:** 257.

57 Wiles NJ, Thomas L, Turner N, et al. Long-term effectiveness and cost-effectiveness of cognitive behavioural therapy as an adjunct to pharmacotherapy for treatment-resistant depression in primary care: follow-up of the CoBalT randomised controlled trial. *Lancet Psychiatry* 2016; **3:** 137–44.

58 Bockting CL, Smid NH, Koeter MW, Spinhoven P, Beck AT, Schene AH. Enduring effects of preventive cognitive therapy in adults remitted from recurrent depression: a 10 year follow-up of a randomized controlled trial. *J Affect Disord* 2015; **185:** 188–94.

59 Wampold BE, Mondin GW, Moody M, Stich F, Benson K, Ahn H. A meta-analysis of outcome studies comparing bona fide psychotherapies: empirically, "all must have prizes". *Psychol Bull* 1997; **112:** 203–15.

60 Marcus DK, O'Connell D, Norris AL, Sawaqdeh A. Is the dodo bird endangered in the 21st century? A meta-analysis of treatment comparison studies. *Clin Psychol Rev* 2014; **34:** 519–30.

61 Lee KH, Siegle GJ, Dahl RE, Hooley JM, Silk JS. Neural responses to maternal criticism in healthy youth. *Soc Cogn Affect Neurosci* 2015; **10:** 902–12.

62 Ressler KJ, Rothbaum BO, Tannenbaum L, et al. Cognitive enhancers as adjuncts to psychotherapy: use of d-cycloserine in phobic individuals to facilitate extinction of fear. *Arch Gen Psychiatry* 2004; **61:** 1136–44.

63 Otto MW, Kredlow MA, Smits JA, et al. Enhancement of psychosocial treatment with d-cycloserine: models, moderators, and future directions. *Biol Psychiatry* 2016; **80:** 274–83.

64 Nader K, Schafe GE, Le Doux JE. Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature* 2000; **406:** 722–26.

65 James EL, Bonsall MB, Hoppitt L, et al. Computer game play reduces intrusive memories of experimental trauma via reconsolidation-update mechanisms. *Psychol Sci* 2015; **26:** 1201–15.

66 Kindt M. A behavioural neuroscience perspective on the aetiology and treatment of anxiety disorders. *Behav Res Ther* 2014; **62:** 24–36.

67 Treanor M, Brown LA, Rissman J, Craske MG. Can memories of traumatic experiences or addiction be erased or modified? A critical review of research on the disruption of memory reconsolidation and its applications. *Perspect Psychol Sci* 2017; **12:** 290–305.

68 Davis M, Ressler K, Rothbaum BO, Richardson R. Effects of d-cycloserine on extinction: translation from preclinical to clinical work. *Biol Psychiatry* 2006; **60:** 369–75.

69 Schramm MJ, Everitt BJ, Milton AL. Bidirectional modulation of alcohol-associated memory reconsolidation through manipulation of adrenergic signaling. *Neuropsychopharmacology* 2016; **41:** 1103–11.

70 Marvar PJ, Goodman J, Fuchs S, Choi DC, Banerjee S, Ressler KJ. Angiotensin type 1 receptor inhibition enhances the extinction of fear memory. *Biol Psychiatry* 2014; **75:** 864–72.

71 Kalueff AV, Stewart AM, Song C, Berridge KC, Graybiel AM, Fentress JC. Neurobiology of rodent self-grooming and its value for translational neuroscience. *Nat Rev Neurosci* 2016; **17:** 45–59.

72 Hales CA, Stuart SA, Anderson MH, Robinson ES. Modelling cognitive affective biases in major depressive disorder using rodents. *Br J Pharmacol* 2014; **171:** 4524–38.

73 Parker RM, Paul ES, Burman OH, Browne WJ, Mendl M. Housing conditions affect rat responses to two types of ambiguity in a reward-reward discrimination cognitive bias task. *Behav Brain Res* 2014; **274:** 73–83.

74 Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013; **46:** 81–95.

75 England MJ, Stith-Butler A, Gonzalez ML. Psychosocial interventions for mental and substance use disorders. A framework for establishing evidence-based standards. Washington, DC: National Academies Press (US), 2015.

76 Murray CJ, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; **380:** 2197–223.

77 Lawrence D, Kisely S, Pais J. The epidemiology of excess mortality in people with mental illness. *Can J Psychiatry* 2010; **55:** 752–60.

78 Wang PS, Angermeyer M, Borges G, et al. Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry* 2007; **6:** 177–85.

79 Patel V, Chowdhary N, Rahman A, Verdeli H. Improving access to psychological treatments: lessons from developing countries. *Behav Res Ther* 2011; **49:** 523–28.

80 WHO. Mental health atlas: 2011. Geneva: World Health Organization, 2011.

81 Saraceno B, van Ommeren M, Batniji R, et al. Barriers to improvement of mental health services in low-income and middle-income countries. *Lancet* 2007; **370:** 1164–74.

82 Vervliet B, Craske MG, Hermans D. Fear extinction and relapse: state of the art. *Annu Rev Clin Psychol* 2013; **9:** 215–48.

83 Kazdin AE, Rabbitt SM. Novel models for delivering mental health services and reducing the burden of mental illness. *Clin Psychol Sci* 2013; **1:** 170–91.

84 Bass J, Neugebauer R, Clougherty KF, et al. Group interpersonal psychotherapy for depression in rural Uganda: 6-month outcomes: randomised controlled trial. *Br J Psychiatry* 2006; **188:** 567–73.

85 Rahman A, Malik A, Sikander S, Roberts C, Creed F. Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: a cluster-randomised controlled trial. *Lancet* 2008; **372:** 902–09.

86 Dawson KS, Bryant RA, Harper M, et al. Problem Management Plus (PM+): a WHO transdiagnostic psychological intervention for common mental health problems. *World Psychiatry* 2015; **14:** 354–57.

87 Craske MG, Rose RD, Lang A, et al. Computer-assisted delivery of cognitive behavioral therapy for anxiety disorders in primary-care settings. *Depress Anxiety* 2009; **26:** 235–42.

88 Andersson G, Titov N. Advantages and limitations of internet-based interventions for common mental disorders. *World Psychiatry* 2014; **13:** 4–11.

89 Arjadi R, Nauta MH, Chowdhary N, Bockting CLH. A systematic review of online interventions for mental health in low and middle income countries: a neglected field. *Glob Ment Health (Camb)* 2015; **2:** e12.

90 Poushter J. Smartphone ownership and internet usage continues to climb in emerging economies. Washington, DC: Pew Research Center, 2016. http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/ (accessed April 29, 2016).

91 Hay P, Chinn D, Forbes D, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for the treatment of eating disorders. *Aust N Z J Psychiatry* 2014; **48:** 977–1008.

92 Ljotsson B, Lundin C, Mitsell K, Carlbring P, Ramklint M, Ghaderi A. Remote treatment of bulimia nervosa and binge eating disorder: a randomized trial of Internet-assisted cognitive behavioural therapy. *Behav Res Ther* 2007; **45:** 649–61.

93 Enebrink P, Högström J, Forster M, Ghaderi A. Internet-based parent management training: a randomized controlled study. *Behav Res Ther* 2012; **50:** 240–49.

94 Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. *World Psychiatry* 2014; **13:** 288–95.

95 Eaton J, McCay L, Semrau M, et al. Scale up of services for mental health in low-income and middle-income countries. *Lancet* 2011; **378:** 1592–603.

96 Nickerson A, Bryant RA, Silove D, Steel Z. A critical review of psychological treatments of posttraumatic stress disorder in refugees. *Clin Psychol Rev* 2011; **31:** 399–417.

97 Singhal A, Rogers EM. Entertainment-education: a communication strategy for social change. Mahwah, NJ: Erlbaum, 1999.

98 Knapp M, McDaid D, Parsonage M. Mental health promotion and mental illness prevention: the economic case. London: Department of Health, 2011.

99 Ori R, Amos T, Bergman H, Soares-Weiser K, Ipser JC, Stein DJ. Augmentation of cognitive and behavioural therapies (CBT) with d-cycloserine for anxiety and related disorders. *Cochrane Database Syst Rev* 2015; **10:** CD007803.

100 Otto MW, Bruce SE, Deckersbach T. Benzodiazepine use, cognitive impairment, and cognitive-behavioral therapy for anxiety disorders: issues in the treatment of a patient in need. *J Clin Psychiatry* 2005; **66** (suppl 2)**:** 34–38.

101 Cuijpers P, Sijbrandij M, Koole SL, Andersson G, Beekman AT, Reynolds CF 3rd. Adding psychotherapy to antidepressant medication in depression and anxiety disorders: a meta-analysis. *World Psychiatry* 2014; **13:** 56–67.

102 Cuijpers P, van Straten A, Warmerdam L, Andersson G. Psychotherapy versus the combination of psychotherapy and pharmacotherapy in the treatment of depression: a meta-analysis. *Depress Anxiety* 2009; **26:** 279–88.

103 Hollon SD, Jarrett RB, Nierenberg AA, Thase ME, Trivedi M, Rush AJ. Psychotherapy and medication in the treatment of adult and geriatric depression: which monotherapy or combined treatment? *J Clin Psychiatry* 2005; **66:** 455–68.

104 Barlow DH, Gorman JM, Shear MK, Woods SW. Cognitive-behavioral therapy, imipramine, or their combination for panic disorder: a randomized controlled trial. *JAMA* 2000; **283:** 2529–36.

105 Merlo E, Milton AL, Goozée ZY, Theobald DE, Everitt BJ. Reconsolidation and extinction are dissociable and mutually exclusive processes: behavioral and molecular evidence. *J Neurosci* 2014; **34:** 2422–31.

106 Guhn A, Dresler T, Andreatta M, et al. Medial prefrontal cortex stimulation modulates the processing of conditioned fear. *Front Behav Neurosci* 2014; **8:** 44.

107 Kennedy SH, Konarski JZ, Segal ZV, et al. Differences in brain glucose metabolism between responders to CBT and venlafaxine in a 16-week randomized controlled trial. *Am J Psychiatry* 2007; **164:** 778–88.

108 Harmer CJ, Duman RS, Cowen PJ. How do antidepressants work? New perspective for refining future treatment approaches. *Lancet Psychiatry* 2017; **4:** 409–18.

109 Beck AT, Rush AJ, Shaw BF, Emery G. Cognitive therapy of depression. New York, NY: Guildford, 1979.

110 Disner SG, Beevers CG, Haigh EA, Beck AT. Neural mechanisms of the cognitive model of depression. *Nat Rev Neurosci* 2011; **12:** 467–77.

111 Shiroma PR, Thuras P, Johns B, Lim KO. Emotion recognition processing as early predictor of response to 8-week citalopram treatment in late-life depression. *Int J Geriatr Psychiatry* 2014; **29:** 1132–39.

112 Moss A, Freeman TP. Mind the gap: bringing together psychopharmacology and cognitive behavioural therapy. The British Psychological Society, 2015. https://thepsychologist.bps.org.uk/mind-gap-bringing-together-psychopharmacology-and-cognitive-behavioural-therapy (accessed Oct 5, 2017).

113 Bouton ME. Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Biol Psychiatry* 2002; **52:** 976–86.

114 Murphy SE, Downham C, Cowen PJ, Harmer CJ. Direct effects of diazepam on emotional processing in healthy volunteers. *Psychopharmacology (Berl)* 2008; **199:** 503–13.

115 Reinecke A, Waldenmaier L, Cooper MJ, Harmer CJ. Changes in automatic threat processing precede and predict clinical changes with exposure-based cognitive-behavior therapy for panic disorder. *Biol Psychiatry* 2013; **73:** 1064–70.

116 Clarke PJ, Browning M, Hammond G, Notebaert L, MacLeod C. The causal role of the dorsolateral prefrontal cortex in the modification of attentional bias: evidence from transcranial direct current stimulation. *Biol Psychiatry* 2014; **76:** 946–52.

117 Browning M, Grol M, Ly V, Goodwin GM, Holmes EA, Harmer CJ. Using an experimental medicine model to explore combination effects of pharmacological and cognitive interventions for depression and anxiety. *Neuropsychopharmacology* 2011; **36:** 2689–97.

118 Başoğlu M, Marks IM, Kiliç C, Brewin CR, Swinson RP. Alprazolam and exposure for panic disorder with agoraphobia. Attribution of improvement to medication predicts subsequent relapse. *Br J Psychiatry* 1994; **164:** 652–59.

119 Cuijpers P, Beekman ATF, Reynolds CF 3rd. Preventing depression: a global priority. *JAMA* 2012; **307:** 1033–34.

120 Campion J, Bhui K, Bhugra D, European Psychiatric Association. European Psychiatric Association (EPA) guidance on prevention of mental disorders. *Eur Psychiatry* 2012; **27:** 68–80.

121 Heckman JJ. Skill formation and the economics of investing in disadvantaged children. *Science* 2006; **312:** 1900–02.

122 Kim-Cohen J, Caspi A, Moffitt TE, Harrington H, Milne BJ, Poulton R. Prior juvenile diagnoses in adults with mental disorder: developmental follow-back of a prospective-longitudinal cohort. *Arch Gen Psychiatry* 2003; **60:** 709–17.

123 National Research Council (US) and Institute of Medicine (US) Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth, and Young Adults: Research Advances and Promising Interventions. Preventing mental, emotional, and behavioural disorders among young people: progress and possibilities. Washington, DC: National Academies Press; 2009.

124 Olds DL, Holmberg JR, Donelan-McCall N, Luckey DW, Knudtson MD, Robinson J. Effects of home visits by paraprofessionals and by nurses on children: follow-up of a randomized trial at ages 6 and 9 years. *JAMA Pediatr* 2014; **168:** 114–21.

125 Rapee RM. The preventative effects of a brief, early intervention for preschool-aged children at risk for internalising: follow-up into middle adolescence. *J Child Psychol Psychiatry* 2013; **54:** 780–88.

126 Stein A, Ramchandani P, Murray L. Impact of parental psychiatric disorder and physical illness. In: Rutter M, Bishop DVM, Pine DS, et al, eds. Rutter's child and adolescent psychiatry, 5th edn. Holboken, NJ: Blackwell Publishers, 2009: 407–20.

127 Carter JC, Stewart DA, Dunn VJ, Fairburn CG. Primary prevention of eating disorders: might it do more harm than good? *Int J Eat Disord* 1997; **22:** 167–72.

128 Fonagy P. Psychotherapy research: do we know what works for whom? *Br J Psychiatry* 2010; **197:** 83–85.

129 Zhou X, Hetrick SE, Cuijpers P, et al. Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: a systematic review and network meta-analysis. *World Psychiatry* 2015; **14:** 207–22.

130 Robling M, Bekkers MJ, Bell K, et al. Effectiveness of a nurse-led intensive home-visitation programme for first-time teenage mothers (Building Blocks): a pragmatic randomised controlled trial. *Lancet* 2016; **387:** 146–55.

131 Ewing DL, Dash S, Thompson EJ, et al. No significant evidence of cognitive biases for emotional stimuli in children at-risk of developing anxiety disorders. *J Abnorm Child Psychol* 2016; **44:** 1243–52.

132 Verhage ML, Schuengel C, Madigan S, et al. Narrowing the transmission gap: a synthesis of three decades of research on intergenerational transmission of attachment. *Psychol Bull* 2016; **142:** 337–66.

133 Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MF. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. *BMJ* 2012; **344:** e2598.

134 Araya R, Fritsch R, Spears M, et al. School intervention to improve mental health of students in Santiago, Chile: a randomized clinical trial. *JAMA Pediatr* 2013; **167:** 1004–10.

135 Juffer F, Bakermans-Kranenburg MJ, van IJzendoorn MH. Promoting positive parenting: an attachment-based intervention. New York, NY: Taylor & Francis Group, 2008.

136 Scott S, Briskman J, O'Connor TG. Early prevention of antisocial personality: long-term follow-up of two randomized controlled trials comparing indicated and selective approaches. *Am J Psychiatry* 2014; **171:** 649–57.

137 Dretzke J, Frew E, Davenport C, et al. The effectiveness and cost-effectiveness of parent training/education programmes for the treatment of conduct disorder, including oppositional defiant disorder, in children. *Health Technol Assess* 2005; **9:** 1–233.

138 Cuijpers P, van Straten A, Smit F, Mihalopoulos C, Beekman A. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. *Am J Psychiatry* 2008; **165:** 1272–80.

139 van Zoonen K, Buntrock C, Ebert DD, et al. Preventing the onset of major depressive disorder: a meta-analytic review of psychological interventions. *Int J Epidemiol* 2014; **43:** 318–29.

140 van der Gaag M, Smit F, Bechdolf A, et al. Preventing a first episode of psychosis: meta-analysis of randomized controlled prevention trials of 12 month and longer-term follow-ups. *Schizophr Res* 2013; **149:** 56–62.

141 McGorry PD, Yung AR, Phillips LJ, et al. Randomized controlled trial of interventions designed to reduce the risk of progression to first-episode psychosis in a clinical sample with subthreshold symptoms. *Arch Gen Psychiatry* 2002; **59:** 921–28.

142 De Silva MJ, Ryan G. Global mental health in 2015: 95% implementation. *Lancet Psychiatry* 2016; **3:** 15–17.

143 Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cogn Behav Ther* 2009; **38:** 196–205.

144 Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. *Clin Psychol Rev* 2012; **32:** 329–42.

145 Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PLoS One* 2010; **5:** e13196.

146 Cheng SK, Dizon J. Computerised cognitive behavioural therapy for insomnia: a systematic review and meta-analysis. *Psychother Psychosom* 2012; **81:** 206–16.

147 Beintner I, Jacobi C, Schmidt UH. Participation and outcome in manualized self-help for bulimia nervosa and binge eating disorder—a systematic review and metaregression analysis. *Clin Psychol Rev* 2014; **34:** 158–76.

148 Riper H, Blankers M, Hadiwijaya H, et al. Effectiveness of guided and unguided low-intensity internet interventions for adult alcohol misuse: a meta-analysis. *PLoS One* 2014; **9:** e99912.

149 Andersson G, Cuijpers P. Pros and cons of online cognitive-behavioural therapy. *Br J Psychiatry* 2008; **193:** 270–71.

150 van Ballegooijen W, Cuijpers P, van Straten A, et al. Adherence to Internet-based and face-to-face cognitive behavioural therapy for depression: a meta-analysis. *PLoS One* 2014; **9:** e100674.

151 Sucala M, Schnur JB, Constantino MJ, Miller SJ, Brackman EH, Montgomery GH. The therapeutic relationship in e-therapy for mental health: a systematic review. *J Med Internet Res* 2012; **14:** e110.

152 Wenze SJ, Miller IW. Use of ecological momentary assessment in mood disorders research. *Clin Psychol Rev* 2010; **30:** 794–804.

153 Smits N, Zitman FG, Cuijpers P, den Hollander-Gijsman ME, Carlier IV. A proof of principle for using adaptive testing in routine outcome monitoring: the efficiency of the mood and anxiety symptoms questionnaire—anhedonic depression CAT. *BMC Med Res Methodol* 2012; **12:** 4.

154 Powers MB, Emmelkamp PMG. Virtual reality exposure therapy for anxiety disorders: a meta-analysis. *J Anxiety Disord* 2008; **22:** 561–69.

155 Turner WA, Casey LM. Outcomes associated with virtual reality in psychological interventions: where are we now? *Clin Psychol Rev* 2014; **34:** 634–44.

156 Mohr DC, Vella L, Hart S, Heckman T, Simon G. The effect of telephone-administered psychotherapy on symptoms of depression and attrition: a meta-analysis. *Clin Psychol (New York)* 2008; **15:** 243–53.

157 Anthes E. Mental health: there's an app for that. *Nature* 2016; **532:** 20–23.

158 Donker T, Petrie K, Proudfoot J, Clarke J, Birch MR, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. *J Med Internet Res* 2013; **15:** e247.

159 Espie CA, Kyle SD, Williams C, et al. A randomized, placebo-controlled trial of online cognitive behavioral therapy for chronic insomnia disorder delivered via an automated media-rich web application. *Sleep* 2012; **35:** 769–81.

160 Bucci S, Barrowclough C, Ainsworth J, et al. Using mobile technology to deliver a cognitive behaviour therapy-informed intervention in early psychosis (Actissist): study protocol for a randomised controlled trial. *Trials* 2015; **16:** 404.

161 Mohr DC, Schueller SM, Riley WT, et al. Trials of intervention principles: evaluation methods for evolving behavioral intervention technologies. *J Med Internet Res* 2015; **17:** e166.

162 Therapist-free therapy. *The Economist* 2011; **398:** 82–83.

163 Christensen H, Cuijpers P, Reynolds CF 3rd. Changing the direction of suicide prevention research: a necessity for true population impact. *JAMA Psychiatry* 2016; **73:** 435–36.

164 National Institute for Heath and Care Excellence. Psychosis and schizophrenia in children and young people: recognition and management. National Institute for Heath and Care Excellence, 2013. https://www.nice.org.uk/guidance/cg155 (accessed Oct 5, 2017).

165 Bradley HA, Rucklidge JJ, Mulder RT. A systematic review of trial registration and selective outcome reporting in psychotherapy randomized controlled trials. *Acta Psychiatr Scand* 2017; **135:** 65–77.

166 Arnberg FK, Alaie I, Parling T, Jonsson U. Recent randomized controlled trials of psychological interventions in healthcare: a review of their quantity, scope, and characteristics. *J Psychosom Res* 2013; **75:** 401–08.

167 Dunn G, Emsley R, Liu H, et al. Evaluation and validation of social and psychological markers in randomised trials of complex interventions in mental health: a methodological research programme. *Health Technol Assess* 2015; **19:** 1–115.

168 Longden E, Read J. Assessing and reporting the adverse effects of antipsychotic medication: a systematic review of clinical studies, and prospective, retrospective, and cross-sectional research. *Clin Neuropharmacol* 2016; **39:** 29–39.

169 Jonsson U, Alaie I, Parling T, Arnberg FK. Reporting of harms in randomized controlled trials of psychological interventions for mental and behavioral disorders: a review of current practice. *Contemp Clin Trials* 2014; **38:** 1–8.

170 Rozental A, Kottorp A, Boettcher J, Andersson G, Carlbring P. Negative effects of psychological treatments: An exploratory factor analysis of the negative effects questionnaire for monitoring and reporting adverse and unwanted events. *PLoS One* 2016; **11:** e0157503.

171 Boettcher J, Rozental A, Andersson G, Carlbring P. Side effects in internet-based interventions for social anxiety disorder. *Internet Interv* 2014; **1:** 3–11.

172 Linden M. How to define, find and classify side effects in psychotherapy: from unwanted events to adverse treatment reactions. *Clin Psychol Psychother* 2013; **20:** 286–96.

173 Mayo-Wilson E, Montgomery P, Hopewell S, Macdonald G, Moher D, Grant S. Developing a reporting guideline for social and psychological intervention trials. *Br J Psychiatry* 2013; **203:** 250–54.

174 Munder T, Brütsch O, Leonhart R, Gerger H, Barth J. Researcher allegiance in psychotherapy outcome research: an overview of reviews. *Clin Psychol Rev* 2013; **33:** 501–11.

175 Lieb K, von der Osten-Sacken J, Stoffers-Winterling J, Reiss N, Barth J. Conflicts of interest and spin in reviews of psychological therapies: a systematic review. *BMJ Open* 2016; **6:** e010606.

176 van der Lem R, de Wever WW, van der Wee NJ, van Veen T, Cuijpers P, Zitman FG. The generalizability of psychotherapy efficacy trials in major depressive disorder: an analysis of the influence of patient selection in efficacy trials on symptom outcome in daily practice. *BMC Psychiatry* 2012; **12:** 192.

177 Freeman D, Dunn G, Startup H, et al. Effects of cognitive behaviour therapy for worry on persecutory delusions in patients with psychosis (WIT): a parallel, single-blind, randomised controlled trial with a mediation analysis. *Lancet Psychiatry* 2015; **2:** 305–13.

178 Sensky T, Turkington D, Kingdon D, et al. A randomized controlled trial of cognitive-behavioral therapy for persistent symptoms in schizophrenia resistant to medication. *Arch Gen Psychiatry* 2000; **57:** 165–72.

179 Cuijpers P, Driessen E, Hollon SD, van Oppen P, Barth J, Andersson G. The efficacy of non-directive supportive therapy for adult depression: a meta-analysis. *Clin Psychol Rev* 2012; **32:** 280–91.

180 Turner DT, van der Gaag M, Karyotaki E, Cuijpers P. Psychological interventions for psychosis: a meta-analysis of comparative outcome studies. *Am J Psychiatry* 2014; **171:** 523–38.

181 Morrison AP, French P, Stewart S, et al. Early detection and intervention evaluation for people at risk of psychosis: multisite randomised controlled trial. *BMJ* 2012; **344:** e2233.

182 Morrison AP, Shryane N, Beck R, et al. Psychosocial and neuropsychiatric predictors of subjective recovery from psychosis. *Psychiatry Res* 2013; **208:** 203–09.

183 Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ* 2010; **340:** c221.

184 Vittengl JR, Jarrett RB, Weitz E, et al. Divergent outcomes in cognitive-behavioral therapy and pharmacotherapy for adult depression. *Am J Psychiatry* 2016; **173:** 481–90.

185 Weitz ES, Hollon SD, Twisk J, et al. Baseline depression severity as moderator of depression outcomes between cognitive behavioral therapy vs pharmacotherapy: an individual patient data meta-analysis. *JAMA Psychiatry* 2015; **72:** 1102–09.

186 Keeley T, Khan H, Pinfold V, et al. Core outcome sets for use in effectiveness trials involving people with bipolar and schizophrenia in a community-based setting (PARTNERS2): study protocol for the development of two core outcome sets. *Trials* 2015; **16:** 47.

187 Morrison AP, Turkington D, Pyle M, et al. Cognitive therapy for people with schizophrenia spectrum disorders not taking antipsychotic drugs: a single-blind randomised controlled trial. *Lancet* 2014; **383:** 1395–403.

188 Ennis L, Wykes T. Impact of patient involvement in mental health research: longitudinal study. *Br J Psychiatry* 2013; **203:** 381–86.

189 Brett J, Staniszewska S, Mockford C, et al. A systematic review of the impact of patient and public involvement on service users, researchers and communities. *Patient* 2014; **7:** 387–95.

190 Lloyd K, Rose D, Fenton M. Identifying uncertainties about the effects of treatments for schizophrenia. *J Ment Health* 2006; **15:** 263–68.

191 Law H, Morrison AP. Recovery in psychosis: a Delphi study with experts by experience. *Schizophr Bull* 2014; **40:** 1347–55.

192 Byrne R, Morrison AP. Service users' priorities and preferences for treatment of psychosis: a user-led Delphi study. *Psychiatr Serv* 2014; **65:** 1167–69.

193 Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986; **51:** 1173–82.

194 VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology* 2009; **20:** 18–26.

195 Emsley R, Dunn G, White IR. Mediation and moderation of treatment effects in randomised controlled trials of complex interventions. *Stat Methods Med Res* 2010; **19:** 237–70.

196 Freeman D, Dunn G, Garety P, et al. Patients' beliefs about the causes, persistence and control of psychotic experiences predict take-up of effective cognitive behaviour therapy for psychosis. *Psychol Med* 2013; **43:** 269–77.

197 Goldsmith LP, Lewis SW, Dunn G, Bentall RP. Psychological treatments for early psychosis can be beneficial or harmful, depending on the therapeutic alliance: an instrumental variable analysis. *Psychol Med* 2015; **45:** 2365–73.

198 Parmar MKB, Carpenter J, Sydes MR. More multiarm randomised trials of superiority are needed. *Lancet* 2014; **384:** 283–84.

199 Relton C, Torgerson D, O'Cathain A, Nicholl J. Rethinking pragmatic randomised controlled trials: introducing the "cohort multiple randomised controlled trial" design. *BMJ* 2010; **340:** c1066.

200 Frank G. The Boulder Model: history, rationale, and critique. *Prof Psychol Res Pr* 1984; **15:** 417–35.

201 Abramowitz JS, Deacon BJ, Whiteside SPH. Exposure therapy for anxiety: principles and practice. New York, NY: Guilford Press, 2012.

202 Mowrer OH. Two-factor learning theory: versions one and two. Learning theory and behavior. Hoboken, NJ: John Wiley & Sons, 1960.

203 Foa EB. Cognitive behavioral therapy of obsessive-compulsive disorder. *Dialogues Clin Neurosci* 2010; **12:** 199–207.

204 Abramson LY, Seligman ME, Teasdale JD. Learned helplessness in humans: critique and reformulation. *J Abnorm Psychol* 1978; **87:** 49–74.

205 Abramson LY, Alloy LB, Metalsky GI. Hopelessness depression—a theory-based subtype of depression. *Psychol Rev* 1989; **96:** 358–72.

206 Seligman ME, Maier SF, Geer JH. Alleviation of learned helplessness in the dog. *J Abnorm Psychol* 1968; **73:** 256–62.

207 Liu RT, Kleiman EM, Nestor BA, Cheek SM. The hopelessness theory of depression: a quarter century in review. *Clin Psychol (New York)* 2015; **22:** 345–65.

208 Pryce CR, Azzinnari D, Spinelli S, Seifritz E, Tegethoff M, Meinlschmidt G. Helplessness: a systematic translational review of theory and evidence for its relevance to understanding and treating depression. *Pharmacol Ther* 2011; **132:** 242–67.

209 Roffman JL, Simon AB, Prasad KM, Truman CJ, Morrison J, Ernst CL. Neuroscience in psychiatry training: how much do residents need to know? *Am J Psychiatry* 2006; **163:** 919–26.

210 Fung LK, Akil M, Widge A, Roberts LW, Etkin A. Attitudes toward neuroscience education in psychiatry: a national multi-stakeholder survey. *Acad Psychiatry* 2015; **39:** 139–46.

211 The Academy of Medical Sciences. Strengthening academic psychiatry in the UK. London; The Academy of Medical Science, 2013.

212 Goodwin GM, Holmes EA, Andersson E, et al. From neuroscience to evidence based psychological treatments—the promise and the challenge, ECNP March 2016, Nice, France. *Eur Neuropsychopharmacol* 2018; **28:** 321–37.

213 Schwabe L, Nader K, Pruessner JC. Reconsolidation of human memory: brain mechanisms and clinical relevance. *Biol Psychiatry* 2014; **76:** 274–80.

214 Brunet A, Orr SP, Tremblay J, Robertson K, Nader K, Pitman RK. Effect of post-retrieval propranolol on psychophysiologic responding during subsequent script-driven traumatic imagery in post-traumatic stress disorder. *J Psychiatr Res* 2008; **42:** 503–06.

215 Brunet A, Poundja J, Tremblay J, et al. Trauma reactivation under the influence of propranolol decreases posttraumatic stress symptoms and disorder: 3 open-label trials. *J Clin Psychopharmacol* 2011; **31:** 547–50.

216 Iyadurai L, Blackwell SE, Meiser-Stedman R, et al. Preventing intrusive memories after trauma via a brief intervention involving Tetris computer game play in the emergency department: a proof-of-concept randomized controlled trial. *Mol Psychiatry* 2017; published online March 28. DOI:10.1038/mp.2017.23.

217 Horsch A, Vial Y, Favrod C, et al. Reducing intrusive traumatic memories after emergency caesarean section: a proof-of-principle randomized controlled study. *Behav Res Ther* 2017; **94:** 36–47.

218 Uher R, Perlis RH, Henigsberg N, et al. Depression symptom dimensions as predictors of antidepressant treatment outcome: replicable evidence for interest-activity symptoms. *Psychol Med* 2012; **42:** 967–80.

219 Eshel N, Roiser JP. Reward and punishment processing in depression. *Biol Psychiatry* 2010; **68:** 118–24.

220 Gray JA. The psychophysiological basis of introversion-extraversion. *Behav Res Ther* 1970; **8:** 249–66.

221 Craske MG, Meuret AE, Ritz T, Treanor M, Dour HJ. Treatment for anhedonia: a neuroscience driven approach. *Depress Anxiety* 2016; **33:** 927–38.

222 Dimidjian S, Barrera M Jr, Martell C, Muñoz RF, Lewinsohn PM. The origins and current status of behavioral activation treatments for depression. *Annu Rev Clin Psychol* 2011; **7:** 1–38.

223 Dichter GS, Felder JN, Petty C, Bizzell J, Ernst M, Smoski MJ. The effects of psychotherapy on neural responses to rewards in major depression. *Biol Psychiatry* 2009; **66:** 886–97.

224 Blackwell SE, Browning M, Mathews A, et al. Positive imagery-based cognitive bias modification as a web-based treatment tool for depressed adults: a randomized controlled trial. *Clin Psychol Sci* 2015; **3:** 91–111.

225 Williams AD, O'Moore K, Blackwell SE, Smith J, Holmes EA, Andrews G. Positive imagery cognitive bias modification (CBM) and internet-based cognitive behavioral therapy (iCBT): a randomized controlled trial. *J Affect Disord* 2015; **178:** 131–41.

226 National Collaborating Centre for Mental Health. Depression: the NICE guideline on the treatment and management of depression in adults, updated edition. Great Britain: The British Psychological Society and The Royal College of Psychiatrists, 2010.

227 American Psychiatric Association. Practice guideline for the treatment of patients with major depressive disorder, 3rd edn. Washington, DC: American Psychiatric Association, 2010.

228 Cramer AO, Waldorp LJ, van der Maas HL, Borsboom D. Comorbidity: a network perspective. *Behav Brain Sci* 2010; **33:** 137–50.

229 Lewinsohn PM, Zinbarg R, Seeley JR, Lewinsohn M, Sack WH. Lifetime comorbidity among anxiety disorders and between anxiety disorders and other mental disorders in adolescents. *J Anxiety Disord* 1997; **11:** 377–94.

230 Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005; **62:** 617–27.

231 Borsboom D, Cramer AO. Network analysis: an integrative approach to the structure of psychopathology. *Annu Rev Clin Psychol* 2013; **9:** 91–121.

232 van der Maas HL, Molenaar PC. Stagewise cognitive development: an application of catastrophe theory. *Psychol Rev* 1992; **99:** 395–417.

233 Borsboom D, Rhemtulla M, Cramer AO, van der Maas HL, Scheffer M, Dolan CV. Kinds versus continua: a review of psychometric approaches to uncover the structure of psychiatric constructs. *Psychol Med* 2016; **46:** 1567–79.

234 Bockting CLH. My optimism wears heavy boots: towards empirically driven tailored interventions, illustrated by depression research. Utrecht; Communicatie Service Centrum, Universiteit Utrecht, 2015 (in Dutch).

235 Kuppens P, Oravecz Z, Tuerlinckx F. Feelings change: accounting for individual differences in the temporal dynamics of affect. *J Pers Soc Psychol* 2010; **99:** 1042–60.

236 GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 2017; **390:** 1211–59.

237 Merikangas KR, He JP, Burstein M, et al. Service utilization for lifetime mental disorders in U.S. adolescents: results of the National Comorbidity Survey-Adolescent Supplement (NCS-A). *J Am Acad Child Adolesc Psychiatry* 2011; **50:** 32–45.

238 Nock MK, Hwang I, Sampson NA, Kessler RC. Mental disorders, comorbidity and suicidal behavior: results from the National Comorbidity Survey Replication. *Mol Psychiatry* 2010; **15:** 868–76.

239 Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005; **62:** 593–602.

240 Moffitt TE, Harrington H, Caspi A, et al. Depression and generalized anxiety disorder: cumulative and sequential comorbidity in a birth cohort followed prospectively to age 32 years. *Arch Gen Psychiatry* 2007; **64:** 651–60.

241 Zinbarg RE, Mineka S, Craske MG, et al. The Northwestern-UCLA youth emotion project: associations of cognitive vulnerabilities, neuroticism and gender with past diagnoses of emotional disorders in adolescents. *Behav Res Ther* 2010; **48:** 347–58.

242 Caspi A, Houts RM, Belsky DW, et al. The p factor: one general psychopathology factor in the structure of psychiatric disorders? *Clin Psychol Sci* 2014; **2:** 119–37.

243 Prenoveau JM, Zinbarg RE, Craske MG, Mineka S, Griffith JW, Epstein AM. Testing a hierarchical model of anxiety and depression in adolescents: a tri-level model. *J Anxiety Disord* 2010; **24:** 334–44.

244 Borsboom D, Cramer AO, Schmittmann VD, Epskamp S, Waldorp LJ. The small world of psychopathology. *PLoS One* 2011; **6:** e27407.

245 Stewart RE, Chambless DL. Cognitive-behavioral therapy for adult anxiety disorders in clinical practice: a meta-analysis of effectiveness studies. *J Consult Clin Psychol* 2009; **77:** 595–606.

246 Wichers M. The dynamic nature of depression: a new micro-level perspective of mental disorder that meets current challenges. *Psychol Med* 2014; **44:** 1349–60.

247 Walz LC, Nauta MH, aan het Rot M. Experience sampling and ecological momentary assessment for studying the daily lives of patients with anxiety disorders: a systematic review. *J Anxiety Disord* 2014; **28:** 925–37.

248 aan het Rot M, Hogenelst K, Schoevers RA. Mood disorders in everyday life: a systematic review of experience sampling and ecological momentary assessment studies. *Clin Psychol Rev* 2012; **32:** 510–23.

249 Serre F, Fatseas M, Swendsen J, Auriacombe M. Ecological momentary assessment in the investigation of craving and substance use in daily life: a systematic review. *Drug Alcohol Depend* 2015; **148:** 1–20.

250 Bockting CL, Elgersma HJ, van Rijsbergen GD, et al. Disrupting the rhythm of depression: design and protocol of a randomized controlled trial on preventing relapse using brief cognitive therapy with or without antidepressants. *BMC Psychiatry* 2011; **11:** 8.

251 van Os J, Lataster T, Delespaul P, Wichers M, Myin-Germeys I. Evidence that a psychopathology interactome has diagnostic value, predicting clinical needs: an experience sampling study. *PLoS One* 2014; **9:** e86652.

252 Wigman JT, van Os J, Borsboom D, et al. Exploring the underlying structure of mental disorders: cross-diagnostic differences and similarities from a network perspective using both a top-down and a bottom-up approach. *Psychol Med* 2015; **45:** 2375–87.

253 Wichers M, Lothmann C, Simons CJP, Nicolson NA, Peeters F. The dynamic interplay between negative and positive emotions in daily life predicts response to treatment in depression: a momentary assessment study. *Br J Clin Psychol* 2012; **51:** 206–22.

254 DeRubeis RJ, Cohen ZD, Forand NR, Fournier JC, Gelfand LA, Lorenzo-Luaces L. The Personalized Advantage Index: translating research on prediction into individualized treatment recommendations. A demonstration. *PLoS One* 2014; **9:** e83875.

255 Barnett S, Moonesinghe SR. Clinical risk scores to guide perioperative management. *Postgrad Med J* 2011; **87:** 535–41.

256 Politi K, Herbst RS. Lung cancer in the era of precision medicine. *Clin Cancer Res* 2015; **21:** 2213–20.

257 Kramer I, Simons CJ, Hartmann JA, et al. A therapeutic application of the experience sampling method in the treatment of depression: a randomized controlled trial. *World Psychiatry* 2014; **13:** 68–77.

258 Holmes EA, Bonsall MB, Hales SA, et al. Applications of time-series analysis to mood fluctuations in bipolar disorder to promote treatment innovation: a case series. *Transl Psychiatry* 2016; **6:** e720.

259 Keller MC, Neale MC, Kendler KS. Association of different adverse life events with distinct patterns of depressive symptoms. *Am J Psychiatry* 2007; **164:** 1521–29.

260 Costantini G, Epskamp S, Borsboom D, et al. State of the aRt personality research: a tutorial on network analysis of personality data in R. *J Res Pers* 2015; **54:** 13–29.

261 Dalle Grave R, Calugi S, Sartirana M, Fairburn CG. Transdiagnostic cognitive behaviour therapy for adolescents with an eating disorder who are not underweight. *Behav Res Ther* 2015; **73:** 79–82.

262 Fairburn CG, Cooper Z, Doll HA, et al. Transdiagnostic cognitive-behavioral therapy for patients with eating disorders: a two-site trial with 60-week follow-up. *Am J Psychiatry* 2009; **166:** 311–19.

263 Farchione TJ, Fairholme CP, Ellard KK, et al. Unified protocol for transdiagnostic treatment of emotional disorders: a randomized controlled trial. *Behav Ther* 2012; **43:** 666–78.

264 Thompson-Hollands J, Sauer-Zavala S, Barlow DH. CBT and the future of personalized treatment: a proposal. *Depress Anxiety* 2014; **31:** 909–11.

265 WHO. Preventing suicide: a global imperative. Geneva: World Health Organization, 2014.

266 McDaid D, Bonin E, Park A, et al. Making the case for investing in suicide prevention interventions: estimating the economic impact of suicide and non-fatal self harm events. *Inj Prev* 2010; **16** (suppl 1)**:** A257–58.

267 O'Connor RC, Platt S, Gordon J. Achievements and challenges in suicidology: conclusions and future directions. In: O'Connor R, Platt S, Gordon J, eds. International handbook of suicide prevention: research, policy and practice. Chichester: Wiley Blackwell, 2011: 625–42.

268 Hawton K, van Heeringen K. Suicide. *Lancet* 2009; **373:** 1372–81.

269 Hawton K, Saunders KE, O'Connor RC. Self-harm and suicide in adolescents. *Lancet* 2012; **379:** 2373–82.

270 Turecki G, Brent DA. Suicide and suicidal behaviour. *Lancet* 2016; **387:** 1227–39.

271 O'Connor RC, Nock MK. The psychology of suicidal behaviour. *Lancet Psychiatry* 2014; **1:** 73–85.

272 Brown GK, Jager-Hyman S. Evidence-based psychotherapies for suicide prevention: future directions. *Am J Prev Med* 2014; **47** (suppl 2)**:** S186–94.

273 Inagaki M, Kawashima Y, Kawanishi C, et al. Interventions to prevent repeat suicidal behavior in patients admitted to an emergency department for a suicide attempt: a meta-analysis. *J Affect Disord* 2015; **175:** 66–78.

274 O'Connor E, Gaynes BN, Burda BU, Soh C, Whitlock EP. Screening for and treatment of suicide risk relevant to primary care: a systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med* 2013; **158:** 741–54.

275 Tarrier N, Taylor K, Gooding P. Cognitive-behavioral interventions to reduce suicide behavior: a systematic review and meta-analysis. *Behav Modif* 2008; **32:** 77–108.

276 Hawton K, Witt KG, Taylor Salisbury TL, et al. Psychosocial interventions for self-harm in adults. *Cochrane Database Syst Rev* 2016; **5:** CD012189.

277 Hawton K, Witt KG, Taylor Salisbury TL, et al. Interventions for self-harm in children and adolescents. *Cochrane Database Syst Rev* 2015; **12:** CD012013.

278 Zalsman G, Hawton K, Wasserman D, et al. Suicide prevention strategies revisited: 10-year systematic review. *Lancet Psychiatry* 2016; **3:** 646–59.

279 van der Feltz-Cornelis CM, Sarchiapone M, Postuvan V, et al. Best practice elements of multilevel suicide prevention strategies: a review of systematic reviews. *Crisis* 2011; **32:** 319–33.

280 Cavanagh JT, Carson AJ, Sharpe M, Lawrie SM. Psychological autopsy studies of suicide: a systematic review. *Psychol Med* 2003; **33:** 395–405.

281 Hawton K, Saunders K, Topiwala A, Haw C. Psychiatric disorders in patients presenting to hospital following self-harm: a systematic review. *J Affect Disord* 2013; **151:** 821–30.

282 van Heeringen K. Towards a psychobiological model of the suicidal process. In: van Heeringen K, ed. Understanding suicidal behaviour: the suicidal process approach to research, treatment and prevention. Chichester: John Wiley & Sons, 2001: 136–59.

283 Bostwick JM, Pankratz VS. Affective disorders and suicide risk: a reexamination. *Am J Psychiatry* 2000; **157:** 1925–32.

284 O'Connor RC. Towards an integrated motivational–volitional model of suicidal behaviour. In: O'Connor RC, Platt S, Gordon J, eds. International handbook of suicide prevention: research, policy and practice. Chichester: Wiley Blackwell, 2011: 181–98.

285 Joiner TE. Why people die by suicide. Boston, MA: Harvard University Press, 2005.

286 Van Orden KA, Witte TK, Cukrowicz KC, Braithwaite SR, Selby EA, Joiner TE Jr. The interpersonal theory of suicide. *Psychol Rev* 2010; **117:** 575–600.

287 O'Connor RC, Smyth R, Ferguson E, Ryan C, Williams JMG. Psychological processes and repeat suicidal behavior: a four-year prospective study. *J Consult Clin Psychol* 2013; **81:** 1137–43.

288 O'Connor RC, Smyth R, Williams JMG. Intrapersonal positive future thinking predicts repeat suicide attempts in hospital-treated suicide attempters. *J Consult Clin Psychol* 2015; **83:** 169–76.

289 Klonsky ED, May AM. The Three-Step Theory (3ST): a new theory of suicide rooted in the "ideation-to-action" framework. *Int J Cogn Ther* 2015; **8:** 114–29.

290 Williams JMG. The cry of pain: understanding suicide and self-harm. London: Penguin, 2001.

291 Jobes DA. The Collaborative Assessment and Management of Suicidality (CAMS): an evolving evidence-based clinical approach to suicidal risk. *Suicide Life Threat Behav* 2012; **42:** 640–53.

292 O'Connor RC, O'Carroll RE, Ryan C, Smyth R. Self-regulation of unattainable goals in suicide attempters: a two year prospective study. *J Affect Disord* 2012; **142:** 248–55.

293 Blasco-Fontecilla H, Oquendo MA. Biomarkers of suicide: predicting the predictable? In: Courtet P (ed). Understanding suicide: from diagnosis to personalized treatment. Cham: Springer, 2016: 77–83.

294 O'Connor DB, Green JA, Ferguson E, O'Carroll RE, O'Connor RC. Cortisol reactivity and suicidal behavior: investigating the role of hypothalamic-pituitary-adrenal axis responses to stress in suicide attempters and ideators. *Psychoneuroendocrinology* 2017; **75:** 183–91.

295 Comtois KA, Linehan MM. Psychosocial treatments of suicidal behaviors: a practice-friendly review. *J Clin Psychol* 2006; **62:** 161–70.

296 Crawford MJ, Thomas O, Khan N, Kulinskaya E. Psychosocial interventions following self-harm: systematic review of their efficacy in preventing suicide. *Br J Psychiatry* 2007; **190:** 11–17.

297 Harris F, Maxwell M, O'Connor R, et al. Exploring synergistic interactions and catalysts in complex interventions: longitudinal, mixed methods case studies of an optimised multi-level suicide prevention intervention in four European countries (Ospi-Europe). *BMC Pub Health* 2016; **16:** 268.

298 Linehan MM, Comtois KA, Murray AM, et al. Two-year randomized controlled trial and follow-up of dialectical behavior therapy vs therapy by experts for suicidal behaviors and borderline personality disorder. *Arch Gen Psychiatry* 2006; **63:** 757–66.

299 Guthrie E, Kapur N, Mackway-Jones K, et al. Randomised controlled trial of brief psychological intervention after deliberate self poisoning. *BMJ* 2001; **323:** 135–38.

300 Rossouw TI, Fonagy P. Mentalization-based treatment for self-harm in adolescents: a randomized controlled trial. *J Am Acad Child Adolesc Psychiatry* 2012; **51:** 1304–13.

301 Ryberg W, Fosse R, Zahl PH, et al. Collaborative Assessment and Management of Suicidality (CAMS) compared to treatment as usual (TAU) for suicidal patients: study protocol for a randomized controlled trial. *Trials* 2016; **17:** 481.

302 Gysin-Maillart A, Schwab S, Soravia L, Megert M, Michel K. A novel brief therapy for patients who attempt suicide: a 24-months follow-up randomized controlled study of the attempted suicide short intervention program (ASSIP). *PLoS Med* 2016; **13:** e1001968.

303 Ougrin D, Tranah T, Stahl D, Moran P, Asarnow JR. Therapeutic interventions for suicide attempts and self-harm in adolescents: systematic review and meta-analysis. *J Am Acad Child Adolesc Psychiatry* 2015; **54:** 97–107.

304 Cuijpers P, de Beurs DP, van Spijker BA, Berking M, Andersson G, Kerkhof AJ. The effects of psychotherapy for adult depression on suicidality and hopelessness: a systematic review and meta-analysis. *J Affect Disord* 2013; **144:** 183–90.

305 Carter G, Page A, Large M, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guideline for the management of deliberate self-harm. *Aust N Z J Psychiatry* 2016; **50:** 939–1000.

306 Hawton K, Witt KG, Salisbury TLT, et al. Psychosocial interventions following self-harm in adults: a systematic review and meta-analysis. *Lancet Psychiatry* 2016; **3:** 740–50.

307 Platt S. Inequalities and suicidal behavior. In: O'Connor RC, Pirkis J, eds. International handbook of suicide prevention, 2nd edn. Chichester: Wiley Blackwell, 2016: 258–83.

308 Bruffaerts R, Demyttenaere K, Hwang I, et al. Treatment of suicidal people around the world. *Br J Psychiatry* 2011; **199:** 64–70.

309 Evans K, Tyrer P, Catalan J, et al. Manual-assisted cognitive-behaviour therapy (MACT): a randomized controlled trial of a brief intervention with bibliotherapy in the treatment of recurrent deliberate self-harm. *Psychol Med* 1999; **29:** 19–25.

310 Milner AJ, Carter G, Pirkis J, Robinson J, Spittal MJ. Letters, green cards, telephone calls and postcards: systematic and meta-analytic review of brief contact interventions for reducing self-harm, suicide attempts and suicide. *Br J Psychiatry* 2015; **206:** 184–90.

311 Armitage CJ, Rahim WA, Rowe R, O'Connor RC. An exploratory randomised trial of a simple, brief psychological intervention to reduce subsequent suicidal ideation and behaviour in patients admitted to hospital for self-harm. *Br J Psychiatry* 2016; **208:** 470–76.

312 O'Connor RC, Ferguson E, Scott F, et al. A brief psychological intervention to reduce repetition of self-harm in patients admitted to hospital following a suicide attempt: a randomised controlled trial. *Lancet Psychiatry* 2017; **4:** 451–60.

313 Hegerl U, Althaus D, Schmidtke A, Niklewski G. The alliance against depression: 2-year evaluation of a community-based intervention to reduce suicidality. *Psychol Med* 2006; **36:** 1225–33.

314 Chesin M, Stanley B. Risk assessment and psychosocial interventions for suicidal patients. *Bipolar Disord* 2013; **15:** 584–93.

315 de Beurs D, Kirtley O, Kerkhof A, Portzky G, O'Connor RC. The role of mobile phone technology in understanding and preventing suicidal behavior. *Crisis* 2015; **36:** 79–82.

316 Mishara B, Kerkhof A. Suicide prevention and new technologies. Evidence based practice. London: Palgrave Macmillan, 2013.

317 Cleary A. Suicidal action, emotional expression, and the performance of masculinities. *Soc Sci Med* 2012; **74:** 498–505.

318 Colucci E, Lester D, Hjelmeland H, Park BCB. Suicide and culture: understanding the context. Göttingen: Hogrefe Publishing, 2013.

319 Scourfield J. Suicidal masculinities. Sociological Research Online, 2005. http://www.socresonline.org.uk/10/2/scourfield.html (accessed Jan 17, 2018).

320 Lezine D. Lived experience and suicide prevention. In: O'Connor RC, Pirkis J, eds. International handbook of suicide prevention, 2nd edn. Chichester: Wiley, 2016: 681–95.

321 Becker CB, Zayfert C, Anderson E. A survey of psychologists' attitudes towards and utilization of exposure therapy for PTSD. *Behav Res Ther* 2004; **42:** 277–92.

322 Dijstelbloem H, Huisman F, Miedema F, Mijnhardt W. Why science does not work as it should and what to do about it., Science in Transition, 2013. http://www.scienceintransition.nl/wp-content/uploads/2013/10/Science-in-Transition-Position-Paper-final.pdf (accessed March 15, 2017).

323 Ioannidis JP. How to make more published research true. *PLoS Med* 2014; **11:** e1001747.

324 Johnsen TJ, Friborg O. The effects of cognitive behavioral therapy as an anti-depressive treatment is falling: a meta-analysis. *Psychol Bull* 2015; **141:** 747–68.

325 Bower P, Gilbody S. Stepped care in psychological therapies: access, effectiveness and efficiency. Narrative literature review. *Br J Psychiatry* 2005; **186:** 11–17.

326 Wang PS, Aguilar-Gaxiola S, Alonso J, et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *Lancet* 2007; **370:** 841–50.

327 Bockting CL, Hollon SD, Jarrett RB, Kuyken W, Dobson K. A lifetime approach to major depressive disorder: the contributions of psychological interventions in preventing relapse and recurrence. *Clin Psychol Rev* 2015; **41:** 16–26.

328 Loerinc AG, Meuret AE, Twohig MP, Rosenfield D, Bluett EJ, Craske MG. Response rates for CBT for anxiety disorders: need for standardized criteria. *Clin Psychol Rev* 2015; **42:** 72–82.

329 Rosen GM, Davison GC. Psychology should list empirically supported principles of change (ESPs) and not credential trademarked therapies or other treatment packages. *Behav Modif* 2003; **27:** 300–12.

330 Meade MS, Emch M. Medical geography, 3rd edn. New York, NY: The Guildford Press, 2012.

331 Naylor C, Parsonage M, McDaid D, Knapp M, Fossey M, Galea A. Long-term conditions and mental health. The cost of co-morbidities. London: The King's Fund and Centre for Mental Health, 2012.

332 Grant J, Cottrell R, Cluzeau F, Fawcett G. Evaluating "payback" on biomedical research from papers cited in clinical guidelines: applied bibliometric study. *BMJ* 2000; **320:** 1107–11.

333 Insel T, Cuthbert B, Garvey M, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry* 2010; **167:** 748–51.

334 Nature Editorials. Therapy deficit. *Nature* 2012; **489:** 473–74.

335 MQ Landscape Analysis. UK Mental Health Research Funding. Clerkenwell: MQ Transforming mental health, 2015.

336 Davidson L, O'Connell MJ, Tondora J, Lawless M, Evans AC. Recovery in serious mental illness: a new wine or just a new bottle? *Prof Psychol Res Pr* 2005; **36:** 480–87.

337 Rachman S. Betrayal: a psychological analysis. *Behav Res Ther* 2010; **48:** 304–11.

338 Wykes T, Haro JM, Belli SR, et al. Mental health research priorities for Europe. *Lancet Psychiatry* 2015; **2:** 1036–42.

Annual Review of Clinical Psychology

Measurement-Based and Data-Informed Psychological Therapy

Wolfgang Lutz,[1] Brian Schwartz,[1] and Jaime Delgadillo[2]

[1]Department of Psychology, University of Trier, Trier, Germany; email: lutzw@uni-trier.de

[2]Clinical and Applied Psychology Unit, Department of Psychology, University of Sheffield, Sheffield, United Kingdom

Annu. Rev. Clin. Psychol. 2022. 18:71–98

First published as a Review in Advance on December 15, 2021

The *Annual Review of Clinical Psychology* is online at clinpsy.annualreviews.org

https://doi.org/10.1146/annurev-clinpsy-071720-014821

Keywords

measurement-based psychological therapy, data-informed, routine outcome monitoring, feedback research, prediction, statistical decision making, clinical decision making, clinical navigation systems

Abstract

Outcome measurement in the field of psychotherapy has developed considerably in the last decade. This review discusses key issues related to outcome measurement, modeling, and implementation of data-informed and measurement-based psychological therapy. First, an overview is provided, covering the rationale of outcome measurement by acknowledging some of the limitations of clinical judgment. Second, different models of outcome measurement are discussed, including pre–post, session-by-session, and higher-resolution intensive outcome assessments. Third, important concepts related to modeling patterns of change are addressed, including early response, dose–response, and nonlinear change. Furthermore, rational and empirical decision tools are discussed as the foundation for measurement-based therapy. Fourth, examples of clinical applications are presented, which show great promise to support the personalization of therapy and to prevent treatment failure. Finally, we build on continuous outcome measurement as the basis for a broader understanding of clinical concepts and data-driven clinical practice in the future.

Contents

INTRODUCTION

> A whole class of loosely related errors made in the clinical case conference arises from forgetting (on the part of the psychologist) or never having learned (in the case of the psychiatrist and social worker) certain elementary statistical or psychometric principles.
>
> Meehl 1973, p. 232

The above quote from Paul Meehl's essay "Why I Do Not Attend Case Conferences" is more than 40 years old; however, the topic of measurement in clinical practice is more current than ever. In recent decades, numerous randomized clinical trials (RCTs) have helped to establish psychological therapies as effective interventions to treat a broad range of psychological disorders (Barkham & Lambert 2021, Cuijpers et al. 2019). As a consequence, the implementation of psychological

treatments has become an important asset in most health care systems around the world. However, research demonstrating their average effectiveness does not mean that psychological treatments work for all patients under all circumstances. Furthermore, negative treatment response is rarely identified or reported in clinical trials and meta-analytic reviews that focus on average effects. As a consequence, monitoring actual clinical progress in routine care and using such data to improve treatment is a necessary supplement to psychological therapy implementation and has the potential to substantially change the way we think about psychotherapy as a science.

Outcome measurement has developed impressively over the past decade. New technological developments such as the Internet and computerized data assessment and feedback tools have made outcome measurement easier to implement than in the past. Recent developments in outcome measurement make use of the information to improve the clinical decision-making process by grounding clinical practice in empirical data. For example, psychometric and demographic data can be used by clinicians to personalize the selection of therapeutic techniques and to monitor a patient's response to therapy in real time (Lutz et al. 2021). Therefore, outcome measurement can be seen as an important and integral part of clinical competence, practice, and training. This is comparable to many other areas in the health care system, where continuous monitoring of health indicators is common in day-to-day clinical practice (e.g., temperature, blood pressure). In this sense, continuous outcome measurement forms the basis of modern measurement-based and data-informed psychological therapy (Delgadillo & Lutz 2020, Lutz et al. 2022). This move toward the adoption of data-informed psychological treatment is strongly justified in view of the limitations of clinical judgment, which have become better understood in recent decades.

Approaches that Guide Decision Making in Psychological Therapy

The assessment of treatment response is an issue that therapists have been grappling with since the early history of psychotherapy. Despite their theoretical differences, influential theorists like Sigmund Freud, B.F. Skinner, Aaron Beck, Carl Rogers, John Bowlby, and others recognized that psychological disturbances cannot be observed directly; rather, they are latent (i.e., hidden) phenomena that can be inferred from behavior, introspection, and self-report. In many traditions of psychological treatment, latent change is assumed to operate at two levels: the process level (i.e., maintaining factors or conflicts) and the consequence level (i.e., symptoms, interpersonal and behavioral problems). As such, the practice of psychotherapy implicitly or explicitly involves the monitoring of these changes to determine whether treatment is working as expected and whether it will ultimately benefit the patient. The central theme of this review concerns the continuous monitoring of change, which is referred to as routine outcome monitoring in contemporary literature, and the monitoring of related processes of change.

Broadly speaking, therapists can monitor treatment response via qualitative (i.e., as part of the therapy dialogue) and quantitative methods (i.e., using validated questionnaires or idiographic measures). Such information must somehow be processed and interpreted to inform treatment decisions. Clinical decision making is based on one of two general approaches: clinical judgment or actuarial methods. The first approach relies on intuition, which is a function of the therapist's way of processing (qualitative/quantitative) data, their clinical experience, theoretical orientation, and consultation with others. This method has been referred to as informal (Grove & Meehl 1996) since it is not usually based on a structured set of decision rules or equations, and it is highly variable within (i.e., changes over time) and between therapists. The second approach, referred to as formal or algorithmic, involves using structured decision rules or equations to combine and interpret data to reach a clinical decision.

The Limitations of Clinical Intuition

In a review of decision making in the field of clinical psychology, Garb (2005) argued that psychologists' judgments are error-prone because of the influence of heuristics and biases, as defined in Tversky & Kahneman's (1974) seminal work. For example, the representativeness heuristic may partly explain the modest interrater reliability of therapists' diagnostic assessments, as these judgments are influenced by how closely a patient resembles the diagnostic prototype that each assessor uses to make a diagnosis (Evans et al. 2002). The influence of selection and confirmation biases in therapists' clinical interviews and interactions is a related issue, as therapists vary widely in the extent to which they gather key information (e.g., symptoms) to determine a diagnosis (Miller et al. 2001). Another bias that has received attention in the field of psychotherapy is clinical overoptimism, as exemplified in the classic study by Hannan et al. (2005) in which therapists' prognostic assessments of their patients significantly underestimated the number of patients who eventually made little improvement and the number of those who deteriorated after therapy (as determined using psychometric patient-reported data). In another example, Walfish et al. (2012) showed that therapists estimate that about 85% of their patients have improved or recovered, a success rate far higher than those found based on measured outcomes in clinical trials and routine care. Furthermore, 90% of clinicians rate themselves in the upper quartile of successful therapists, and none consider themselves below average. This overly positive self-assessment is known as the better-than-average effect and can also be found in other areas and professions (e.g., Zell et al. 2020). In part, these biases of clinical intuition are a result of therapists having to rely on indirect (i.e., proxy indicators of latent factors) and often subjective information (i.e., the patient's experiences narrated verbally) to make judgments. Experts in the field of decision making agree that accurate intuition can be developed in situations that provide regular and highly objective feedback, enabling judges to refine their pattern-recognition abilities (Kahneman & Klein 2009). Hence, it is not surprising that clinical intuition is not highly reliable in psychotherapy in the absence of systematic routine outcome monitoring and empirically derived decision-support tools.

Therefore, there is a strong case for the additional use of formal and empirical methods. Of course, formal methods of routine outcome monitoring are also imperfect, but they offer additional empirical data to support clinical decision making. The aim of this article is to review the history of formal routine outcome monitoring methods, their strengths and limitations, their implementation, and their future development. Over the years, several terms have been used to describe this line of research, including practice-oriented research, patient-focused research, practice-based evidence, and routine outcome monitoring (Barkham & Lambert 2021; Castonguay et al. 2013, 2021; Lutz et al. 2021). In this review, we use the terms patient-focused feedback research and measurement-based as well as data-informed psychological therapy. These terms highlight the clinical decision-support function of such endeavors and the new trend to use such data to personalize treatment (Delgadillo & Lutz 2020).

MEASURING OUTCOMES

Supporting clinical decision making with empirical data can be seen as a major advance in the field of psychotherapy. This process can rely on sparse data (i.e., collected pre- and posttherapy) or continuous and intensive data collection (i.e., collected during therapy). In this section we first present a theoretical framework for data-informed psychological therapy. Then we describe the classical pre–post treatment method before moving on to multiple assessments, session-by-session assessments, and intensive longitudinal assessments.

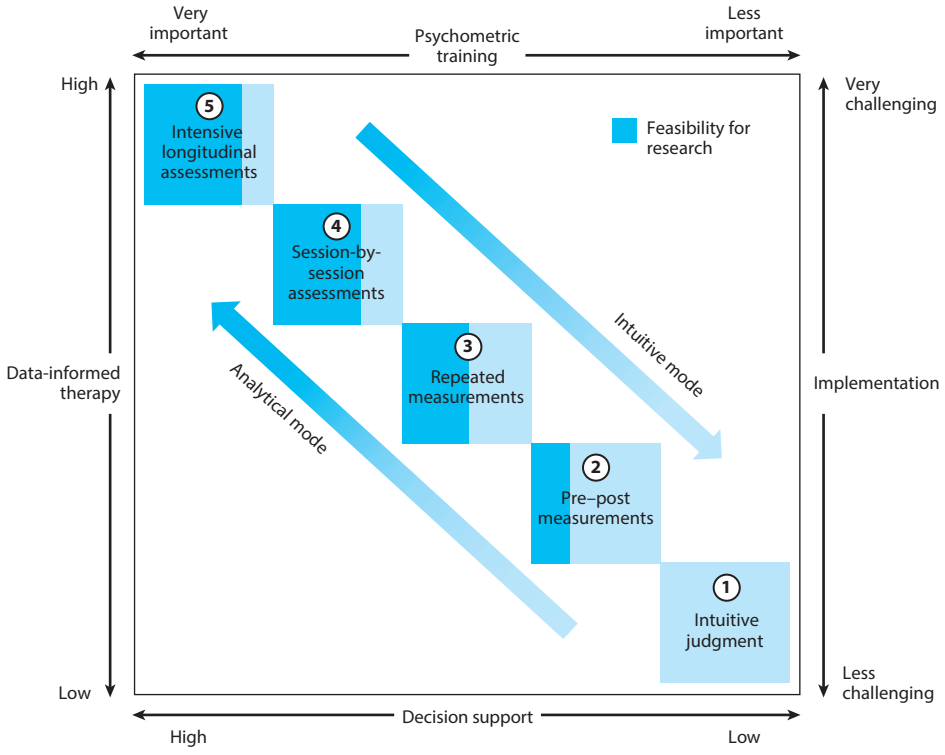


Figure 1

Matrix of assessment modes in measurement-based and data-informed psychological therapy. Darker shades of blue indicate greater feasibility for research.

Practice-Based Evidence and Clinical Decision Making

Figure 1 displays a matrix of assessment modes in clinical practice and their relation to empirically supported clinical decision making. The matrix is based on Hammond's (1978) modes of inquiry in evaluation research and includes five modes and six dimensions. It has been adapted to describe modes and dimensions of measurement-based and data-informed psychological therapy. The six dimensions include (*a*) the potential for data-informed psychological therapy based on the information provided by the assessment structure, ranging from very low to very high (left vertical axis in **Figure 1**), (*b*) the extent of empirically based support to generate practice-based evidence for clinical decision making (lower horizontal axis in **Figure 1**), (*c*) the challenge of implementing specific data collection modes into a clinical routine, ranging from less challenging to very challenging (right vertical axis in **Figure 1**), (*d*) the intensity of psychometric training, which is necessary to make use of the collected information (upper horizontal axis in **Figure 1**), (*e*) the usefulness of the data collection strategy for practice-oriented research (darker shades of blue in **Figure 1** indicate greater usefulness), and (*f*) the mode of cognition, ranging from analytical to intuitive thought (blue arrows in **Figure 1**).

Mode 1 represents the predominant use of intuitive clinical decision making, not influenced by outcome measures. This is the traditional psychotherapy mode and likely represents the most common mode to date. Therapists working in this mode adapt their clinical approach based on experience with similar patients while often simultaneously following guidelines set out by

professional organizations, employers, or policy makers. This may be a conscious choice (i.e., a preference for intuitive decision making) or may be due to lack of training, resources, or time to implement routine outcome monitoring. Nevertheless, this means that clinical decision making is mainly conducted without routine outcome assessments. Furthermore, in mode 1, empirically based decision support is nonexistent, and practice-based evidence is not systematically generated.

Pre–Post and Repeated Assessments

Mode 2 represents the first step of implementing data into clinical practice. Data collection and clinical decision making are based on a pre–post treatment data collection design (or direct collection of posttreatment information only). This mode allows conclusions to be drawn about the effectiveness of treatments under routine conditions and is relatively easy to implement in clinical practice. However, this form of assessment is often limited by substantial missing data, as patients often drop out, leave treatment early, or do not fill out the final questionnaire for other reasons. This becomes an even larger issue when longer-term follow-up assessments are added. Mode 3 includes repeated measurements over the course of treatment—for example, every four or five sessions or weeks. Such a repeated-measurement setup has greater potential to support clinical decision making and research. For example, expected treatment response (ETR) models (covered later on) can be developed and used to monitor and adjust therapy in real time. These modes require basic training and knowledge of psychometrics to make use of the data for clinical decisions regarding individual patients, groups, and services.

Session-by-Session Assessments

Most studies on routine outcome monitoring and the effectiveness of psychometric feedback in clinical practice focus on assessments on a weekly or session-by-session basis (see mode 4 in **Figure 1**). Such an assessment structure provides the context for the effective use of measurement-based treatments, especially in combination with clinical support tools (CSTs) for patients at high risk of treatment failure (e.g., Lutz et al. 2022). As indicated in **Figure 1**, here the potential for data-informed psychological treatment is high. However, more psychometric knowledge and therapist training are needed to make use of the information and integrate it into clinical practice. Of course, the implementation of a session-by-session outcome assessment structure is more challenging; however, the potential value of the generated information for further research is also high.

Furthermore, this mode includes a data-generation strategy, which has laid the foundation for several new developments in clinical research over the last decade. Usually, clinical concepts are developed based on theory, or intuitively based on expert clinical experience, and then tested in RCTs (or meta-analyses) applied to evaluate the effectiveness of treatment orientations or packages (Cristea et al. 2021). However, there are limitations to thinking about clinical concepts of psychological therapies exclusively based on RCTs and meta-analyses, as was recently pointed out by Baldwin & Imel (2020). The categorization of different variants or orientations of psychological therapy can be difficult (in comparison, for example, to most somatic treatments) and is usually based on arbitrary boundaries and theoretical arguments about the predominance of certain change processes being more or less relevant in specific forms of treatment. Therefore, this line of research requires a complementary approach that can be routinely implemented in clinical practice and that continuously uses research results to improve treatment outcomes in routine clinical practice (e.g., Howard et al. 1996, Lambert et al. 2001). Therapists not only deliver treatment but simultaneously collect data by assessing patients' progress session by session over the course of treatment (Lutz et al. 2019). Large databases generated in this way can then be analyzed

using advanced statistical analyses (e.g., growth curve modeling, multilevel modeling, machine learning algorithms) to develop clinical decision-support tools. This allows therapists not only to track progress (and potential side effects) on an individual level but also to measure the variability of change as a function of patient characteristics, treatment interventions, therapists, and clinics/services (Barkham & Lambert 2021). For example, on an aggregated level, the data can serve to investigate side effects, treatment dosage, therapist effects, and clinic effects (e.g., McKay & Jensen-Doss 2021). The session-by-session assessment structure is accompanied by statistical and methodological developments to analyze the nested structure of longitudinal data (e.g., Raudenbush & Bryk 2002). Such multilevel models allow a better disentanglement of variance components at the session, patient, therapist, and service levels. They also allow the disentanglement of between-patient and within-patient variance components to study mechanisms of change (e.g., Hamaker & Wichers 2017).

On a larger scale, such session-by-session information holds the potential to establish a community or national data collection system to which independent research groups can be granted access. Such endeavors have just begun to emerge in the field, as exemplified by the Improving Access to Psychological Therapies (IAPT) initiative in England (see Clark 2018). In combination with technological innovations, session-by-session assessments and clinically informative data-driven decision-support tools can facilitate the development of large scientist-practitioner infrastructures.

Intensive Longitudinal Assessments

Despite the advances that session-by-session assessments have made possible for data-informed therapy, in recent years, even more fine-grained measurements have been applied in psychotherapy research. Proponents of these intensive longitudinal assessments have argued that higher resolution is necessary to obtain useful information about dynamic changes in patients' everyday experiences.

Data from such high-frequency measurements can be obtained via ecological momentary assessment (EMA), a real-time, within-subject assessment allowing measurements several times a day (mode 5 in **Figure 1**). EMA has also been referred to as ambulatory assessment or experience sampling (Ebner-Priemer & Trull 2009). During the last two decades, technological innovations have drastically facilitated these measurements, allowing laborious paper-and-pencil diaries to be replaced by technological devices such as smartphones, smartwatches, and other mobile high-tech devices (Miller 2012). EMA circumvents the problem of retrospective bias, is more ecologically valid as it can be conducted within patients' daily lives, and has higher resolution and can thus better represent intrapersonal processes (Ebner-Priemer & Trull 2009). It can simultaneously collect psychological, physiological, and behavioral variables, allowing the examination of situation-specific relationships and feedback to the patient in real time. Furthermore, EMA is face-valid, convenient, and unobtrusive for patients (Miller 2012).

EMA is most often used to collect self-report questionnaires, which patients complete on their own or with the help of a smartphone. The ability to collect psychometric measures multiple times a day supports data-informed therapy and—in principle—allows for a high level of decision support (**Figure 1**). In psychotherapy research, intensive longitudinal assessments have been used to improve the prediction of early change (Husen et al. 2016), treatment response (Fisher et al. 2019, Wichers et al. 2012), and treatment dropout (Lutz et al. 2018). At the same time, more therapist training in psychometrics is necessary compared to lower-frequency assessment methods, and implementation of the assessments is more challenging, often leading to higher dropout rates during assessments. Furthermore, when implementing EMA, decisions must be made about the duration,

frequency, scope (number of items), and timing of assessments. To date, intensive longitudinal assessments have been implemented very heterogeneously, and a consensus on EMA survey design standards is lacking. However, some authors have already offered suggestions to improve transparency and user acceptability (e.g., Fisher et al. 2021). An important limitation of intensive longitudinal assessments concerns patient burden (completion of questionnaires): Longer questionnaires, rather than higher assessment frequency, increase the burden and affect data quantity and quality (Eisele et al. 2022).

In addition to the high-frequency collection of psychometric measures, EMA can be used to collect biological and physiological measures (e.g., heart rate, electrodermal activity, motor activity, sleep). For example, wristbands that collect physiological parameters are already better than chance at predicting epileptic seizures (Meisel et al. 2020). Smartwatches and associated mobile sensors enable the passive and continuous collection of huge amounts of data within patients' daily lives with minimal patient burden (e.g., Jacobson et al. 2019, Wright & Woods 2020). This type of data collection has been referred to as digital phenotyping and has recently been applied in psychotherapy research to identify diagnostic groups, symptom severity, and patterns of change as well as to inform treatment selection (e.g., Hehlmann et al. 2021, Jacobson et al. 2019, Webb et al. 2022).

Intensive longitudinal data can also help to optimize the assessment of therapy processes during treatment. This allows intraindividual changes to be assessed with a higher resolution and standardized change measurements to be supplemented by idiographic approaches. Intensive longitudinal assessments are also necessary for the continuous evaluation of video and audio recordings of a therapy session. For example, physical movements of the patient and therapist (Ramseyer & Tschacher 2011), speech content and prosodic features (Imel et al. 2014, Paz et al. 2021), and emotions based on gestures and facial expressions (Baur et al. 2020) can be recorded automatically. This allows a patient, a patient–therapist dyad, and the entire therapeutic process to be captured not only unidimensionally based on known pretherapy self-report items but also multidimensionally based on high-frequency data from different modalities (e.g., audio, video, self-report, physiology).

Recently, data-driven tools have been developed to provide therapists with personalized feedback by dynamic visualizations of intensive longitudinal data (e.g., Bringmann et al. 2021, Hehlmann et al. 2021). However, these approaches still have to contend with a number of problems in addition to the challenging implementation and large training effort. These data require the use of sophisticated statistical analysis methods, which are currently still being developed and refined (Bringmann 2021). In addition, many analytical options and decisions need to be made during data preparation and evaluation, whereby small changes can lead to significant differences in the results. Therefore, these approaches are currently more likely to be found in the context of pilot and proof-of-concept studies.

Summary

Practice-based evidence and clinical decision making are based on data assessed before, during, and after treatment. In measurement-based and data-informed psychological therapy, outcomes are measured and observed at varying frequencies: (*a*) Pre–post assessments are easy to implement but are limited by missing data and can represent only simple changes, (*b*) repeated measures throughout treatment can help to model patterns of change, to inform treatment decisions, and to enable psychometric feedback, (*c*) session-by-session assessments allow progress to be tracked on an individual level to develop CSTs and examine variability in change as a function of patient, therapist, and/or treatment characteristics, and (*d*) intensive longitudinal assessments can improve the representation of high-frequency intrapersonal processes in patients' daily lives. Currently,

however, these high-frequency measurements still face several problems, and therefore session-by-session assessment remains the best-suited method to assess treatment progress and to inform decisions in routine care.

MODELING PATTERNS OF CHANGE

Interest in how patients change over time has been documented for several decades in the field of psychotherapy, although earlier investigations were hampered by unreliable methods of assessing longitudinal patterns of change. In the last 35 years, however, a large body of empirical research on this topic has accumulated. To a great extent, the development of brief psychometric measures that could be regularly collected during psychotherapy led to important methodological and conceptual advances. Patients are not sufficiently described by their diagnoses because they are quite different from each other across several other demographic, clinical, and interpersonal characteristics that are related to treatment response (e.g., Lutz et al. 2021). These individual differences influence the longitudinal trends that characterize symptomatic change over time as well as symptom fluctuations that can be used to develop empirically supported clinical decision rules. These concepts and related findings are discussed below.

Dose–Response

Research on trajectories of change can be traced back to Kenneth Howard and colleagues' seminal examination of the dose–response effect (Howard et al. 1986). Alluding to the pharmacological notion of a relationship between the dose of medication and its expected effects, Howard et al. (1986) aggregated routine outcomes data from 2,431 psychotherapy patients across 15 studies and modeled the statistical relationship between the number of treatment sessions (dose) and symptomatic improvements. This relationship was nonlinear and characterized by a negatively accelerating (log-linear) trend, where most of the improvements were observed in the earlier sessions, showing diminishing improvements thereafter. They proposed that an optimal dose of therapy could be defined as an interval between the point at which at least 50% of cases respond to treatment and the point after which response rates plateau, which in their study was around 8–26 sessions. Conceptually, the dose–response notion is based on the assumption that therapy sessions cause change and have a cumulative effect (more sessions are better), but the potency of this effect diminishes over time. This study motivated a surge of investigations over the following decades in clinical samples with various diagnoses and treatment modalities (e.g., Baldwin et al. 2009, Barkham et al. 1996). Despite the diversity of statistical methods applied in these studies, a systematic review of over 20 dose–response studies (Robinson et al. 2020a) concluded that a curvilinear relationship between treatment duration and outcomes has been extensively replicated, although the "optimal dose" apparently varies according to clinical heterogeneity (e.g., impairment level), treatment intensities (e.g., guided self-help versus psychotherapy), and treatment settings (e.g., student counseling, outpatient care, inpatient care). Consistent with these assumptions, more recent investigations (Robinson et al. 2020b) of highly standardized treatments and more diagnostically homogeneous samples report differential dose–response patterns according to diagnosis [e.g., posttraumatic stress disorder (PTSD) requires lengthier treatment than generalized anxiety disorder] and treatment intensity [e.g., the dose effect of guided self-help plateaus sooner than that of cognitive behavioral therapy (CBT)].

Additional research has revealed that not all patients follow a uniform dose–response relation and that subgroups of patients follow different latent trends of change (e.g., Lutz et al. 2014, Owen et al. 2015). Such findings motivated the proposal of an alternative perspective, the "good-enough level" (GEL) model. Barkham et al. (1996) argued that the classic log-linear model could be a

statistical artifact that results from the aggregation of data from cases that in fact have heterogeneous treatment response patterns (i.e., early responders, gradual responders, nonresponders with lengthy treatments, and those that drop out). Rather than assuming uniformity across all patients, the GEL model assumes that the number of sessions needed to attain improvement varies from case to case, resulting in a process of "responsive regulation" of treatment duration (Barkham et al. 2006, Stiles et al. 1998). Several studies have provided support for this model, although studies comparing the goodness of fit between the dose–response and GEL models have tended to yield mixed findings (e.g., Stulz et al. 2013). A systematic review of 15 studies in this area found empirical support for some of the assumptions of the GEL model, such as the observations that higher intake severity tends to require lengthier treatments and that some subgroups of cases show curvilinear change while others show linear change (Bone et al. 2021a).

Overall, the dose–response and GEL models have empirical support and together indicate that (*a*) subgroups of patients change in similar ways, (*b*) change is most often nonlinear, (*c*) intake severity influences the linearity and duration of treatment, (*d*) although some patients respond early and others more gradually, the net benefit of therapy tends to occur within a predictable window of time, and (*e*) patients who remain in treatment beyond a typical optimal dose without showing reliable change are more likely to be nonresponders. The precise operationalization of the latter two points, however, remains a matter of controversy and debate in the field (e.g., Nordmo et al. 2020, Robinson et al. 2020a). Consequently, predicting and monitoring which patients are likely to require shorter or lengthier interventions is a goal that is empirically justified and clinically important.

Early Response

Early response refers to symptomatic improvement during the initial phase of treatment (usually the first month[1]). Over the last three decades, numerous studies have investigated early response and its prognostic value in different conditions, such as major depressive disorder, panic disorder, generalized anxiety disorder, and eating disorders, and in samples of patients with heterogeneous presenting problems (e.g., Chang et al. 2021, Delgadillo et al. 2014, Lutz et al. 2014, Moggia et al. 2020). A systematic review aggregated available data from 15 studies in a random-effects meta-analysis and reported a large pooled effect size ($g = 0.87$; Beard & Delgadillo 2019). Early responders had significantly better posttreatment outcomes compared to patients without early response. The detected effect seems to be larger in anxiety measures ($g = 1.37$) compared to depression measures ($g = 0.76$). Furthermore, early response seems to be a phenomenon that occurs in a variety of treatment models, such as CBT, the cognitive behavioral analysis system of psychotherapy (CBASP), behavioral activation, interpersonal psychotherapy (IPT), psychodynamic therapy, guided self-help, Internet-based CBT, and group therapy (Beard & Delgadillo 2019, Moggia et al. 2020).

Expected Treatment Response

The development of ETR models was key to the rise of measurement-based care. The ETR concept was first proposed by Lutz et al. (1999), who used archival data from a large routine outcome monitoring database and generated an expected recovery curve for individual patients based on

[1]Early response is often conceptualized by reliable and/or clinically significant change (Jacobson & Truax 1991). However, several studies have also used growth mixture models to identify patients' early change patterns (e.g., Lutz et al. 2014, Moggia et al. 2020).

growth curve modeling and seven predictors (well-being, symptoms, life functioning, past use of therapy, problem duration, treatment expectations, global assessment of functioning). Finch et al. (2001) applied a similar ETR model including 80% tolerance intervals around the predicted ETR curve to identify the 10% of cases that were at risk of deteriorating (outside the upper tolerance interval, also referred to as a failure boundary). This model was applied as an empirical decision rule in several progress feedback studies to identify patients who are on track (meaning their symptom change is within the expected course of treatment) or not on track (meaning their symptom trajectory is worse than expected and crosses the upper failure boundary) (e.g., Lambert 2017).

After these early applications, further ETR approaches were developed to model change and support clinical decision making. For example, Lutz et al. (2005) modeled ETR for patients using the machine learning algorithm called nearest neighbors (NN). The NN approach allows the generation of an ETR model for each patient based on a similar reference group (their nearest neighbors) with similar intake characteristics. The predictive accuracy of this advanced ETR model increased further when information on early response was included and ETR curves were dynamically adapted over the course of treatment (Bone et al. 2021b, Lutz et al. 2019). Furthermore, methods modeling latent growth trajectories for outcome and suicidal ideation (e.g., growth mixture models) have also been developed to generate ETRs based on change patterns of similar subgroups (e.g., Kyron et al. 2019, Lutz et al. 2014, Saunders et al. 2016, Stulz et al. 2013). Other variations on the ETR concept include systems that model expected change patterns for type of treatment. For example, Lutz et al. (2006) generated individual NN ETR models for patients receiving CBT and patients receiving an integrative CBT and IPT treatment to predict the optimal modality before patients begin treatment. Further ETR models have been developed to monitor and provide feedback on treatment dropout, common factors and therapeutic processes such as the patient–therapist alliance, outcome expectations, empathy, and substance use (e.g., Crits-Christoph et al. 2012, McClintock et al. 2017, Miller et al. 2005).

Nonlinear Change: Sudden Gains and Losses

As described above, an individual patient's treatment response can substantially deviate from aggregated or predicted group trends, and dismissing this phenomenon as error variance leads to a loss of information (e.g., Krause 2018). The investigation of sudden gains and losses is an area of research dealing with nonlinear change patterns (e.g., Tang & DeRubeis 1999). A sudden gain or loss is defined as a significant symptomatic positive (gain) or negative (loss) change between two consecutive sessions during the course of treatment. In their seminal work on the topic, Tang & DeRubeis (1999) identified 79% of patients with (at least one) sudden gain in a group of patients with a positive outcome in CBT for depression. In most cases, these gains occur early in treatment; this finding suggests that the phenomenon is related to the early response concept described above. However, gains also occur between later sessions, and some therapists seem to be better than others at facilitating and initiating such gains (Deisenhofer et al. 2021). Aderka et al. (2012) conducted a meta-analysis including 16 studies that identified sudden gains not only in major depressive disorder but also in social anxiety disorder, PTSD, and other anxiety problems as well as different treatment approaches (e.g., CBT, IPT, supportive expressive therapy). Overall, the analysis showed that sudden gains had an effect on primary outcome measures at posttreatment ($g = 0.62$) as well as at follow-up ($g = 0.56$). Although only a few studies have been conducted on sudden losses, their findings are still useful for the development of clinical decision rules (e.g., Krüger et al. 2014, Lutz et al. 2013). Patients with sudden losses seem to benefit less from treatment than patients with sudden gains or patients with no significant shifts. Furthermore, sudden losses are less frequent and more equally distributed over the course of treatment than sudden gains.

Monitoring Clinical Processes and Mechanisms

In most applications of measurement-based care, outcome measures are tracked and feedback is provided to therapists. Therefore, mainly treatment outcome predictors and moderators are relevant, and they are applied in the development of decision rules. While predictors and moderators provide information about for whom treatment works, mediators inform us about how treatment works (e.g., Kraemer et al. 2001). However, in recent years, several investigations using repeated assessments of process and outcome variables over the course of therapy (e.g., session by session) and multilevel modeling have allowed advancements in the establishment of mechanisms and outcome relations by separating within- and between-patient variance components of the process–outcome relation (e.g., Falkenström et al. 2016). Therefore, several within- and between-patient associations among processes (e.g., coping skills, therapeutic alliances/ruptures, emotional involvement, competence and skills) and symptom change in psychological therapy have been successfully investigated (e.g., Crits-Christoph & Connolly Gibbons 2021; Gómez Penedo et al. 2022; Lorenzo-Luaces & DeRubeis 2018; Rubel et al. 2017, 2020; Strunk et al. 2012; Zilcha-Mano 2017).

Furthermore, measures on change mechanisms have been added to applications of measurement-based care to enrich the clinical usefulness of such systems (e.g., McAleavey et al. 2021, Miller et al. 2005, Lutz et al. 2019). Several authors have recently argued that psychotherapy practice and research should focus on therapy processes instead of treatment schools or the classical illness model of psychopathology and should move toward a process-based and transdiagnostic approach with a focus on monitoring and targeting central evidence-based therapeutic procedures and processes (e.g., Goldfried 2019, Hofmann & Hayes 2019, Lutz & Schwartz 2021).

Complex Models and Critical Instability

Critical transitions represent another line of research on nonlinear patterns of change (Scheffer et al. 2009). These transitions refer to sudden changes in the state of a complex system, which thereby tips from one state (e.g., symptoms meeting diagnostic criteria for a mental disorder) to another distinct state (e.g., full remission of symptoms). This phenomenon has already been observed in ecosystems, the climate, the financial market, and neurological disorders (e.g., Lenton et al. 2008, May et al. 2008).

Conceptually comparable, critical instability can be seen as a universal early warning signal (EWS) that indicates the system's increasing instability and thus the increasing probability of a change of state (Scheffer et al. 2009). These warning signals show the critical slowing down of a system (i.e., decelerated recovery from small perturbations) that occurs when a system is approaching a transition. For example, a flurry of obsessional thoughts and related emotions may increase temporarily, immediately preceding the transition toward a state of remission of symptoms of obsessive–compulsive disorder (Heinzel et al. 2014). This can be measured in time-series data by, for example, detecting increasing autocorrelation and regression at low lags, increasing variance, changing spectral properties, and increasing skewness and kurtosis (Dakos et al. 2012). However, few studies have investigated intraindividual change in EWS (Wichers et al. 2016). For example, Olthof et al. (2020) found EWSs in a large sample to be associated with sudden gains and losses. In a feasibility study on seven cases, time-varying change point autoregressive models detected gradual and abrupt changes in continuously measured physiological stress levels that might improve outcome prediction (Hehlmann et al. 2021). However, actual intraindividual changes of EWS and their predictive power need to be investigated in future studies (Bringmann 2021).

Network models of psychopathology are also examples of complex models. The network approach assumes that symptoms (e.g., affect, thoughts, behavior) influence each other rather than

being triggered by an underlying latent disease factor (e.g., Borsboom 2017, Hofmann & Hayes 2019, Wright & Woods 2020). Network analysis models the connections (edges) between the symptoms (nodes) and allows the calculation of so-called centrality measures, which indicate the centrality of a variable within the network and thus its potential influence on other nodes. This approach was applied to identify relevant psychopathological variables via their centrality in a network and the dynamics among symptoms (Contreras et al. 2019). Beyond psychopathology, it was used in process research to identify bridges between intersession processes and symptom stress (Kaiser & Laireiter 2018) as well as in outcome research to predict treatment dropout (Lutz et al. 2018) and recovery from major depressive disorder (van Borkulo et al. 2015).

Recently, however, the fit of network models and centrality measures to psychological data has been questioned. Especially the idea of betweenness of a node, but also the assumption of symptoms as distinct nodes, has been discussed (Bringmann 2021). Furthermore, networks of symptoms have been found to be unstable—a finding that calls into question their reliability and validity—and this methodological heterogeneity challenges the comparability and interpretability of the results (Bringmann 2021). Therefore, more sophisticated models need to be developed, and the acceptability of new applications needs to be evaluated under real-world clinical conditions (Epskamp 2020). For example, in a pilot study on 12 patient–therapist dyads, Frumkin et al. (2021) found that patients were willing to participate in these assessments, were interested in the results, and tended to evaluate the data and models as useful, while therapists were less open to this method and were not yet convinced of the added value of the results.

Summary

Four findings on patterns of change are especially important for outcome monitoring in clinical practice: (*a*) Change, outcome, and dropout are influenced by patients' intake characteristics, (*b*) change is influenced by patients' early response to treatment, (*c*) change is influenced by nonlinear phenomena such as sudden gains and sudden losses, and (*d*) probably most importantly, individual patient change can vary widely, deviating from general trends or symptom trajectories derived from clinical samples. This information can be used to identify patients at risk for treatment failure.

CLINICAL DECISION TOOLS

Identifying empirically derived decision rules to support clinical decision making is one of the central goals of clinical research (e.g., Meehl 1973). In this section, we provide examples of some of the key developments arising from routinely collected data, which enable data-informed psychological therapy and decision making.

Rational Versus Empirical Decision Rules

Two distinct approaches to develop decision rules using empirical data have been applied in studies that monitor progress and tailor treatments to patients' needs (cf. Lambert et al. 2002). Rationally derived models are based on clinical assessments and the predefined classification of patients. A widely used model is based on the concepts of reliable change and clinically significant improvement (Jacobson & Truax 1991). Typically, routine outcome monitoring is applied to enable a prognostic assessment (e.g., likelihood of response to treatment) based on symptomatic change since the initial intake assessment. Thereby, reliable change criteria are used as a classification rule. Studies applying these methods show that these classification rules reliably predict treatment response in diverse treatment settings and clinical populations (e.g., Delgadillo

et al. 2014, Flood et al. 2019, Lutz et al. 2006). Another example of a rationally derived model is the routine outcome monitoring method applied in IAPT services in England. In these services, all patients complete depression and anxiety questionnaires on a session-to-session basis, and therapists apply conventional cutoff scores to assess treatment progress. Symptom reductions below the cutoff on at least one of the two measures indicate a favorable response, while a reduction below both measures' cutoffs is considered a full recovery (Clark 2018).

In contrast, empirically derived methods are based on prediction models and the above-described ETR concept and associated developments. Some newer applications also include prediction models and critical failure boundaries (beyond which the probability of negative treatment outcome is greater than positive outcome) for individual patients by dynamically recalculating failure boundaries based on current progress data—that is, assessments from previous treatment sessions (e.g., Bone et al. 2021b, Lutz et al. 2019). For example, Bone et al. (2021b) used session-by-session self-report depression and anxiety measures from 42,992 patients in IAPT services to train a dynamic prediction model (adapting failure boundaries over the course of treatment) using iterative logistic regression analysis. Subsequently, the model was evaluated on an external test sample of 30,026 patients. This dynamic prediction model improved the accuracy of empirically derived decision rules to identify patients at risk for treatment failure.

Prognostic Indices and Treatment Selection Models

Prognostic models can complement the therapist's clinical impression and thus support decision making. Data-informed prognostic indices (PIs) and prediction models based on patient information have recently been developed using machine learning approaches (see review by Chekroud et al. 2021). These PIs and prediction algorithms can be used as criteria for treatment selection to identify the optimal treatment package or treatment strategy that suits the individual best. For example, Lorenzo-Luaces et al. (2017) found that patients with a worse prognosis (i.e., an unfavorable PI) benefited more from CBT than from treatment as usual or brief therapy, while the treatment alternative made no difference if the overall prognosis was good.

Another option for treatment selection is to estimate the therapy outcome for each treatment alternative, and to recommend the treatment with the most favorable prognosis, that is, with the best outcome prediction. For this purpose, Lutz et al. (2005) used an NN model to predict patient-specific differential response to treatments. Another concept, the Personalized Advantage Index (PAI), was introduced by DeRubeis et al. (2014); it represents the difference of two predictions and quantifies the estimated superiority of one treatment model over another one (a PAI of zero means that both treatments should be equally effective for a given patient). Patients who received their recommended treatment have been shown to have better outcomes in studies evaluating a wide range of treatment packages, such as cognitive, interpersonal, person-centered, and psychodynamic therapy as well as antidepressant medication and eye movement desensitization and reprocessing (Cohen et al. 2021, Delgadillo et al. 2020, Deisenhofer et al. 2018, Huibers et al. 2015, Webb et al. 2019). Another line of research focuses not on predictions of different treatment packages but rather on expected change in clinical modules, strategies, and mechanisms (e.g., such as problem-solving and motivation-oriented interventions) to tailor clinical interventions within a treatment package (e.g., Gómez Penedo et al. 2022; Lutz et al. 2019; Ng et al. 2021; Rubel et al. 2018, 2020).

In addition, recent studies have begun to cross-validate treatment selection models using independent holdout data (e.g., Delgadillo & Gonzalez Salas Duhne 2020, Schwartz et al. 2021) and data from another study (cross-trial validation; van Bronswijk et al. 2021), or to apply these models prospectively in clinical trials (Delgadillo et al. 2021, Lutz et al. 2022).

In the future, predictions and recommendations based on pretherapy assessments collected at one point in time could be complemented by EMA data (Fisher et al. 2019, Webb et al. 2022; see mode 5 on intensive longitudinal assessments above). Furthermore, variables from different data sources can be integrated multimodally to improve predictions and recommendations (e.g., clinical and video data; Atzil-Slonim et al. 2021).

Therapist–Patient Matching

Variability in treatment outcomes between therapists (also known as therapist effects) has been extensively documented in the psychotherapy literature (Baldwin & Imel 2013). Even if they apply the same treatment protocol (e.g., CBT) for the same target problem (e.g., depression), some therapists attain impressive clinical outcomes compared to their peers, while others are less effective than average. Studies indicate that therapist effects may be partly related to patients' features, such that highly effective therapists are especially helpful for patients with severe levels of distress and risk of self-harm (Saxon & Barkham 2012). Furthermore, therapist effects also appear to be related to therapists' interpersonal skills (Heinonen & Nissen-Lie 2020) and particularly their ability to work effectively in highly challenging cases (Anderson et al. 2016). It logically follows that matching some patients to specific therapists could be a potentially effective method of treatment allocation. Recent research shows that, in fact, patients with certain combinations of demographic (e.g., employment status) and clinical features (e.g., comorbidity of depression and anxiety, symptom severity) respond better to therapy with some therapists rather than other therapists (Delgadillo et al. 2020). On this basis, using archival clinical data, it is possible to predict the likely outcome for new patients assigned to specific therapists (i.e., percent probability of recovery), potentially enabling an evidence-based approach to patient–therapist matching (Constantino et al. 2021).

Clinical Support Tools

In the context of routine outcome monitoring, CSTs have been developed. They implement several findings from psychotherapy process research to support clinical decision making. As monitoring symptom change alone does not provide any information on how to adjust the treatment strategy, the goal of such clinical problem-solving tools is to alert therapists to potential obstacles to treatment progress and to provide suggestions for interventions to improve treatment for patients at risk of treatment failure (Boswell et al. 2015). Numerous CST models have been developed and applied in the field of psychotherapy.

One such model is Lambert's assessment of signal cases (e.g., Lambert 2017), which guides therapists to routinely monitor four domains using a validated patient-reported questionnaire that covers therapeutic alliance, social support, motivation, and negative life events. If a patient is not on track according to feedback from routine outcome measures, the therapist is prompted to assess problems across one or more of these domains and to apply clinical skills that target the perspective problem (e.g., rupture repair for alliance deficits, a decisional balance exercise for motivational deficits). A meta-analysis of controlled trials that supplemented progress feedback with this clinical problem-solving approach concluded that the addition of CSTs significantly improves clinical outcomes (Shimokawa et al. 2010). Subsequent secondary analyses of data from these trials indicated that not-on-track signals were significantly associated with elevated scores on the domains of social support and adverse life events (Probst et al. 2020, White et al. 2015).

A unidimensional approach to CSTs is featured in the Partners for Change Outcome Management System (PCOMS; e.g., Miller et al. 2005), in which clinicians routinely monitor a measure of the therapeutic alliance (Session Rating Scale) in addition to treatment outcomes (Outcome

Rating Scale). This is based on the well-known association between alliance and treatment outcomes (e.g., see Flückiger et al. 2020) and follows the hypothesis that poor progress is likely to be related to alliance problems and potentially rectified by close attention to the alliance. However, a meta-analysis of clinical trials applying the PCOMS system found that subgroup analyses excluding studies that did not use the Session Rating Scale (i.e., did not systematically monitor the alliance) did not influence the overall meta-analytic finding that simple outcome monitoring (using the Outcome Rating Scale) improves treatment outcomes (Østergård et al. 2020).

Another multidomain model (Lutz et al. 2019) involves the routine monitoring of five domains: (*a*) risk/suicidality, (*b*) motivation/therapy goals, (*c*) therapeutic alliance, (*d*) social support and critical life events, and (*e*) emotion regulation/self-regulation. Each domain is assessed at every fifth session using a battery of validated patient-reported questionnaires. Like previous CST systems, this model primes the therapist to assess potential problems across these domains in cases where clinical outcomes are classed as not on track using computerized feedback. In an analysis of archival data from over 400 patients who were monitored using this battery of measures, Schilling et al. (2021) found that patients classed as not on track had significantly elevated scores on the domains of risk/suicidality and social support and critical life events.

Summary

The collection and systematic investigation of outcome measures in routine care have generated several new options to make use of continuous data assessments in clinical practice. Research findings allow the generation of clinically meaningful decision rules for the selection of treatment options as well as the monitoring of treatment and the early detection of negative developments. In particular, the development of CSTs to supplement outcome monitoring seems to constitute a step forward to improve psychological therapy for those patients who are at greatest risk of poor treatment response.

IMPLEMENTING MEASUREMENT-BASED AND DATA-INFORMED PSYCHOLOGICAL THERAPY

Continuous assessment of change during treatment enables clinicians to adapt their decisions about the best available treatment option, and it also allows the immediate application of research findings into clinical practice. A particular strength of data-rich health research and medicine is the ability to develop individualized diagnosis and treatment options. This has also been demonstrated in other public health areas; examples include the development of tailor-made immune cells and tumor therapies, the search for genes that can influence the risk of heart disease and life-threatening pulmonary hypertension, and therapy options for Parkinson's disease (e.g., Han et al. 2020, Stoker & Barker 2020).

Feedback-Informed Treatment

Important developments, such as brief psychometrically validated self-report measures, their repeated administration in routine care, and the advancement of methods such as the ETR models discussed above, have enabled the development of a new paradigm of measurement-based care. This approach uses psychometric and statistical methods to support the delivery of effective care in a way that enables therapists to make timely decisions that are supported by evidence rather than by clinical judgment alone. Furthermore, since most commonly used measurement-based care approaches use patient-reported measures (of symptoms, alliance, processes, goals, etc.), this approach places the patient's perspective at the center of the therapeutic process.

Feedback-informed treatment (FIT) is one of the most well-established examples of measurement-based care (FIT is also known as outcome feedback or patient-focused feedback research) (Howard et al. 1996, Lambert et al. 2001). FIT involves regularly monitoring a patient's response to treatment by comparing their response to predicted trajectories of improvement (clinical norms). These clinical norms are derived from ETR models (e.g., Finch et al. 2001) or other advanced statistical approaches, such as NN analysis (Lutz et al. 2006) or dynamic prediction models that are recalibrated from session to session (Bone et al. 2021b). Typically, patients are asked to complete questionnaires about their symptoms before every therapy session, and the results are entered into a computerized FIT system. The system compares each patient's symptoms to those observed in hundreds of similar patients and then provides a prognosis: Either the patient is on track (likely to recover) or not on track (unlikely to recover). This feedback is provided to therapists every week, which prompts them to quickly identify and resolve obstacles to improvement in cases classified as not on track.

There are now over 50 clinical trials and several meta-analyses indicating that FIT systems can help improve treatment outcomes, prevent dropout, and improve the efficiency of psychological treatment (for comprehensive narrative and meta-analytic reviews, see de Jong et al. 2021, Lambert et al. 2018). In particular, the evidence is stronger for studies that offered evidence-based psychological treatments for depression and/or anxiety problems, supplemented by a FIT system, and it appears to generalize to different countries and health care systems. Furthermore, supplementing FIT technology with CSTs has been shown to enhance the effects of FIT (Lambert et al. 2018).

Moderator analyses of FIT trials reveal that the effects of feedback are enhanced by using ETR-based models to track treatment response and supplementing these systems with training for therapists and with CSTs (de Jong et al. 2021). Providing feedback to therapists and patients, rather than only to therapists, also seems to enhance its effectiveness (de Jong et al. 2014). Importantly, the extent to which individual clinicians adhere to the use of feedback systems determines the extent to which measurement-based therapy will be effective (see Lutz et al. 2021). As such, there is already compelling evidence demonstrating that FIT is more effective than usual psychological care (purely guided by clinical judgment) in terms of symptomatic improvement and dropout prevention (de Jong et al. 2021), but adequate adherence by therapists varies considerably even in clinical trials. Thus, the current challenges and research questions in this field concern how to optimize the implementation of feedback systems in routine practice. Numerous obstacles to implementation have been identified in prior studies, such as organizational, technological, practical, and attitudinal barriers (for an overview, see Lutz et al. 2021). Several examples of successful implementation and strategies to optimize adoption by therapists have also been documented, and we refer interested readers to a previously published collection of implementation-focused case studies (de Jong 2016).

Precision Mental Health Care

The debate about the development and implementation of personalized medicine (also called precision medicine) has also influenced psychotherapy (research) in recent years. Several new developments have emerged within this tradition of outcome prediction and monitoring, which can be summarized as precision mental health research. This branch of research is intertwined with the traditions and improvements in continuous outcome measurement (e.g., Delgadillo & Lutz 2020, DeRubeis et al. 2014, Huibers et al. 2015). The overarching goal is to use evidence-based strategies to support decision making in clinical practice (e.g., Bickman 2020, Chekroud et al. 2021, Cohen et al. 2021, Page et al. 2019, Zilcha-Mano 2019).

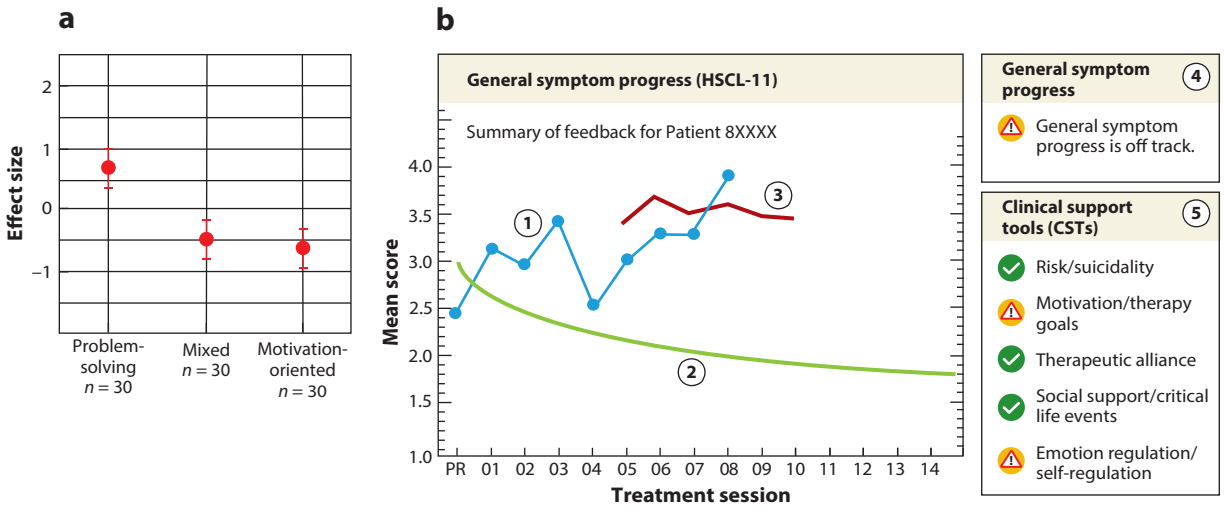


Figure 2

(*a*) Example of a patient-specific pretherapy strategy and (*b*) an adaptive personalized recommendation during treatment as displayed in the clinical decision-support system. (①) Patient scores at the beginning of each session, (②) expected recovery curve, and (③) failure boundary. (④) As soon as the patient's score exceeds the failure boundary on the HSCL-11, the therapist receives a warning signal, which is defined in more detail in the CSTs, and (⑤) CSTs are divided into five domains. The exclamation mark indicates the domains in which the patient has specific problems. The therapist is able to click on these icons to gain access to the activated tools. The check mark indicates that the patient has few or no problems in a given area. Abbreviations: CST, clinical support tool; HSCL-11, Hopkins Symptom Checklist-11; PR, preassessment.

One example of a comprehensive treatment selection and tracking system, which is also augmented by tools from e-mental health, is the Trier Treatment Navigator (TTN). This system includes a personalized treatment recommendation at the beginning as well as adjustments based on continuous outcome assessments during the course of treatment (Lutz et al. 2019, 2022). For each patient, the TTN generates individual predictions based on a large, previously treated patient sample for several important indicators—for instance, dropout risk and the optimal treatment strategy to start the treatment (see **Figure 2*a***). To predict the optimal treatment strategy, the most similar patients (based on the NN algorithm) who have already been treated are identified for an individual patient out of a large archival patient sample ($N = 1{,}234$). These NN groups of patients are generated for two treatment strategies or a combination of both (problem-solving-oriented, relationship- and motivation-oriented, or mixed approach in the first 10 sessions), and effect sizes for each group achieved in the first 10 sessions are generated. Effect sizes of the three treatment strategies and confidence intervals are then mapped, and the therapist can directly evaluate whether a particular treatment strategy is predicted to have a clear advantage for a given patient (i.e., the problem-solving approach for the patient example in **Figure 2*a***).

After the start of treatment, the navigation and monitoring system includes personalized treatment adjustment based on a dynamic risk index and CSTs. The system alerts therapists to the risk of an adverse treatment outcome. **Figure 2*b*** shows the course of treatment (session by session) and the visual feedback for a sample patient. The feedback graph is trained using data from the 30 most similar neighbors' average course of treatment and the above-described dynamic risk index. This risk index is also based on the ETRs of the most similar previously treated patients and is adaptively recalculated after each session. A patient progress score above the risk index indicates a significantly increased risk of a negative outcome. In such cases, therapists receive

a warning signal as in the example in **Figure 2*b***: "(!) General symptom progress is off track." In such cases, therapists also receive feedback on the potential problem area(s) (risk/suicidality, motivation/therapy goals, therapeutic alliance, social support/critical life events, and emotion regulation/self-regulation) in which the patient has indicated high scores. Furthermore, the system supports the implementation of helpful alternative clinical interventions by a number of support tools (e.g., videos on how to perform specific techniques, worksheets, audio files for download). In the example in **Figure 2*b***, the patient shows high scores in the areas of motivation/therapy goals and emotion regulation/self-regulation, and the feedback system advises the therapist to click on the support tools in these two areas.

Figure 2 shows the sample treatment of a 35-year-old, single male patient with major depressive disorder (single episode) and subthreshold personality disorder. At the beginning of therapy, the patient showed great ambivalence toward treatment and only started because of his friends' and family members' suggestions. He attributed all his problems to his environment, and, for example, in session 7 he was getting extremely upset about a friend who had canceled a vacation. This event led him to express strong doubts about whether therapy could help him at all. The navigation system proposed several techniques concerning doubts about treatment (CST motivation/therapy goals), and the therapist used these techniques to question the patient's perspective and attributions concerning the cause of his problems. Together, they developed alternative explanations for the origin of his problems and were able to strengthen his motivation for treatment by including additional motivational interviewing techniques.

In an RCT with 538 patients, the TTN was prospectively evaluated (Lutz et al. 2022). Patients showed an increased effect size of about 0.3 when therapists followed the recommended treatment strategy in the first 10 sessions. Moreover, the linear mixed models revealed therapist symptom awareness and therapist attitude and confidence as significant predictors of outcome as well as therapist-rated usefulness of feedback as a significant moderator of the feedback–outcome and the not on track–outcome associations. These results demonstrate the importance of prospective studies and the necessity of a high-quality implementation of digital decision-support tools in clinical practice.

Implementation

From the beginning, the integration of psychometric measures into clinical practice has been associated with barriers and implementation issues. Historically, psychotherapists have often been hesitant about or even critical of implementing measures on a routine basis (e.g., Boswell et al. 2015, de Jong 2016, Douglas et al. 2016). This phenomenon has been described as the scientist–practitioner gap in the psychological therapies (e.g., Lilienfeld et al. 2015). This contrasts with the patient perspective, from which outcome evaluation is usually well received. For example, when patients are asked whether they find it important to monitor the results of psychotherapeutic treatments, such as by using questionnaires, more than 90% seem to agree or partially agree (e.g., Lutz et al. 2021).

Furthermore, therapists' behavior and attitudes, perceived usefulness, and commitment to outcome measurement have some impact on the effectiveness of such measurement in clinical practice (e.g., de Jong et al. 2021, Lutz et al. 2022). Several factors contribute to the hesitant reception of outcome measurement in clinical practice. For example, the technical equipment, financial support, and necessary time might be not available in daily practice. Furthermore, therapists' general aversion to the use of technology in their practice has often been mentioned as a reason in the literature. In addition, clinicians might not trust the ecological validity of measures, thinking that empirical findings do not reflect their everyday practice (e.g., Boswell et al. 2015, Gilbody et al.

2002). Sometimes, measures can also be perceived as controlling, or concerns may arise with regard to data security (e.g., Mütze et al. 2021). All these factors culminate in an underuse of outcome monitoring in clinical practice, and clinicians lack the training and support necessary to make good use of the information (e.g., Boswell et al. 2015). As described throughout this review, in the future a cultural shift to measurement-based and data-informed psychological therapy including outcome measures as well as feedback and navigation systems is one of the most important steps to improve clinical services in mental health.

Summary

Measurement-based and data-informed psychological therapy evolved from efforts to extend the external validity of outcome research, which is traditionally based on efficacy research and RCTs. However, after a treatment effect has been established, observational, noninterventional designs and the monitoring of real-world applications are necessary to investigate implementation issues. The main focus of these Phase IV surveillance investigations is the transportability of a treatment into real-life conditions, as well as the identification of side effects, subgroup differences, and treatment failures (e.g., Suvarna 2010). The implementation of outcome monitoring supplemented by feedback and navigation systems takes this investigation a step further by not only monitoring real-world effects but also improving effects in real time.

CONCLUSION

Common to all endeavors described in this review is the idea of providing clinicians with personalized recommendations for their everyday clinical decision making based on continuous data collection. The generation of large practice-based data sets allows the development of new data disaggregation strategies to more precisely define relevant reference groups for each patient, thereby improving the applicability of information to the individual patient and supporting therapists in their clinical practice with an expert decision-support system.

So far, clinical judgment has been largely based on theory and intuition. However, we now have adequate assessment tools that are psychometrically reliable, brief, clinically useful, and sensitive to change (Lutz et al. 2021). Furthermore, feedback and decision-support tools are available that, when used, have been shown to lead to more evidence-based and effective treatments. Clinicians can make use of feedback to identify patients with a high likelihood of a negative treatment development and to obtain empirically supported recommendations about treatment strategies that could improve outcomes. In the future, further research efforts in this area should focus on implementation. Furthermore, the improvement of freely available and easy-to-apply measures remains a priority along with efforts to replicate results under large-scale routine care conditions. Of course, further improvements are also necessary in terms of item overlap and standardization of measures (e.g., Fried 2017). Important future research perspectives include (*a*) new statistical methods (e.g., machine learning) to analyze large cross-sectional as well as intensive longitudinal data sets, (*b*) improved research on processes and mechanisms of change, (*c*) better dissemination and cross-cultural adaptation (Kazdin 2018), and (*d*) better implementation and testing of clinical decision-support systems to identify and treat patients at risk for treatment failure.

Our vision for the future is to support clinical work not only with one algorithm but with multiple algorithms integrated into expert decision-support systems, which include algorithms and problem-solving modules to address several difficulties encountered in clinical practice, such as how to personalize the choice of treatment modality or techniques for the individual; how to predict and prevent dropout, symptomatic deterioration, and side effects; and how to predict and prevent relapse. The aim is to support human decision making, not to replace it with automated

interventions. Well-trained clinicians should receive empirically based support to solve the complex problems that are commonly encountered in the field of psychological therapy.

SUMMARY POINTS

1. Measurement-based and data-informed psychological therapy uses algorithmic decision tools to overcome some of the limitations of clinical judgment and intuition. This is enabled by the accumulation of large-scale routine clinical data sets.
2. Studies using longitudinal data show that the course of psychotherapy is influenced by patients' baseline characteristics as well as early symptomatic changes, and it often follows a nonlinear pattern. Furthermore, individual trajectories can deviate significantly from aggregated clinical population trends.
3. Intensive (i.e., daily) longitudinal assessments can improve the representation of high-frequency intrapersonal processes in patients' daily lives. However, because of their complexity and existing data analysis problems, less frequent session-by-session assessments are easier to implement and use in clinical practice.
4. Data-informed decision tools can guide treatment selection, monitoring of treatment progress, early detection of not-on-track patients, and adaptive recommendation of clinical strategies.
5. The effectiveness of measurement-based psychological therapy depends on therapists' behaviors and attitudes toward outcome measurement and feedback, as well as their use of technology. Therefore, to realize the full potential of data-informed therapy, future generations of psychotherapists need to be trained to understand psychometrics and to embrace technology as a means of making smarter and more effective decisions.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The authors are most thankful for the tremendously helpful proofreading by Kaitlyn Poster (University of Trier) and Paulina Gonzalez Salas Duhne (University of Sheffield). We also thank Björn Bennemann and Jana Wasserheß (University of Trier) for providing helpful comments, supporting figure design, and compiling the reference list.

LITERATURE CITED

Aderka IM, Nickerson A, Bøe HJ, Hofmann SG. 2012. Sudden gains during psychological treatments of anxiety and depression: a meta-analysis. *J. Consult. Clin. Psychol.* 80(1):93–101

Anderson T, McClintock AS, Himawan L, Song X, Patterson CL. 2016. A prospective study of therapist facilitative interpersonal skills as a predictor of treatment outcome. *J. Consult. Clin. Psychol.* 84(1):57–66

Atzil-Slonim D, Juravski D, Bar-Kalifa E, Gilboa-Schechtman E, Tuval-Mashiach R, et al. 2021. Using topic models to identify clients' functioning levels and alliance ruptures in psychotherapy. *Psychotherapy* 58(2):324–39

Baldwin SA, Berkeljon A, Atkins DC, Olsen JA, Nielsen SL. 2009. Rates of change in naturalistic psychotherapy: contrasting dose-effect and good-enough level models of change. *J. Consult. Clin. Psychol.* 77(2):203–11

Baldwin SA, Imel ZE. 2013. Therapist effects: findings and methods. In *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change*, ed. MJ Lambert, pp. 258–97. New York: Wiley. 6th ed.

Baldwin SA, Imel ZE. 2020. Studying specificity in psychotherapy with meta-analysis is hard. *Psychother. Res.* 30(3):294–96

Barkham M, Connell J, Stiles WB, Miles JNV, Margison F, et al. 2006. Dose-effect relations and responsive regulation of treatment duration: the good enough level. *J. Consult. Clin. Psychol.* 74(1):160–67

Barkham M, Lambert MJ. 2021. The efficacy and effectiveness of psychological therapies. See Barkham et al. 2021, pp. 129–86

Barkham M, Lutz W, Castonguay LG, eds. 2021. *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change*. New York: Wiley. 7th ed.

Barkham M, Rees A, Stiles WB, Shapiro DA, Hardy GE, Reynolds S. 1996. Dose-effect relations in time-limited psychotherapy for depression. *J. Consult. Clin. Psychol.* 64(5):927–35

Baur T, Clausen S, Heimerl A, Lingenfelser F, Lutz W, André E. 2020. NOVA: a tool for explanatory multimodal behavior analysis and its application to psychotherapy. In *Lecture Notes in Computer Science*, Vol. 11962: *MultiMedia Modeling: 26th International Conference, MMM 2020, Daejeon, South Korea, January 5–8, 2020: Proceedings*, ed. W-H Cheng, pp. 577–88. Cham, Switz.: Springer

Beard JIL, Delgadillo J. 2019. Early response to psychological therapy as a predictor of depression and anxiety treatment outcomes: a systematic review and meta-analysis. *Depress. Anxiety* 36(9):866–78

Bickman L. 2020. Improving mental health services: a 50-year journey from randomized experiments to artificial intelligence and precision mental health. *Adm. Policy Ment. Health* 47:795–843

Bone C, Delgadillo J, Barkham M. 2021a. A systematic review and meta-analysis of the good-enough level (GEL) literature. *J. Couns. Psychol.* 68(2):219–31

Bone C, Simmonds-Buckley M, Thwaites R, Sandford D, Merzhvynska M, et al. 2021b. Dynamic prediction of psychological treatment outcomes: development and validation of a prediction model using routinely collected symptom data. *Lancet Digit. Health* 3(4):e231–40

Borsboom D. 2017. A network theory of mental disorders. *World Psychiatry* 16(1):5–13

Boswell JF, Kraus DR, Miller SD, Lambert MJ. 2015. Implementing routine outcome monitoring in clinical practice: benefits, challenges, and solutions. *Psychother. Res.* 25(1):6–19

Bringmann LF. 2021. Person-specific networks in psychopathology: past, present, and future. *Curr. Opin. Psychol.* 41:59–64

Bringmann LF, van der Veen DC, Wichers M, Riese H, Stulp G. 2021. ESMvis: a tool for visualizing individual experience sampling method (ESM) data. *Qual. Life Res.* 30:3179–88

Castonguay LG, Barkham M, Lutz W, McAleavey A. 2013. Practice-oriented research: approaches and applications. In *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change*, ed. MJ Lambert, pp. 85–133. New York: Wiley. 6th ed.

Castonguay LG, Barkham M, Youn SJ, Page AC. 2021. Practice-based evidence: findings from routine clinical settings. See Barkham et al. 2021, pp. 187–220

Chang PGRY, Delgadillo J, Waller G. 2021. Early response to psychological treatment for eating disorders: a systematic review and meta-analysis. *Clin. Psychol. Rev.* 86(1):102032

Chekroud AM, Bondar J, Delgadillo J, Doherty G, Wasil A, et al. 2021. The promise of machine learning in predicting treatment outcomes in psychiatry. *World Psychiatry* 20(2):154–70

Clark DM. 2018. Realizing the mass public benefit of evidence-based psychological therapies: the IAPT program. *Annu. Rev. Clin. Psychol.* 14:159–83

Cohen ZD, Delgadillo J, DeRubeis RJ. 2021. Personalized treatment approaches. See Barkham et al. 2021, pp. 667–700

Constantino MJ, Boswell JF, Coyne AE, Swales TP, Kraus DR. 2021. Effect of matching therapists to patients vs assignment as usual on adult psychotherapy outcomes: a randomized clinical trial. *JAMA Psychiatry* 78(9):960–69

Contreras A, Nieto I, Valiente C, Espinosa R, Vazquez C. 2019. The study of psychopathology from the network analysis perspective: a systematic review. *Psychother. Psychosom.* 88(2):71–83

Cristea IA, Vecchi T, Cuijpers P. 2021. Top-down and bottom-up pathways to developing psychological interventions. *JAMA Psychiatry* 78(6):593–94
Crits-Christoph P, Connolly Gibbons MB. 2021. Psychotherapy process-outcome research: advances in understanding causal connections. See Barkham et al. 2021, pp. 259–92
Crits-Christoph P, Ring-Kurtz S, Hamilton JL, Lambert MJ, Gallop R, et al. 2012. A preliminary study of the effects of individual patient-level feedback in outpatient substance abuse treatment programs. *J. Subst. Abuse Treat.* 42(3):301–9
Cuijpers P, Reijnders M, Huibers MJH. 2019. The role of common factors in psychotherapy outcomes. *Annu. Rev. Clin. Psychol.* 15:207–31
Dakos V, Carpenter SR, Brock WA, Ellison AM, Guttal V, et al. 2012. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLOS ONE* 7(7):e41010
de Jong K. 2016. Challenges in the implementation of measurement feedback systems. *Adm. Policy Ment. Health* 43(3):467–70
de Jong K, Conijn JM, Gallagher RAV, Reshetnikova AS, Heij M, Lutz MC. 2021. Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: a multilevel meta-analysis. *Clin. Psychol. Rev.* 85:102002
de Jong K, Timman R, Hakkaart-Van Roijen L, Vermeulen P, Kooiman K, et al. 2014. The effect of outcome monitoring feedback to clinicians and patients in short and long-term psychotherapy: a randomized controlled trial. *Psychother. Res.* 24(6):629–39
Deisenhofer A-K, Delgadillo J, Rubel JA, Böhnke JR, Zimmermann D, et al. 2018. Individual treatment selection for patients with posttraumatic stress disorder. *Depress. Anxiety* 35(6):541–50
Deisenhofer A-K, Rubel JA, Bennemann B, Aderka IM, Lutz W. 2021. Are some therapists better at facilitating and consolidating sudden gains than others? *Psychother. Res.* **https://doi.org/10.1080/10503307.2021.1921302**
Delgadillo J, Ali S, Fleck K, Agnew C, Southgate A, et al. 2021. Stratified care vs stepped care for depression: a cluster randomized clinical trial. *JAMA Psychiatry* 79(2):101–8
Delgadillo J, Gonzalez Salas Duhne P. 2020. Targeted prescription of cognitive-behavioral therapy versus person-centered counseling for depression using a machine learning approach. *J. Consult. Clin. Psychol.* 88(1):14–24
Delgadillo J, Lutz W. 2020. A development pathway towards precision mental health care. *JAMA Psychiatry* 77(9):889–90
Delgadillo J, McMillan D, Lucock M, Leach C, Ali S, Gilbody S. 2014. Early changes, attrition, and dose-response in low intensity psychological interventions. *Br. J. Clin. Psychol.* 53(1):114–30
Delgadillo J, Rubel J, Barkham M. 2020. Towards personalized allocation of patients to therapists. *J. Consult. Clin. Psychol.* 88(9):799–808
DeRubeis RJ, Cohen ZD, Forand NR, Fournier JC, Gelfand LA, Lorenzo-Luaces L. 2014. The Personalized Advantage Index: translating research on prediction into individualized treatment recommendations. A demonstration. *PLOS ONE* 9(1):e83875
Douglas S, Button S, Casey SE. 2016. Implementing for sustainability: promoting use of a measurement feedback system for innovation and quality improvement. *Adm. Policy Ment. Health* 43(3):286–91
Ebner-Priemer UW, Trull TJ. 2009. Ecological momentary assessment of mood disorders and mood dysregulation. *Psychol. Assess.* 21(4):463–75
Eisele G, Vachon H, Lafit G, Kuppens P, Houben M, et al. 2022. The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. *Assessment* 29(2):136–51
Epskamp S. 2020. Psychometric network models from time-series and panel data. *Psychometrika* 85(1):206–31
Evans DL, Herbert JD, Nelson-Gray RO, Gaudiano BA. 2002. Determinants of diagnostic prototypicality judgments of the personality disorders. *J. Pers. Disord.* 16(1):95–106
Falkenström F, Ekeblad A, Holmqvist R. 2016. Improvement of the working alliance in one treatment session predicts improvement of depressive symptoms by the next session. *J. Consult. Clin. Psychol.* 84(8):738–51
Finch AE, Lambert MJ, Schaalje BG. 2001. Psychotherapy quality control: the statistical generation of expected recovery curves for integration into an early warning system. *Clin. Psychol. Psychother.* 8(4):231–42

Fisher AJ, Bosley HG, Fernandez KC, Reeves JW, Soyster PD, et al. 2019. Open trial of a personalized modular treatment for mood and anxiety. *Behav. Res. Ther.* 116:69–79
Fisher AJ, Soyster P, Ashlock L. 2021. Machine learning algorithms for generating early warning signals in real time. *Biol. Psychiatry* 89(9):S58–59
Flood N, Page A, Hooke G. 2019. A comparison between the clinical significance and growth mixture modelling early change methods at predicting negative outcomes. *Psychother. Res.* 29(7):947–58
Flückiger C, Rubel J, Del Re AC, Horvath AO, Wampold BE, et al. 2020. The reciprocal relationship between alliance and early treatment symptoms: a two-stage individual participant data meta-analysis. *J. Consult. Clin. Psychol.* 88(9):829–43
Fried EI. 2017. The 52 symptoms of major depression: lack of content overlap among seven common depression scales. *J. Affect. Disord.* 208:191–97
Frumkin MR, Piccirillo ML, Beck ED, Grossman JT, Rodebaugh TL. 2021. Feasibility and utility of idiographic models in the clinic: a pilot study. *Psychother. Res.* 31(4):520–34
Garb HN. 2005. Clinical judgment and decision making. *Annu. Rev. Clin. Psychol.* 1:67–89
Gilbody SM, House AO, Sheldon TA. 2002. Psychiatrists in the UK do not use outcomes measures: national survey. *Br. J. Psychiatry* 180(2):101–3
Goldfried MR. 2019. Obtaining consensus in psychotherapy: What holds us back? *Am. Psychol.* 74(4):484–96
Gómez Penedo JM, Schwartz B, Giesemann J, Rubel JA, Deisenhofer A-K, Lutz W. 2022. For whom should psychotherapy focus on problem coping? A machine learning algorithm for treatment personalization. *Psychother. Res.* 32(2):151–64
Grove WM, Meehl PE. 1996. Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical-statistical controversy. *Psychol. Public Policy Law* 2(2):293–323
Hamaker EL, Wichers M. 2017. No time like the present: discovering the hidden dynamics in intensive longitudinal data. *Curr. Dir. Psychol. Sci.* 26(1):10–15
Hammond KR. 1978. Toward increasing competence of thought in public policy formation. In *Judgement and Decision in Public Policy Formation*, ed. KR Hammond, pp. 11–32. New York: Routledge
Han S, Shuen WH, Wang W-W, Nazim E, Toh HC. 2020. Tailoring precision immunotherapy: coming to a clinic soon? *ESMO Open* 5(Suppl. 1):e000631
Hannan C, Lambert MJ, Harmon C, Nielsen SL, Smart DW, et al. 2005. A lab test and algorithms for identifying clients at risk for treatment failure. *J. Clin. Psychol.* 61(2):155–63
Hehlmann MI, Schwartz B, Lutz T, Gómez Penedo JM, Rubel JA, Lutz W. 2021. The use of digitally assessed stress levels to model change processes in CBT—a feasibility study on seven case examples. *Front. Psychiatry* 12:613085
Heinonen E, Nissen-Lie HA. 2020. The professional and personal characteristics of effective psychotherapists: a systematic review. *Psychother. Res.* 30(4):417–32
Heinzel S, Tominschek I, Schiepek G. 2014. Dynamic patterns in psychotherapy—discontinuous changes and critical instabilities during the treatment of obsessive compulsive disorder. *Nonlinear Dyn. Psychol. Life Sci.* 18(2):155–76
Hofmann SG, Hayes SC. 2019. The future of intervention science: process-based therapy. *Clin. Psychol. Sci.* 7(1):37–50
Howard KI, Kopta SM, Krause MS, Orlinsky DE. 1986. The dose-effect relationship in psychotherapy. *Am. Psychol.* 41(2):159–64
Howard KI, Moras K, Brill PL, Martinovich Z, Lutz W. 1996. Evaluation of psychotherapy: efficacy, effectiveness, and patient progress. *Am. Psychol.* 51(10):1059–64
Huibers MJH, Cohen ZD, Lemmens LH, Arntz A, Peeters FP, et al. 2015. Predicting optimal outcomes in cognitive therapy or interpersonal psychotherapy for depressed individuals using the Personalized Advantage Index approach. *PLOS ONE* 10(11):e0140771
Husen K, Rafaeli E, Rubel JA, Bar-Kalifa E, Lutz W. 2016. Daily affect dynamics predict early response in CBT: feasibility and predictive validity of EMA for outpatient psychotherapy. *J. Affect. Disord.* 206:305–14
Imel ZE, Barco JS, Brown HJ, Baucom BR, Baer JS, et al. 2014. The association of therapist empathy and synchrony in vocally encoded arousal. *J. Couns. Psychol.* 61(1):146–53

Jacobson NC, Weingarden H, Wilhelm S. 2019. Digital biomarkers of mood disorders and symptom change. *NPJ Digit. Med.* 2:3
Jacobson NS, Truax P. 1991. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J. Consult. Clin. Psychol.* 59(1):12–19
Kahneman D, Klein G. 2009. Conditions for intuitive expertise: a failure to disagree. *Am. Psychol.* 64(6):515–26
Kaiser T, Laireiter A-R. 2018. Process-symptom-bridges in psychotherapy: an idiographic network approach. *J. Pers.-Oriented Res.* 4(2):49–62
Kazdin AE. 2018. *Innovations in Psychosocial Interventions and Their Delivery: Leveraging Cutting-Edge Science to Improve the World's Mental Health*. New York: Oxford Univ. Press
Kraemer HC, Stice E, Kazdin A, Offord D, Kupfer D. 2001. How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. *Am. J. Psychiatry* 158(6):848–56
Krause MS. 2018. Associational versus correlational research study design and data analysis. *Qual. Quant.* 52(6):2691–707
Krüger A, Ehring T, Priebe K, Dyer AS, Steil R, Bohus M. 2014. Sudden losses and sudden gains during a DBT-PTSD treatment for posttraumatic stress disorder following childhood sexual abuse. *Eur. J. Psychotraumatol.* 5(1):24470
Kyron MJ, Hooke GR, Page AC. 2019. Assessing interpersonal and mood factors to predict trajectories of suicidal ideation within an inpatient setting. *J. Affect. Disord.* 252:315–24
Lambert MJ. 2017. Maximizing psychotherapy outcome beyond evidence-based medicine. *Psychother. Psychosom.* 86(2):80–89
Lambert MJ, Hansen NB, Finch AE. 2001. Patient-focused research: using patient outcome data to enhance treatment effects. *J. Consult. Clin. Psychol.* 69(2):159–72
Lambert MJ, Whipple JL, Bishop MJ, Vermeersch DA, Gray GV, Finch AE. 2002. Comparison of empirically-derived and rationally-derived methods for identifying patients at risk for treatment failure. *Clin. Psychol. Psychother.* 9(3):149–64
Lambert MJ, Whipple JL, Kleinstäuber M. 2018. Collecting and delivering progress feedback: a meta-analysis of routine outcome monitoring. *Psychotherapy* 55(4):520–37
Lenton TM, Held H, Kriegler E, Hall JW, Lucht W, et al. 2008. Tipping elements in the Earth's climate system. *PNAS* 105(6):1786–93
Lilienfeld SO, Ritschel LA, Lynn SJ, Cautin RL, Latzman RD. 2015. Science-practice gap. In *The Encyclopedia of Clinical Psychology*, ed. RL Cautin, SO Lilienfeld, pp. 2548–55. Hoboken, NJ: Wiley
Lorenzo-Luaces L, DeRubeis RJ. 2018. Miles to go before we sleep: advancing the understanding of psychotherapy by modeling complex processes. *Cogn. Ther. Res.* 42(2):212–17
Lorenzo-Luaces L, DeRubeis RJ, van Straten A, Tiemens B. 2017. A prognostic index (PI) as a moderator of outcomes in the treatment of depression: a proof of concept combining multiple variables to inform risk-stratified stepped care models. *J. Affect. Disord.* 213:78–85
Lutz W, de Jong K, Rubel JA, Delgadillo J. 2021. Measuring, predicting and tracking change in psychotherapy. See Barkham et al. 2021, pp. 89–133
Lutz W, Deisenhofer A-K, Rubel J, Bennemann B, Giesemann J, et al. 2022. Prospective evaluation of a clinical decision support system in psychological therapy. *J. Consult. Clin. Psychol.* 90(1):90–106
Lutz W, Ehrlich T, Rubel J, Hallwachs N, Röttger M-A, et al. 2013. The ups and downs of psychotherapy: sudden gains and sudden losses identified with session reports. *Psychother. Res.* 23(1):14–24
Lutz W, Hofmann SG, Rubel J, Boswell JF, Shear MK, et al. 2014. Patterns of early change and their relationship to outcome and early treatment termination in patients with panic disorder. *J. Consult. Clin. Psychol.* 82(2):287–97
Lutz W, Leach C, Barkham M, Lucock M, Stiles WB, et al. 2005. Predicting change for individual psychotherapy clients on the basis of their nearest neighbors. *J. Consult. Clin. Psychol.* 73(5):904–13
Lutz W, Martinovich Z, Howard KI. 1999. Patient profiling: an application of random coefficient regression models to depicting the response of a patient to outpatient psychotherapy. *J. Consult. Clin. Psychol.* 67(4):571–77
Lutz W, Rubel JA, Schwartz B, Schilling V, Deisenhofer A-K. 2019. Towards integrating personalized feedback research into clinical practice: development of the Trier Treatment Navigator (TTN). *Behav. Res. Ther.* 120:103438

Lutz W, Saunders SM, Leon SC, Martinovich Z, Kosfelder J, et al. 2006. Empirically and clinically useful decision making in psychotherapy: differential predictions with treatment response models. *Psychol. Assess.* 18(2):133–41

Lutz W, Schwartz B. 2021. Trans-theoretical clinical models and the implementation of precision mental health care. *World Psychiatry* 20(3):380–81

Lutz W, Schwartz B, Hofmann SG, Fisher AJ, Husen K, Rubel JA. 2018. Using network analysis for the prediction of treatment dropout in patients with mood and anxiety disorders: a methodological proof-of-concept study. *Sci. Rep.* 8(1):7819

May RM, Levin SA, Sugihara G. 2008. Complex systems: ecology for bankers. *Nature* 451(7181):893–95

McAleavey AA, Nordberg SS, Moltu C. 2021. Initial quantitative development of the Norse Feedback system: a novel clinical feedback system for routine mental healthcare. *Qual. Life Res.* 30:3097–115

McClintock AS, Perlman MR, McCarrick SM, Anderson T, Himawan L. 2017. Enhancing psychotherapy process with common factors feedback: a randomized, clinical trial. *J. Couns. Psychol.* 64(3):247–60

McKay D, Jensen-Doss A. 2021. Harmful treatments in psychotherapy. *Clin. Psychol.: Sci. Pract.* 28(1):2–4

Meehl PE. 1973. Why I do not attend case conferences. In *Psychodiagnosis: Selected Papers*, ed. PE Meehl, pp. 225–302. Minneapolis: Univ. Minn. Press

Meisel C, El Atrache R, Jackson M, Schubach S, Ufongene C, Loddenkemper T. 2020. Machine learning from wristband sensor data for wearable, noninvasive seizure forecasting. *Epilepsia* 61(12):2653–66

Miller G. 2012. The smartphone psychology manifesto. *Perspect. Psychol. Sci.* 7(3):221–37

Miller PR, Dasher R, Collins R, Griffiths P, Brown F. 2001. Inpatient diagnostic assessments: 1. Accuracy of structured versus unstructured interviews. *Psychiatry Res.* 105(3):255–64

Miller SD, Duncan BL, Sorrell R, Brown GS. 2005. The partners for change outcome management system. *J. Clin. Psychol.* 61(2):199–208

Moggia D, Lutz W, Arndt A, Feixas G. 2020. Patterns of change and their relationship to outcome and follow-up in group and individual psychotherapy for depression. *J. Consult. Clin. Psychol.* 88(8):757–73

Mütze K, Witthöft M, Lutz W, Bräscher A-K. 2021. Matching research and practice: prediction of individual patient progress and dropout risk for basic routine outcome monitoring. *Psychother. Res.* **https://doi.org/10.1080/10503307.2021.1930244**

Ng MY, Schleider JL, Horn RL, Weisz JR. 2021. Psychotherapy for children and adolescents: from efficacy to effectiveness, scaling, and personalizing. See Barkham et al. 2021, pp. 621–66

Nordmo M, Monsen JT, Høglend PA, Solbakken OA. 2020. Investigating the dose–response effect in open-ended psychotherapy. *Psychother. Res.* 31(7):859–69

Olthof M, Hasselman F, Strunk G, van Rooij M, Aas B, et al. 2020. Critical fluctuations as an early-warning signal for sudden gains and losses in patients receiving psychotherapy for mood disorders. *Clin. Psychol. Sci.* 8(1):25–35

Østergård OK, Randa H, Hougaard E. 2020. The effect of using the Partners for Change Outcome Management System as feedback tool in psychotherapy—a systematic review and meta-analysis. *Psychother. Res.* 30(2):195–212

Owen J, Adelson J, Budge S, Wampold B, Kopta M, et al. 2015. Trajectories of change in psychotherapy. *J. Clin. Psychol.* 71(9):817–27

Page AC, Camacho KS, Page JT. 2019. Delivering cognitive behaviour therapy informed by a contemporary framework of psychotherapy treatment selection and adaptation. *Psychother. Res.* 29(8):971–73

Paz A, Rafaeli E, Bar-Kalifa E, Gilboa-Schectman E, Gannot S, et al. 2021. Intrapersonal and interpersonal vocal affect dynamics during psychotherapy. *J. Consult. Clin. Psychol.* 89(3):227–39

Probst T, Kleinstäuber M, Lambert MJ, Tritt K, Pieh C, et al. 2020. Why are some cases not on track? An item analysis of the assessment for signal cases during inpatient psychotherapy. *Clin. Psychol. Psychother.* 27(4):559–66

Ramseyer F, Tschacher W. 2011. Nonverbal synchrony in psychotherapy: coordinated body movement reflects relationship quality and outcome. *J. Consult. Clin. Psychol.* 79(3):284–95

Raudenbush SW, Bryk AS. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, CA: Sage

Robinson L, Delgadillo J, Kellett S. 2020a. The dose-response effect in routinely delivered psychological therapies: a systematic review. *Psychother. Res.* 30(1):79–96

Robinson L, Kellett S, Delgadillo J. 2020b. Dose-response patterns in low and high intensity cognitive behavioral therapy for common mental health problems. *Depress. Anxiety* 37(3):285–94
Rubel JA, Fisher AJ, Husen K, Lutz W. 2018. Translating person-specific network models into personalized treatments: development and demonstration of the Dynamic Assessment Treatment Algorithm for Individual Networks (DATA-IN). *Psychother. Psychosom.* 87(4):249–51
Rubel JA, Rosenbaum D, Lutz W. 2017. Patients' in-session experiences and symptom change: session-to-session effects on a within- and between-patient level. *Behav. Res. Ther.* 90:58–66
Rubel JA, Zilcha-Mano S, Giesemann J, Prinz J, Lutz W. 2020. Predicting personalized process-outcome associations in psychotherapy using machine learning approaches—a demonstration. *Psychother. Res.* 30(3):300–9
Saunders R, Cape J, Fearon P, Pilling S. 2016. Predicting treatment outcome in psychological treatment services by identifying latent profiles of patients. *J. Affect. Disord.* 197:107–15
Saxon D, Barkham M. 2012. Patterns of therapist variability: therapist effects and the contribution of patient severity and risk. *J. Consult. Clin. Psychol.* 80(4):535–46
Scheffer M, Bascompte J, Brock WA, Brovkin V, Carpenter SR, et al. 2009. Early-warning signals for critical transitions. *Nature* 461(7260):53–59
Schilling VNLS, Zimmermann D, Rubel JA, Boyle KS, Lutz W. 2021. Why do patients go off track? Examining potential influencing factors for being at risk of psychotherapy treatment failure. *Qual. Life Res.* 30:3287–98
Schwartz B, Cohen ZD, Rubel JA, Zimmermann D, Wittmann WW, Lutz W. 2021. Personalized treatment selection in routine care: integrating machine learning and statistical algorithms to recommend cognitive behavioral or psychodynamic therapy. *Psychother. Res.* 31(1):33–51
Shimokawa K, Lambert MJ, Smart DW. 2010. Enhancing treatment outcome of patients at risk of treatment failure: meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *J. Consult. Clin. Psychol.* 78(3):298–311
Stiles WB, Honos-Webb L, Surko M. 1998. Responsiveness in psychotherapy. *Clin. Psychol.: Sci. Pract.* 5(4):439–58
Stoker TB, Barker RA. 2020. Recent developments in the treatment of Parkinson's Disease. *F1000Research* 9(Faculty Rev.):862
Strunk DR, Cooper AA, Ryan ET, DeRubeis RJ, Hollon SD. 2012. The process of change in cognitive therapy for depression when combined with antidepressant medication: predictors of early intersession symptom gains. *J. Consult. Clin. Psychol.* 80(5):730–38
Stulz N, Lutz W, Kopta SM, Minami T, Saunders SM. 2013. Dose-effect relationship in routine outpatient psychotherapy: Does treatment duration matter? *J. Couns. Psychol.* 60(4):593–600
Suvarna V. 2010. Phase IV of drug development. *Perspect. Clin. Res.* 1(2):57–60
Tang TZ, DeRubeis RJ. 1999. Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *J. Consult. Clin. Psychol.* 67(6):894–904
Tversky A, Kahneman D. 1974. Judgment under uncertainty: heuristics and biases. *Science* 185(4157):1124–31
van Borkulo C, Boschloo L, Borsboom D, Penninx BWJH, Waldorp LJ, Schoevers RA. 2015. Association of symptom network structure with the course of corrected depression. *JAMA Psychiatry* 72(12):1219–26
van Bronswijk SC, Bruijniks SJE, Lorenzo-Luaces L, DeRubeis RJ, Lemmens LHJM, et al. 2021. Cross-trial prediction in psychotherapy: external validation of the Personalized Advantage Index using machine learning in two Dutch randomized trials comparing CBT versus IPT for depression. *Psychother. Res.* 31(1):78–91
Walfish S, McAlister B, O'Donnell P, Lambert MJ. 2012. An investigation of self-assessment bias in mental health providers. *Psychol. Rep.* 110(2):639–44
Webb CA, Forgeard M, Israel ES, Lovell-Smith N, Beard C, Björgvinsson T. 2022. Personalized prescriptions of therapeutic skills from patient characteristics: an ecological momentary assessment approach. *J. Consult. Clin. Psychol.* 90(1):51–60
Webb CA, Trivedi MH, Cohen ZD, Dillon DG, Fournier JC, et al. 2019. Personalized prediction of antidepressant v. placebo response: evidence from the EMBARC study. *Psychol. Med.* 49(7):1118–27
White MM, Lambert MJ, Ogles BM, Mclaughlin SB, Bailey RJ, Tingey KM. 2015. Using the Assessment for Signal Clients as a feedback tool for reducing treatment failure. *Psychother. Res.* 25(6):724–34

Wichers M, Groot PC, Psychosystems, ESM Group, EWS Group. 2016. Critical slowing down as a personalized early warning signal for depression. *Psychother. Psychosom.* 85(2):114–16

Wichers M, Peeters F, Rutten BPF, Jacobs N, Derom C, et al. 2012. A time-lagged momentary assessment study on daily life physical activity and affect. *Health Psychol.* 31(2):135–44

Wright AGC, Woods WC. 2020. Personalized models of psychopathology. *Annu. Rev. Clin. Psychol.* 16:49–74

Zell E, Strickhouser JE, Sedikides C, Alicke MD. 2020. The better-than-average effect in comparative self-evaluation: a comprehensive review and meta-analysis. *Psychol. Bull.* 146(2):118–49

Zilcha-Mano S. 2017. Is the alliance really therapeutic? Revisiting this question in light of recent methodological advances. *Am. Psychol.* 72(4):311–25

Zilcha-Mano S. 2019. Major developments in methods addressing for whom psychotherapy may work and why. *Psychother. Res.* 29(6):693–708

Psychotherapy and Psychosomatics

Innovations

Psychother Psychosom 2020;89:345–356
DOI: 10.1159/000510006

Received: June 15, 2020
Accepted: July 6, 2020
Published online: August 13, 2020

Sequential Combination of Cognitive-Behavioral Treatment and Well-Being Therapy in Depressed Patients with Acute Coronary Syndromes: A Randomized Controlled Trial (TREATED-ACS Study)

Chiara Rafanelli[a] Sara Gostoli[a] Sara Buzzichelli[b] Jenny Guidi[a] Laura Sirri[a] Pamela Gallo[c] Enrica Marzola[b] Serena Bergerone[d] Gaetano Maria De Ferrari[d] Renzo Roncuzzi[e] Giuseppe Di Pasquale[c] Giovanni Abbate-Daga[b] Giovanni A. Fava[f]

[a]Department of Psychology, University of Bologna, Bologna, Italy; [b]Eating Disorders Center for Treatment and Research, Department of Neuroscience, University of Turin, Turin, Italy; [c]Division of Cardiology, Maggiore Hospital, Bologna, Italy; [d]Division of Cardiology, Internal Medicine Department, Città della Salute e della Scienza, University of Turin, Turin, Italy; [e]Division of Cardiology, Bellaria Hospital, Bologna, Italy; [f]Department of Psychiatry, University at Buffalo, Buffalo, NY, USA

Keywords
Acute coronary syndrome · Cognitive-behavioral therapy · Depression · Sequential treatment · Well-being therapy

Abstract
Introduction: Randomized controlled trials (RCT) of psychotherapeutic interventions have addressed depression and demoralization associated with acute coronary syndromes (ACS). The present trial introduces psychological well-being, an increasingly recognized factor in cardiovascular health, as a therapeutic target. ***Objective:*** This study was designed to determine whether the sequential combination of cognitive-behavioral therapy (CBT) and well-being therapy (WBT) may yield more favorable outcomes than an active control group (clinical management; CM) and to identify subgroups of patients at greater risk for cardiac negative outcomes. ***Methods:*** This multicenter RCT compared CBT/WBT sequential combination versus CM, with up to 30 months of follow-up. One hundred consecutive depressed and/or demoralized patients (out of 740 initially screened by cardiologists after a first episode of ACS) were randomized to CBT/WBT associated with lifestyle suggestions ($n = 50$) and CM ($n = 50$). The main outcome measures included: severity of depressive symptoms according to the Clinical Interview for Depression, changes in subclinical psychological distress, well-being, and biomarkers, and medical complications and events. ***Results:*** CBT/WBT sequential combination was associated with a significant improvement in depressive symptoms compared to CM. In both groups, the benefits persisted at follow-up, even though the differences faded. Treatment was also related to a significant amelioration of biomarkers (platelet count, HDL, and D-dimer), whereas the 2 groups showed similar frequencies of adverse cardiac events. ***Con-***

karger@karger.com
www.karger.com/pps

Chiara Rafanelli
Department of Psychology, University of Bologna
Viale Berti Pichat 5
IT–40127 Bologna (Italy)
chiara.rafanelli@unibo.it

clusions: Addressing psychological well-being in the psychotherapeutic approach to ACS patients with depressive symptoms was found to entail important clinical benefits. It is argued that lifestyle changes geared toward cardiovascular health may be facilitated by a personalized approach that targets well-being.

Introduction

There is extensive evidence that the presence of depressive symptoms in acute coronary syndromes (ACS) is associated with poor therapeutic adherence, a higher frequency of relapses, and increased mortality [1]. Mood disturbances may consist of major or minor depressive episodes, chronic depression, and demoralization [1–3], which is characterized by a sense of subjective incompetence [4].

The relationship of depression to ACS has generated the hypothesis that treatment of mood disturbances may yield improved medical and psychological outcomes. A number of randomized controlled trials (RCT) have indicated the effectiveness of antidepressant drugs compared to placebo in relieving depression, yet a favorable effect on cardiovascular events was not detected [1] or could not be generalized [5]. Similar findings have been reported for the application of cognitive-behavioral therapy (CBT) to ACS [6], pioneered by the ENRICHD trial [7].

Psychotherapeutic approaches, however, have been mainly shifted to the side of psychological dysfunction and have neglected psychological well-being. There is increasing evidence of the role of positive psychological assets on lifestyle and cardiovascular health [8].

In this trial, the sequential use of distress and well-being psychotherapeutic strategies was selected. The first phase of treatment (CBT) was concerned with distress associated with hospitalization and medical events. In the second phase, well-being therapy (WBT), a specific psychotherapeutic approach for modulating psychological well-being [9], was introduced and suggestions for lifestyle modifications geared to cardiovascular health were provided [10]. The sequential combination of CBT and WBT has been found to yield enduring clinical benefits in the setting of psychiatric disorders [9, 10], with particular reference to recurrent depression [11].

The aim of the trial was to evaluate the efficacy of the sequential combination of CBT and WBT, compared to clinical management (CM), in terms of depressive symptoms (primary outcome), psychological distress, and well-being, as well as cardiovascular events, biomarkers, and mortality (secondary outcomes), both after treatment and up to a 30-month follow-up. The identification of subgroups of patients at greater risk for cardiac negative outcomes was included.

Materials and Methods

Sample

Participants were patients hospitalized for a first episode of acute myocardial infarction or unstable angina at the Cardiology Divisions of Maggiore Hospital (Bologna, Italy) and Molinette Hospital (Torino, Italy). Myocardial infarction was documented based on cardiac symptoms (presence of acute chest, epigastric, neck, jaw, or arm pain or discomfort or pressure without an apparent noncardiac source) and signs (acute congestive heart failure or cardiogenic shock in the absence of non-CHD causes) associated with ECG findings (characteristic evolutionary ST-T changes or new Q waves) and/or cardiac biomarkers (blood measures of myocardial necrosis, specifically CK, CK-MB, CK-MBm, or troponin, and cTn). Instable angina was documented based on cardiac symptoms (chest pain lasting less than 20 min) with likely ECG findings (ST-segment depression and an abnormal T-wave) in absence of myocardial necrosis biomarkers.

Medically eligible patients underwent a psychological evaluation by 2 clinical psychologists with expertise in the field of psychosomatic aspects of cardiovascular diseases about 30 days after ACS. The inclusion criteria were: a current diagnosis of major/minor depression or dysthymia according to DSM-IV-TR [12] and/or demoralization according to Diagnostic Criteria for Psychosomatic Research (DCPR) criteria [13]. The exclusion criteria included a positive history of bipolar disorder (DSM-IV-TR), major depression with psychotic features, a positive history of substance abuse/dependence during the previous 12 months, suicide risk, and current use of antidepressants and/or psychotherapy.

A psychological evaluation was performed in 288 patients with a first episode of ACS, and the first 100 depressed and/or demoralized consecutive patients were enrolled (Fig. 1).

Assessment

Medical Variables

Data on ACS, traditional cardiac risk factors (smoking habit, hypertension, dyslipidemia, a family history of cardiovascular disease, diabetes mellitus, and left ventricular ejection fraction <40), medications, and comorbidities were collected from medical records. The cardiologists involved in this study evaluated the patients at intake and once every 6 months to monitor changes in the clinical course of cardiac disease. Data from electrocardiograms, echocardiograms, X-rays, blood pressure and blood samples (cholesterol levels, creatinine, glycosylated hemoglobin, C-reactive protein, and coagulation/fibrinolysis biomarkers) were provided at intake. The Global Registry of Acute Coronary Events (GRACE) risk index [14] was calculated during hospital admission for ACS to determine the risk of morbidity and mortality both in hospital and 6 months after discharge. From the beginning of the psychological treatment and up to a 30-month follow-up after the end of the intervention, information about cardiac negative outcomes,

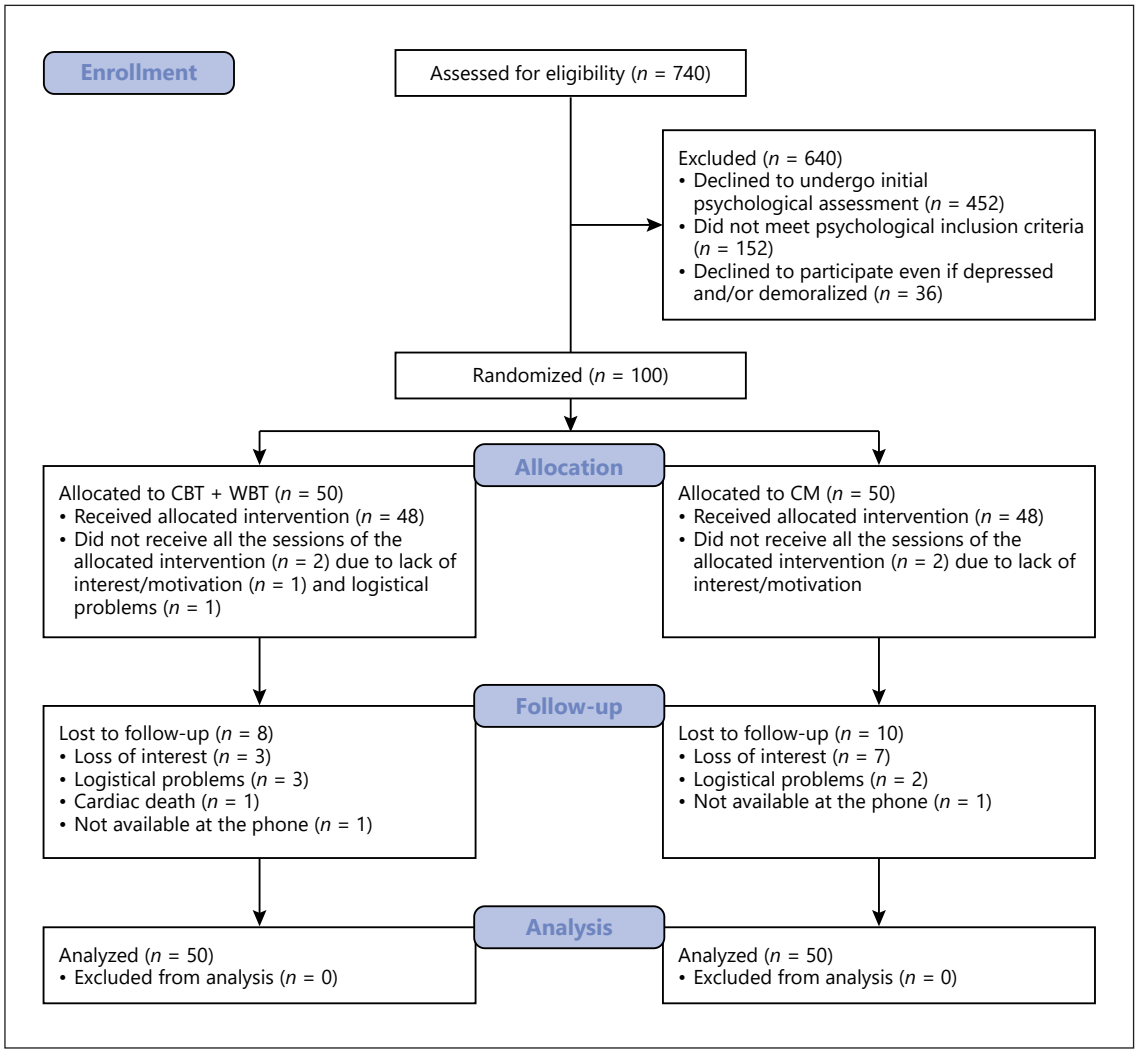


Fig. 1. CONSORT flow diagram of this study.

such as rehospitalizations due to cardiac complications, acute myocardial infarction, unstable angina, angioplasty, cardiac surgery, and cardiac mortality after the first ACS, was collected.

Psychological Variables

Psychological assessment included both observer-rated and self-reported measures before the beginning of the interventions (baseline, pretreatment), at the end (posttreatment), and 3, 6, 12, and 30 months after the end of treatment. The Structured Clinical Interview for DSM-IV-TR, Axis I Disorders [15], was used to investigate the presence of major/minor depression and dysthymia. The Semi-Structured Interview based on the DCPR (SSI-DCPR) [16] was administered to assess the presence of demoralization [17]. This interview has shown excellent interrater reliability, with κ values ranging from 0.69 to 0.97 [18]. The 20-item change version of the Clinical Interview for Depression (CID) [19, 20], a modified version of the Hamilton Rating Scale for Depression [21, 22], was used to perform a comprehensive assessment of affective symptoms. It contains 20 items rated on a 7-point Likert scale, with specification of each anchor point based on the severity, frequency, and/or quality of the symptoms. The higher the score, the worse the psychological condition. The CID has been shown to be a sensitive assessment tool in clinical trials [20]. The Symptom Questionnaire (SQ) [23, 24] is a 92-item self-report questionnaire that yields 4 main scales, i.e., depression, anxiety, hostility-irritability, and somatization. The higher the score, the higher the psychological distress. The Psychological Well-Being scales (PWB) [25–26], an 84-item questionnaire, was used to evaluate 6 psychological well-being dimensions (autonomy, environmental mastery, personal growth, positive relations, purpose in life, and self-acceptance). Higher scores correspond to greater psychological well-being.

Study Design

This study is a 2-center RCT with a longitudinal and prospective design. The enrolled patients were randomly assigned to either CBT/WBT or CM and assessed at the beginning and the end of the CBT/WBT or CM sessions, and at subsequent follow-ups up to 30 months after the conclusion of the interventions. Treatment allo-

cation was accomplished through random computerized assignment that allocated 50% of the patients to each treatment group, with assignments concealed until the time of group assignment. Patients were assessed by 2 clinical psychologists, who were blind to treatment assignment, at pretreatment and posttreatment, and 3, 6, 12, and 30 months after the end of treatment. Both the sequential combination of CBT/WBT and the CM were performed by psychotherapists who had received specific training. Both interventions consisted of 12 weekly, 45-min sessions. The sequential administration of CBT (8 sessions) and WBT (4 sessions) was based on a written protocol [9–10]. The WBT techniques were used to improve or balance one or more of the 6 dimensions of psychological well-being (environmental mastery, purpose in life, personal growth, autonomy, self-acceptance, and positive relations with others), and they were supplemented with suggestions for lifestyle modifications geared toward cardiovascular health, including treatment adherence.

CM entails the same amount of time and attention from a professional figure than the experimental group, but specific interventions (such as exposure strategies, diary work, and cognitive restructuring) were proscribed [27]. Such a form of active control – unlike in previous trials that have used treatment as usual [6] – allows discrimination of specific and nonspecific ingredients of the psychotherapeutic approach. It consists of empathic listening, review of the patient's clinical status and providing opportunities for disclosure of distress and worries, and encouragement of treatment adherence.

Statistical Analyses

Data were analyzed using SPSS 20.0 (SPSS Inc., Chicago, IL, USA). The quality of data collection was monitored regularly to assure accuracy and completeness. For all tests performed, significance level was set at 0.05 (two-tailed). The sample size was estimated using Piface software, which identified a minimum of 16 participants per arm to detect the expected superiority of CBT/WBT on CM [11], with a power of 80% and a significance level of 5%. Thus, with 50 patients per group we expected a "large" effect size (Cohen's $d = 0.8$) [28].

A multivariate ANOVA was used to examine differences in dimensional psychological variables (i.e., CID-20 total score and PWB and SQ scale scores) between patients assigned to CBT/WBT and CM at preintervention.

A mixed-model ANOVA (repeated measures) was performed to test differences between groups (CBT/WBT or CM) on the CID-20 total score, PWB scales, and SQ scales scores at different follow-up evaluations. All analyses were performed by using intention-to-treat analysis, where missing values were managed by means of a multiple-imputations procedure. Greenhouse-Geisser correction was applied when appropriate. All analyses were adjusted for cardiac illness severity (i.e., GRACE index for the 6-month probability of cardiac mortality) [14].

Each biomarker was dichotomized around the baseline median of the sample in order to identify subgroups of patients at a higher cardiovascular risk. The McNemar test (applied to contingency tables) was used to identify significant changes over time in the frequencies of DSM, DCPR diagnoses, and subgroups of patients at a higher cardiovascular risk.

Survival analyses (Cox Regression and Kaplan-Meier) to identify cardiac events and mortality that occurred between pretreatment and the 30-month follow-up were performed.

Results

Baseline Profile of the Sample

The first 100 consecutive depressed and/or demoralized patients 1 month after ACS were enrolled, yielding 50 patients in each treatment group. The mean age of the sample was 58.8 years (SD = 10.5, range 40–84). The participants were mainly men (69%), married (69%), employed (58%), and graduated from high school (44%). No significant differences based on group allocation were found (Table 1).

As for the cardiac profile of the sample, ST-elevation myocardial infarction (STEMI) was the most frequent form of ACS (66%) and almost all of the patients (94%) underwent percutaneous transluminal coronary angioplasty – 77% with the application of a single stent and 17% with 2 or more stents. The most frequent cardiovascular risk factors registered at hospital admission were dyslipidemia (58%) and hypertension (52%). No differences concerning ACS-related aspects or GRACE risk scores were found when comparing CBT/WBT versus CM (Table 1).

Among the medications prescribed at discharge, the most frequent were statins (96%), β-blockers (96%), and platelet aggregation inhibitors (96%). Patients allocated to CM were prescribed significantly more frequently β-blockers, calcium antagonists, and α-adrenergic receptor inhibitors compared to the CBT/WBT group (Table 1). The sample presented with a number of medical comorbidities; the most frequent were gastrointestinal (43%) and endocrine diseases (14%). As for comorbid medical diagnoses and levels of biomarkers assessed at baseline, the 2 groups did not show any significant difference (Table 1). From the psychological point of view, the most frequent diagnosis was demoralization (91%), followed by minor depression (56%). The 2 groups did not show any statistical difference, except for PWB "personal growth" scores ($F = 4.45$; df = 1, 98; $p = 0.038$) and frequency of depression/demoralization comorbidity ($\chi^2 = 4.86$; df = 1; $p = 0.028$), which were significantly higher among the CBT/WBT patients (Table 1).

Pre-/Postintervention Modifications

Psychological Variables

Forty-eight patients completed the CBT/WBT treatment, and 48 patients attended CM sessions. Two patients in each group dropped out early, mainly due to a lack of interest or motivation. Forty and 38 patients, respectively, completed follow-up evaluations (Fig. 1).

Table 1. Baseline sociodemographic, medical, and psychological profile of the sample

Variable	CBT/WBT group (n = 50)	CM group (n = 50)
Mean age (SD), years	57.64 (9.99)	60.02 (10.94)
Sex, *n* (%)		
Males	31 (62)	38 (76)
Females	19 (38)	12 (24)
Marital status, *n* (%)		
Single	4 (8)	7 (14)
Married	33 (66)	36 (72)
Separated	5 (10)	4 (8)
Divorced	2 (4)	1 (2)
Widow/widower	6 (12)	2 (4)
Occupation, *n* (%)		
Employed	34 (68)	24 (48)
Unemployed	1 (2)	4 (8)
Retired	13 (26)	19 (38)
Homemaker	2 (4)	3 (6)
Education, *n* (%)		
Primary school	5 (10)	5 (10)
Middle school	16 (32)	18 (36)
High school	19 (38)	25 (50)
University	8 (16)	1 (2)
Postgraduate education	2 (4)	1 (2)
Type of ACS, *n* (%)		
STEMI acute myocardial infarction	33 (66)	33 (66)
NSTEMI acute myocardial infarction	14 (28)	13 (26)
Unstable angina	3 (6)	4 (8)
Medical procedure for ACS, *n* (%)		
Single PTCA	38 (76)	39 (78)
PTCA with 2 or more stents	9 (18)	8 (16)
None	3 (6)	3 (6)
Drug-eluting stent	24 (51.1)	18 (38.3)
Cardiovascular risk factors, *n* (%)		
Dyslipidemia	31 (62)	27 (54)
Hypertension	27 (54)	25 (50)
Smoker (current)	22 (44)	20 (40)
Familiarity	17 (34)	11 (22)
Diabetes	10 (20)	9 (18)
LVEF <40	4 (8)	3 (6)
Mean GRACE risk index at admission (mortality) (SD)		
In-hospital risk, %	3.51 (8.58)	4.56 (7.90)
6-month risk, %	6.60 (11.60)	8.69 (10.57)
Mean GRACE risk index at admission (mortality + AMI) (SD)		
In-hospital risk, %	15.50 (9.85)	16.56 (10.49)
6-month risk, %	25.30 (12.73)	27.50 (15.00)
Medications, *n* (%)		
Cholesterol reducers	49 (98)	47 (94)
β-blockers*	46 (92)	50 (100)
Platelet aggregation inhibitors	48 (96)	48 (96)
Cardioaspirin	47 (94)	48 (96)
Vasodilators	36 (72)	35 (70)
Angiotensin-converting enzyme inhibitors	31 (62)	35 (70)
Polyunsaturated fatty acids – omega-3	11 (22)	10 (20)
Antihyperglycemics	6 (12)	8 (16)
Diuretics	6 (12)	5 (10)
Angiotensin receptor blockers	5 (10)	4 (8)
Calcium antagonists*	1 (2)	6 (12)

Table 1 (continued)

Variable	CBT/WBT group (n = 50)	CM group (n = 50)
α-adrenergic receptor inhibitors*	0 (0)	4 (8)
Antihyperuricemics	0 (0)	2 (4)
Antiarrhythmic	1 (2)	0 (0)
Heart rate reducers	0 (0)	1 (2)
7 or more medications*	11 (22)	23 (46)
Medical comorbidities, n (%)		
Digestive system diseases	18 (36)	25 (50)
Endocrine diseases	9 (18)	5 (10)
Circulatory/cardiac comorbidities	2 (4)	4 (8)
Prostatic and male reproductive system diseases	3 (6)	2 (4)
Urinary system diseases	2 (4)	2 (4)
Orthopedic diseases	1 (2)	3 (6)
Asthma	3 (6)	1 (2)
Chronic obstructive pulmonary disease	2 (4)	1 (2)
Stroke/aneurysm	2 (4)	1 (2)
Heteroplasia/neoplasia	2 (4)	1 (2)
Hyperuricemia	0 (0)	3 (6)
Glaucoma	1 (2)	0 (0)
Multiple sclerosis	1 (2)	0 (0)
Cluster headache	1 (2)	0 (0)
Cushing disease	1 (2)	0 (0)
Sarcoidosis	1 (2)	0 (0)
Thalassemia	0 (0)	1 (2)
Rheumatoid arthritis	0 (0)	1 (2)
2 or more medical comorbidities	12 (24)	13 (26)
Mean biomarkers (SD)		
Hemoglobin, g/dL	13.91 (1.21)	13.93 (1.33)
Platelets, $n \times 10^3/mm^3$	235.42 (57.64)	232.96 (50.20)
Creatinine, mg/dL	0.94 (1.78)	0.95 (0.20)
Triglycerides, mg/dL	115.96 (52.91)	121.69 (58.68)
HDL cholesterol, mg/dL	51.98 (16.59)	46.51 (12.01)
LDL cholesterol, mg/dL	87.40 (25.48)	93.96 (29.25)
Total cholesterol, mg/dL	156.44 (31.07)	160.90 (37.45)
Glycated hemoglobin, mmol/mol	41.20 (8.36)	42.97 (10.21)
Fibrinogen, mg/dL	347.84 (66.04)	356.49 (68.28)
D-dimer, mg/L FEU	0.68 (1.39)	0.45 (0.39)
HRV[a], ms	51.10 (27.66)	41.50 (12.29)
C-reactive protein		
BO, mg/dL	0.19 (0.21)	0.39 (0.69)
TO, mg/L	0.28 (0.39)	0.64 (1.16)
Mean SQ (SD)		
Anxiety	8.60 (4.73)	7.24 (4.67)
Depression	7.92 (4.77)	6.90 (4.87)
Somatization	9.82 (5.65)	7.82 (5.12)
Hostility	4.70 (4.00)	5.34 (4.36)
Mean PWB (SD)		
Autonomy	62.20 (9.18)	61.80 (9.25)
Environmental mastery	55.28 (11.52)	55.32 (10.65)
Personal growth*	60.48 (9.88)	56.18 (10.50)
Positive relations with others	61.26 (13.26)	60.20 (10.68)
Purpose in life	56.80 (11.51)	56.22 (11.59)
Self-acceptance	54.48 (11.63)	55.80 (13.68)
Mean CID-20 (SD)		
CID-20 total score	38.18 (8.48)	36.20 (8.57)
Depression (DSM), n (%)	35 (70)	27 (54)

Table 1 (continued)

Variable	CBT/WBT group (n = 50)	CM group (n = 50)
Major depression	2 (4)	3 (6)
Minor depression	32 (64)	24 (48)
Dysthymia	1 (2)	0 (0)
History of depression (DSM), n (%)	34 (68)	26 (52)
Demoralization (DCPR), n (%)	47 (94)	44 (88)
History of demoralization (DCPR), n (%)	36 (72)	32 (64)
Comorbidities, n (%)		
Depression + demoralization*	32 (64)	21 (42)
Chronicity of depression/demoralization, n (%)		
Current + previous episode of depression	26 (52)	19 (38)
Current + previous episode of demoralization	35 (70)	31 (62)

ACS, acute coronary syndrome; AMI, acute myocardial infarction; CBT, Cognitive-Behavioral Therapy; CID-20, 20-item Clinical Interview for Depression; CM, clinical management; DCPR, diagnostic criteria for psychosomatic research; GRACE, Global Registry of Acute Coronary Events; HRV, heart rate variability; LVEF, left ventricular ejection fraction; NSTEMI, non-ST-segment elevation myocardial infarction; PTCA, percutaneous transluminal coronary angioplasty; PWB, Psychological Well-Being scales; SQ, Symptom Questionnaire; STEMI, ST-segment elevation myocardial infarction; WBT, Well-Being Therapy; BO, Bologna; TO, Torino. * $p \leq 0.05$.
[a] Assessed only in Torino.

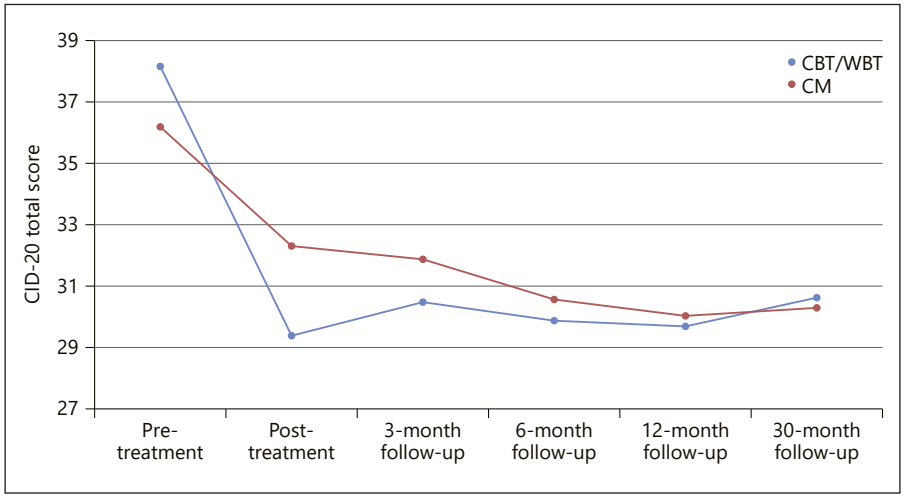


Fig. 2. CID-20 total scores at different time points (intention-to-treat analysis).

As for the CID-20 total score, a significant interaction between group allocation and time was found (F = 2.75; df = 3.85; $p < 0.05$; Fig. 2). Significant decreases in symptom scores from pre- to posttreatment were found in both the CBT/WBT ($p < 0.001$) and CM ($p < 0.01$) groups. However, the effect sizes for score modifications were strong in the CBT/WBT treatment group (Cohen's d = 1.161 and 1.393, respectively) and weak/medium among CM patients (Cohen's d = 0.492 and 0.589, respectively) (Table 2). Patients allocated to CBT/WBT reported significant lower scores posttreatment (p = 0.040) compared to those assigned to CM. Starting from the 3-month follow-up, the CID-20 score differences between the 2 groups were no longer significant. The benefits, however, tended to persist in both groups.

No significant interactions were found between time and group allocation in relation to SQ and PWB mean scores, except for hostility as assessed by the SQ (F = 3.12;

Psychother Psychosom 2020;89:345–356
DOI: 10.1159/000510006

Table 2. Effects of treatment groups on psychological characteristics

Variable	Pretreatment	Posttreatment	3-month follow-up	6-month follow-up	12-month follow-up	30-month follow-up	Time × group			Cohen's *d**	Within-group score change*, [a]
							F	df	*p* value		
Intention-to-treat analysis											
CBT/WBT group (*n* = 50), mean (SD)											
PWB autonomy	62.20 (9.18)	64.58 (9.42)	64.54 (9.24)	64.40 (9.12)	65.50 (8.53)	64.93 (9.67)	0.173	3.846	0.948	−0.26	−2.38 (−5.51 to 0.76)
PWB environmental mastery	55.28 (11.52)	57.33 (12.93)	59.48 (11.32)	58.02 (11.83)	58.36 (12.15)	58.69 (10.97)	0.309	4.353	0.886	−0.17	−2.09 (−5.61 to 1.43)
PWB personal growth	60.48 (9.88)	61.46 (9.92)	61.95 (9.91)	60.79 (9.58)	60.55 (9.54)	59.94 (9.34)	0.982	4.253	0.420	−0.10	−0.93 (−3.91 to 2.06)
PWB positive relations	61.26 (13.26)	61.82 (13.50)	61.88 (12.86)	60.60 (13.08)	61.27 (12.08)	60.48 (11.60)	0.709	4.183	0.592	−0.04	−0.57 (−3.33 to 2.19)
PWB purpose in life	56.80 (11.51)	57.31 (11.21)	58.35 (10.09)	57.88 (10.85)	57.42 (9.81)	57.63 (9.70)	1.104	3.803	0.353	−0.04	−0.49 (−4.14 to 3.17)
PWB self-acceptance	54.48 (11.63)	55.70 (14.36)	57.59 (13.51)	55.83 (14.19)	56.66 (11.92)	56.15 (13.90)	1.593	4.325	0.170	−0.09	1.30 (−4.48 to 1.89)
SQ anxiety	8.60 (4.73)	7.04 (5.23)	6.60 (4.87)	6.67 (4.19)	6.62 (4.51)	6.00 (4.35)	1.008	4.180	0.405	0.31	1.54 (−0.10 to 3.19)
SQ depression	7.92 (4.77)	7.21 (5.42)	6.38 (5.03)	7.06 (5.22)	6.91 (5.08)	5.99 (4.64)	0.605	4.180	0.667	0.14	0.70 (−0.98 to 2.37)
SQ somatization	9.82 (5.65)	8.80 (5.73)	8.67 (5.42)	8.96 (5.02)	9.49 (5.19)	8.17 (5.00)	0.787	3.981	0.534	0.18	1.04 (−0.75 to 2.84)
SQ hostility	4.70 (4.00)	5.19 (4.96)	5.18 (4.46)	4.41 (3.71)	5.32 (4.71)	3.81 (3.37)	3.121	4.288	0.013	−0.11	−0.51 (−1.91 to 0.89)
CID-20 total score	38.18 (8.48)	29.39 (6.55)	30.48 (5.81)	29.89 (5.88)	29.70 (6.51)	30.64 (7.02)	2.748	3.853	0.030	1.16	8.73 (5.39 to 12.07)
CM group (*n* = 50), mean (SD)											
PWB autonomy	61.80 (9.25)	62.82 (8.77)	63.20 (8.51)	63.21 (9.00)	64.57 (9.34)	63.71 (9.26)				−0.11	−1.02 (−4.16 to 2.11)
PWB environmental mastery	55.32 (10.65)	56.69 (8.81)	57.81 (10.15)	57.51 (8.78)	58.03 (11.19)	58.81 (8.10)				−0.14	−1.33 (−4.85 to 2.19)
PWB personal growth	56.18 (10.50)	56.54 (8.70)	56.67 (9.65)	57.10 (8.90)	57.64 (10.24)	57.00 (8.85)				−0.04	−0.41 (−3.40 to 2.57)
PWB positive relations	60.20 (10.68)	59.90 (10.93)	59.93 (12.13)	58.78 (10.82)	58.95 (11.54)	60.56 (10.78)				0.03	0.31 (−2.45 to 3.07)
PWB purpose in life	56.22 (11.59)	54.97 (9.41)	55.47 (10.32)	55.96 (10.12)	55.63 (10.82)	57.76 (8.16)				0.12	1.23 (−2.42 to 4.89)
PWB self-acceptance	55.80 (13.68)	56.03 (11.52)	57.86 (12.84)	58.32 (12.39)	59.69 (13.38)	59.94 (10.52)				−0.02	−0.15 (−3.34 to 3.04)
SQ anxiety	7.24 (4.67)	6.39 (4.41)	6.13 (4.21)	7.10 (5.14)	6.33 (5.09)	5.69 (4.07)				0.19	0.87 (−0.78 to 2.51)
SQ depression	6.90 (4.87)	5.94 (4.22)	5.83 (4.75)	6.80 (5.45)	6.22 (5.09)	5.83 (4.18)				0.21	0.98 (−0.69 to 2.66)
SQ somatization	7.82 (5.12)	8.24 (4.90)	7.87 (4.58)	8.15 (5.64)	7.90 (5.38)	7.61 (4.72)				−0.08	−0.44 (−2.23 to 1.36)
SQ hostility	5.34 (4.36)	4.12 (3.78)	4.71 (3.92)	6.01 (4.73)	5.17 (4.14)	4.56 (4.11)				0.30	1.24 (−0.16 to 2.64)
CID-20 total score	36.20 (8.57)	32.30 (7.26)	31.89 (7.11)	30.59 (7.28)	30.03 (7.05)	30.30 (6.82)				0.49	3.97 (0.63 to 7.31)

All analyses were adjusted for the GRACE index (6-month probability of cardiac mortality). CID-20, 20-item Clinical Interview for Depression; CBT, Cognitive-Behavioral Therapy; CM, clinical management; PWB, Psychological Well-Being scales; SQ, Symptom Questionnaire; WBT, Well-Being Therapy. * Pre-/posttreatment scores change. [a] Values are expressed as mean differences (95% CI).

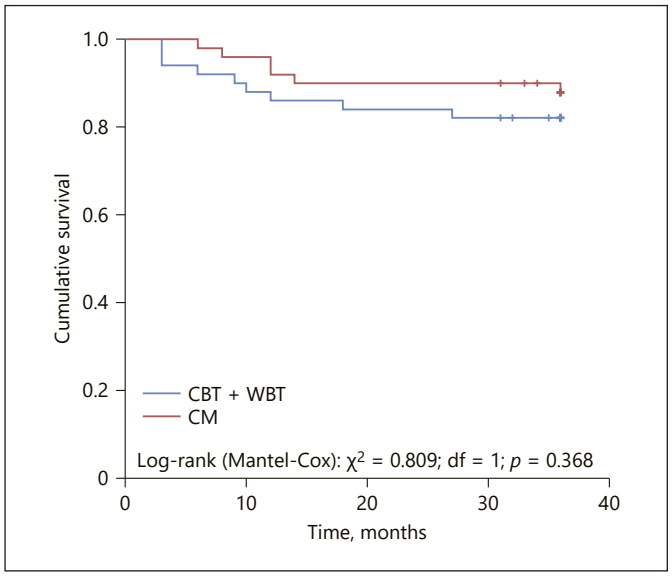


Fig. 3. Survival curves of the CBT/WBT and CM groups.

Color version available online

df = 4.29; $p < 0.05$), with CM group showing significantly higher scores at the 6-month follow-up than CBT/WBT ($p = 0.039$; Table 2).

Biomarkers

At the 3-month post-intervention follow-up, we observed a significant reduction of the frequencies of patients with biomarker levels considered to be at risk (below or above the median) only among patients allocated to the CBT/WBT group. In particular, we found a significant decrease in cases with a high platelet count (from 52 to 36%; $p < 0.05$; median = $226 \times 10^3/mm^3$), lower HDL cholesterol (from 52 to 34%; $p < 0.05$; median = 47 mg/dL), and a higher D-dimer level (from 56 to 40%; $p < 0.05$; median = 0.31 mg/L FEU) in patients assigned to CBT/WBT compared to those receiving CM. No significant decrease in patients with risky levels of biomarkers was observed in the CM group.

Survival Analyses

Within 36 months from baseline, 15% of the total sample had an adverse cardiac outcome. As for cardiac morbidity and mortality, we did not find any significant difference between the CBT/WBT and CM groups in terms of survival. Indeed, among the patients allocated to CBT/WBT 16% ($n = 8$) had nonfatal cardiac events and 1 patient (2%) had a cardiac death (occurring after 18 months from baseline), whereas among the CM patients 10% ($n = 5$) had nonfatal events and 1 patient (2%) had a cardiac death (after 36 months from baseline). Nonetheless, the CBT/WBT patients displayed most of the negative cardiac outcomes within the first 9 months, with almost half of them (4 out of 9) relapsing during treatment sessions. On the contrary, CM participants were more likely to relapse after a longer period (starting after 8 months from baseline) (Fig. 3).

Stratifying the sample by group allocation, among CBT/WBT patients the probabilities of cardiac death, both in hospital (Wald = 4.235; df = 1; HR = 1.040; 95% CI 1.002–1.079; $p = 0.040$) and at 6 months postdischarge (Wald = 4.594; df = 1; HR = 1.031; 95% CI 1.003–1.060; $p = 0.032$) as calculated with GRACE indices, were found to predict a worse cardiac prognosis. On the contrary, in the CM group adverse cardiac outcomes were predicted by baseline scores of depression, as assessed by CID (Wald = 5.540; df = 1; HR = 1.204; 95% CI 1.031–1.404; $p = 0.019$).

Discussion

To our knowledge, this is the first RCT demonstrating a significant improvement in depressive symptoms and biomarkers in patients with ACS following sequential CBT/WBT when compared with CM. This study provides new important clinical insights regarding the treatment of depression in the setting of ACS. The sequential combination of CBT/WBT was effective in significantly decreasing depressive symptoms compared to CM. In both groups the benefits persisted at follow-up, even though the differences between them faded (Fig. 2). It is noteworthy the different trend observed in the 2 groups concerning hostility, since it represents a key variable in the literature on the psychological issues embedded in depressive states [29] and it has been found to have a negative effect on the cardiac prognosis [30].

Medical outcomes did not differ between the 2 groups, yet among the CBT/WBT patients a negative cardiac prognosis was associated with a greater severity of the cardiac illness (as indicated by the GRACE indexes and the timing of relapses), whereas in the CM group it was associated with the severity of baseline depressive symptomatology. Moreover, patients who were assigned to the treatment group displayed significant decreases in placement according to normative values of platelet counts, HDL cholesterol, and D-dimer. There is evidence that these biomarkers may indicate a prognostic significance of the occurrence of cardiovascular events [31–33].

The findings are important in view of the methodology that was used. The patients were not assessed during hospitalization but rather after 1 month, when stress linked to hospitalization and the impact of acute illness are likely to subside and the evaluation of depressive symptoms is likely to be more reliable [34]. The impact of the CBT/WBT sequential combination was not compared to treatment as usual, as occurred in other studies [6], but rather to CM, where patients received the nonspecific elements of psychotherapy [27, 35]. Indeed, also CM yielded significant improvement in affective symptoms. This indicates that nonspecific support after ACS may be important, but specific psychotherapeutic strategies are associated with greater benefits and it underlines the need to schedule booster sessions (i.e., WBT or brief CBT) in order to reinforce progress or address potential obstacles to the continuance of the positive changes made during the therapy.

WBT is a short-term psychotherapeutic strategy that emphasizes self-observation of psychological well-being via the use of a structured diary, cognitive restructuring of interfering thoughts and/or behaviors, and homework assignments [9, 10]. The working hypothesis was that lifestyle changes could only be achieved with a personalized approach that targets psychological well-being [9]. Based on examples taken from post-ACS everyday life, the patients allocated to CBT/WBT were instructed on how to overcome specific obstacles concerning lifestyle (i.e., specific strategies for medication adherence, scheduling of gradual physical exercises, and dietary modification according to specific prescriptions following hospital guidelines). In the phase that immediately follows ACS, interventions that bring the person out of negative functioning and distress may be important, and this was the target of the first phase of psychotherapy (CBT). However, facilitating progression toward restoration of the positive ("there is life after ACS") and appreciation of healthy lifestyle is another target that requires specific interventions (WBT). The results of this investigation confirm previous studies on the role of psychotherapeutic strategies in the setting of ACS [6] and provide a valid alternative/integration to pharmacological strategies, which carry the disadvantages of side effects of antidepressant drugs [36–37], with particular reference to cardiovascular safety [38]. The sequential psychotherapeutic strategy that was used may also be applied after pharmacological treatment of depression, if appropriate, and may have potential in extending therapeutic benefits beyond the time of medication administration, as it has been found to be the case in psychiatric settings [39].

This therapeutic approach may be potentially extended to cardiovascular rehabilitation in view of the suitability of WBT for the rehabilitation process [40] and the adverse prognostic role of an unhealthy lifestyle and depressive symptoms in these settings [41–43]. A number of clinical situations (delayed recovery after treatment, discrepancy between cardiovascular status/functioning, presence of a psychological comorbidity, problems with lifestyle and risky behavior, and presence of stressful circumstances) may be addressed by the sequential strategy we have outlined.

The findings of this investigation targeting psychological well-being in ACS should be seen as preliminary and await proper replication studies. It should also be noted that more than a quarter of the ACS patients diagnosed with depression and/or demoralization (36 out of 136; 26.5%) refused to join the RCT. This percentage, however, is lower than the refusal rates found in the literature on secondary prevention programs, which range from 31.4 [44] to 72.2% [45] among depressed patients. Moreover, about half of the 740 patients initially screened by the cardiologists refused to undergo psychological assessment and almost half of those who agreed refused to join the trial or revoked the initial consent. The results are thus likely to reflect a self-selected population. Nonetheless, they indicate a road to the practice of lifestyle medicine [46] that is worth perusing.

Acknowledgement

We thank all of the patients who agreed to be screened for the TREATED-ACS study and those who joined the trial. In addition, we thank Drs. Laura Alessi, Laura Staccini, Antonio Piolanti, Rachele Ceschin, Alessandra Munno, Emanuela Offidani, Corine Panepinto, Fedra Ottolini, Letizia Riva, and Daniela Calabrese, Mrs. Loretta Lollini, and all of the medical staff working at the Maggiore and Molinette hospitals for providing support.

Statement of Ethics

This study was approved by the institutional review board of the ethics committees of both centers (identifier: Studio CE 09058). Written informed consent was secured from all of the patients for both the initial psychological evaluation and trial participation, after the procedures had been fully explained to them. The participants did not receive any compensation. The authors assert that all of the procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Conflict of Interest Statement

The authors have no conflict of interests to declare.

Funding Sources

This work was supported by a grant from the Compagnia di San Paolo di Torino, Italy, to C.R.

Author Contributions

C.R., G.A.-D., and G.A.F. conceptualized and designed this study. C.R., S.G., G.A.-D., and G.A.F. collected, analyzed, and interpreted the data. C.R., S.G., and G.A.F. wrote the first draft of this paper. S.G. performed the statistical analyses. All of the authors critically revised this work for important intellectual content and provided administrative, technical, or material support. C.R., G.A.-D. and G.A.F. supervised the whole process.

References

1 Carney RM, Freedland KE. Depression and coronary heart disease. Nat Rev Cardiol. 2017 Mar;14(3):145–55.
2 Rafanelli C, Roncuzzi R, Milaneschi Y, Tomba E, Colistro MC, Pancaldi LG, et al. Stressful life events, depression and demoralization as risk factors for acute coronary heart disease. Psychother Psychosom. 2005 Apr;74(3):179–84.
3 Kuhlmann SL, Arolt V, Haverkamp W, Martus P, Ströhle A, Waltenberger J, et al. Prevalence, 12-month prognosis, and clinical management need of depression in coronary heart disease patients. Psychother Psychosom. 2019 Sep;88(5):300–11.
4 de Figueiredo JM, Frank JD. Subjective incompetence, the clinical hallmark of demoralization. Compr Psychiatry. 1982 Jul-Aug; 23(4):353–63.
5 Kim JM, Stewart R, Lee YS, Lee HJ, Kim MC, Kim JW, et al. Effect of escitalopram vs placebo treatment for depression on long-term cardiac outcomes in patients with acute coronary syndrome: a randomized clinical trial. JAMA. 2018 Jul;320(4):350–8.
6 Reavell J, Hopkinson M, Clarkesmith D, Lane DA. Effectiveness of cognitive behavioral therapy for depression and anxiety in patients with cardiovascular disease: A systematic review and meta-analysis. Psychosom Med. 2018 Oct;80(8):742–53.
7 Berkman LF, Blumenthal J, Burg M, Carney RM, Catellier D, Cowan MJ, et al.; Enhancing Recovery in Coronary Heart Disease Patients Investigators (ENRICHD). Effects of treating depression and low perceived social support on clinical events after myocardial infarction: the Enhancing Recovery in Coronary Heart Disease Patients (ENRICHD) Randomized Trial. JAMA. 2003 Jun;289(23):3106–16.
8 Kubzansky LD, Huffman JC, Boehm JK, Hernandez R, Kim ES, Koga HK, et al. Positive psychological well-being and cardiovascular disease: JACC health promotion series. J Am Coll Cardiol. 2018 Sep;72(12):1382–96.
9 Fava GA. Well-being therapy: Treatment manual and clinical applications. Basel, Switzerland: Karger Medical & Scientific Publishers; 2016. https://doi.org/10.1159/isbn.978-3-318-05822-2.
10 Guidi J, Rafanelli C, Fava GA. The clinical role of well-being therapy. Nord J Psychiatry. 2018 Aug;72(6):447–53.
11 Fava GA, Rafanelli C, Grandi S, Conti S, Belluardo P. Prevention of recurrent depression with cognitive behavioral therapy: preliminary findings. Arch Gen Psychiatry. 1998 Sep; 55(9):816–20.
12 American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th ed., revised. Washington (DC): American Psychiatric Association; 2000.
13 Fava GA, Freyberger HJ, Bech P, Christodoulou G, Sensky T, Theorell T, et al. Diagnostic criteria for use in psychosomatic research. Psychother Psychosom. 1995;63(1):1–8.
14 Tang EW, Wong CK, Herbison P. Global Registry of Acute Coronary Events (GRACE) hospital discharge risk score accurately predicts long-term mortality post acute coronary syndrome. Am Heart J. 2007 Jan;153(1):29–35.
15 First MB, Spitzer RL, Gibbon M, Williams J. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, research version. New York (NY): Biometrics Research, New York State Psychiatric Institute; 2002.
16 Porcelli P, Sonino N. Psychological factors affecting medical conditions: a new classification for DSM-V. Adv Psychosom Med. Basel: Karger; 2007.
17 Tecuta L, Tomba E, Grandi S, Fava GA. Demoralization: a systematic review on its clinical characterization. Psychol Med. 2015 Mar; 45(4):673–91.
18 Galeazzi GM, Ferrari S, Mackinnon A, Rigatelli M. Interrater reliability, prevalence, and relation to ICD-10 diagnoses of the Diagnostic Criteria for Psychosomatic Research in consultation-liaison psychiatry patients. Psychosomatics. 2004 Sep-Oct;45(5):386–93.
19 Paykel ES. The clinical interview for depression. Development, reliability and validity. J Affect Disord. 1985 Jul;9(1):85–96.
20 Guidi J, Fava GA, Bech P, Paykel E. The Clinical Interview for Depression: a comprehensive review of studies and clinimetric properties. Psychother Psychosom. 2011 Dec;80(1): 10–27.
21 Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol. 1967 Dec;6(4):278–96.
22 Carrozzino D, Patierno C, Fava GA, Guidi J. The Hamilton Rating Scales for Depression: a critical review of clinimetric properties of different versions. Psychother Psychosom. 2020 May;89(3):133–50.
23 Kellner R. A symptom questionnaire. J Clin Psychiatry. 1987 Jul;48(7):268–74.
24 Benasi G, Fava GA, Rafanelli C. Kellner's Symptom Questionnaire, a highly sensitive patient-reported outcome measure: systematic review of clinimetric properties. Psychother Psychosom. 2020 Mar;89(2):74–89.
25 Ryff CD. Happiness is everything, or is it? Explorations on the meaning of psychological well-being. J Pers Soc Psychol. 1989;57(6): 1069–81.
26 Ryff CD. Psychological well-being revisited: advances in the science and practice of eudaimonia. Psychother Psychosom. 2014 Dec; 83(1):10–28.
27 Guidi J, Brakemeier EL, Bockting CL, Cosci F, Cuijpers P, Jarrett RB, et al. Methodological recommendations for trials of psychological interventions. Psychother Psychosom. 2018 Sep;87(5):276–84.
28 Cohen J. Statistical power analysis for the behavioral sciences. New York (NY): Routledge Academic; 1988.
29 Brummett BH, Babyak MA, Barefoot JC, Bosworth HB, Clapp-Channing NE, Siegler IC, et al. Social support and hostility as predictors of depressive symptoms in cardiac patients one month after hospitalization: a prospective study. Psychosom Med. 1998 Nov-Dec;60(6): 707–13.
30 Rafanelli C, Gostoli S, Tully PJ, Roncuzzi R. Hostility and the clinical course of outpatients with congestive heart failure. Psychol Health. 2016 Oct;31(2):228–38.
31 von Känel R. Acute mental stress and hemostasis: when physiology becomes vascular harm. Thromb Res. 2015 Feb;135(1):S52–5.
32 Vinholt PJ, Hvas AM, Frederiksen H, Bathum L, Jørgensen MK, Nybo M. Platelet count is associated with cardiovascular disease, cancer and mortality: a population-based cohort study. Thromb Res. 2016 Dec;148:136–42.

33 Ishida M, Itoh T, Nakajima S, Ishikawa Y, Shimoda Y, Kimura T, et al. A low early high-density lipoprotein cholesterol level is an independent predictor of in-hospital death in patients with acute coronary syndrome. Intern Med. 2019 Feb;58(3):337–43.
34 Fava GA, Sonino N. Depression associated with medical illness. CNS Drugs. 1996 Oct; 5(3):175–89.
35 Fava GA, Guidi J, Rafanelli C, Rickels K. The clinical inadequacy of the placebo model and the development of an alternative conceptual framework. Psychother Psychosom. 2017 Nov;86(6):332–40.
36 Carvalho AF, Sharma MS, Brunoni AR, Vieta E, Fava GA. The safety, tolerability and risks associated with the use of newer generation antidepressant drugs: a critical review of the literature. Psychother Psychosom. 2016; 85(5):270–88.
37 Fava GA, Rafanelli C. Iatrogenic factors in psychopathology. Psychother Psychosom. 2019 Jun;88(3):129–40.
38 Grace SL, Medina-Inojosa JR, Thomas RJ, Krause H, Vickers-Douglas KS, Palmer BA, et al. Antidepressant use by class: association with major adverse cardiac events in patients with coronary artery disease. Psychother Psychosom. 2018 Mar;87(2):85–94.
39 Guidi J, Tomba E, Fava GA. The sequential integration of pharmacotherapy and psychotherapy in the treatment of major depressive disorder: a meta-analysis of the sequential model and a critical review of the literature. Am J Psychiatry. 2016 Feb;173(2):128–37.
40 Nierenberg B, Mayersohn G, Serpa S, Holovatyk A, Smith E, Cooper S. Application of well-being therapy to people with disability and chronic illness. Rehabil Psychol. 2016 Feb;61(1):32–43.
41 Rafanelli C, Roncuzzi R, Finos L, Tossani E, Tomba E, Mangelli L, et al. Psychological assessment in cardiac rehabilitation. Psychother Psychosom. 2003 Nov-Dec;72(6):343–9.
42 Gostoli S, Roncuzzi R, Urbinati S, Morisky DE, Rafanelli C. Unhealthy behaviour modification, psychological distress, and 1-year survival in cardiac rehabilitation. Br J Health Psychol. 2016 Nov;21(4):894–916.
43 Gostoli S, Roncuzzi R, Urbinati S, Rafanelli C. Clinical and subclinical distress, quality of life and psychological well-being after cardiac rehabilitation. Appl Psychol Health Well-Being. 2017 Nov;9(3):349–69.
44 Turk-Adawi KI, Oldridge NB, Tarima SS, Stason WB, Shepard DS. Cardiac rehabilitation enrollment among referred patients: patient and organizational factors. J Cardiopulm Rehabil Prev. 2014 Mar-Apr;34(2):114–22.
45 Zullo MD, Gathright EC, Dolansky MA, Josephson RA, Cheruvu VK, Hughes JW. The influence of depression on utilization of cardiac rehabilitation post-myocardial infarction: a study of 158,991 Medicare beneficiaries. J Cardiopulm Rehabil Prev. 2017 Jan;37(1):22–9.
46 Rippe JM. Are we ready to practice lifestyle medicine? Am J Med. 2019 Jan;132(1):6–8.

Available online at www.sciencedirect.com

ScienceDirect

Comprehensive Psychiatry 74 (2017) 61–69

COMPREHENSIVE PSYCHIATRY

www.elsevier.com/locate/comppsych

Does psychological well-being change following treatment? An exploratory study on outpatients with eating disorders

Elena Tomba[a,*], Lucia Tecuta[a], Romana Schumann[b], Donatella Ballardini[b]

[a]Department of Psychology, University of Bologna, Bologna, Italy
[b]Eating Disorder Clinic "Centro Gruber", Bologna, Italy

Abstract

Psychological well-being changes following cognitive-behavioral therapy-based treatment were investigated in outpatients with eating disorders (ED). While it is known that CBT reduces symptomatology in EDs, less is known about how changes in positive functioning may ensue. One-hundred and eighty five ED outpatients were analyzed for pre-treatment and post-treatment changes in psychological well-being (PWB) by last observation carried forward – Wilcoxon signed rank tests. Significant gains in all PWB dimensions were found, with moderate effect size correlations in environmental mastery ($r = -.418$), personal growth ($r = -.351$) and self-acceptance ($r = -.341$). A subsample of patients in remission ($n = 51$) was selected and compared to healthy controls in PWB post-treatment scores through Mann–Whitney U tests. Remitted patients showed significantly lower psychological well-being in two dimensions compared to controls: PWB-positive relations ($r = -.360$) and PWB-self-acceptance ($r = -.288$). However, more than 50% of ED outpatients in remission had PWB scores that fell below the 50th percentile of healthy controls in all psychological well-being dimensions, despite significant treatment response. Several mechanisms of psychological well-being change following CBT-based treatment are discussed. The assessment of treatment outcome in EDs may benefit from considering changes in positive functioning such as psychological well-being, in addition to the standard measurement of BMI, symptomatology and behavioral parameters. CBT-based treatment outcomes may be strengthened by promoting the development of optimal domains particularly in the interpersonal realm, such as building of quality and warm relationships and focusing on enhancing self-acceptance.

1. Introduction

Widely used criteria for treatment response in eating disorders are typically focused on the reduction of psychopathological symptomatology as well as modifications of physical and behavioral aspects [1,2]. Recently, the concept of euthymia, a term used in psychiatry to define when patients no longer meet criteria of a disorder, has been revisited by authors who recommend the inclusion of positive gains as criteria for mental health outcomes [3]. Indeed, gains in positive functioning are frequently not considered, despite findings indicating the persistence of impairment in positive qualities such as psychological well-being in various psychiatric illnesses including eating disorders (EDs) [4]. Such impairments may be associated with increased vulnerability to future adversity [5,6].

The focus on the improvement of positive characteristics falls within the realm of positive psychology, whose purpose is to broaden definitions of mental health towards the inclusion of optimal functioning and a focus on building the best qualities. Positive psychology represents a clear paradigm shift from the historical negativity bias of the fields of psychiatry and psychology originally based on the disease model [7] and from a preoccupation only with repairing the worst aspects of life [8]. Such a shift has profound clinical implications. Bringing an individual out of negative functioning is one form of success, while facilitating progression towards the restoration of positive functioning is quite another [8,9]. Indeed, the World Happiness Report 2016 Update [10] found that autonomy, positive affect, generosity and social support were highly associated with quality of life and overall well-being, a correlation which was non-existent with negative affect, commonly the sole target for psychiatric disorder treatment.

Several definitions of well-being in psychology have been attempted, the most widespread being various conceptualizations

* Corresponding author at: Department of Psychology, Viale Berti, Pichat 5, 40127 Bologna, Italy.
E-mail address: elena.tomba@unibo.it (E. Tomba).

http://dx.doi.org/10.1016/j.comppsych.2017.01.001
0010-440X/

of *subjective well-being*, that is an individuals' perceptions and evaluations of their own lives in terms of their psychological and social functioning and affective states. Subjective well-being is frequently limited to and operationalized as the presence of positive emotions and satisfaction [8]. Alternatively, one of the most rigorously tested models of *psychological well-being*, introduced by Ryff [11], synthesizes clinical and personality theorists' conceptions of positive functioning, constituting a more comprehensive model of well-being. Ryff's model expands subjective well-being definitions towards a broader conceptualization of positive functioning. The author's eudaimonic model [11] derived from Jahoda's mental health criteria [12] considers specific domains that contribute to the development of optimal functioning and the fulfillment of one's potential including six interrelated psychological dimensions: autonomy, self-acceptance, a sense of continued growth and development, the belief that life is purposeful and meaningful, quality relationships with others, and the capacity to master effectively one's environment. A number of clinical studies on psychiatric populations have yielded substantial support to this model, finding that psychological well-being was frequently impaired [9,13,14]. Unlike other well-being definitions, this model of psychological well-being has not only been operationalized into an empirically validated instrument, Ryff's Psychological Well-Being Scales [11] but is also the theoretical basis for a specific psychotherapeutic intervention, well-being therapy [5], recently tested in EDs [15].

However, to date when positive mental health has been considered in eating disorders, there has been an exclusive focus on well-being in terms of subjective well-being and quality of life [16–19]. In our previous controlled study, a paucity of optimal positive functioning, in terms of psychological well-being, correlated with eating disorder symptomatology in ED out-patients. The lack of psychological well-being was independent of the presence of psychological distress and the severity of the disorder [4]. Furthermore, very few studies on EDs have focused on gains in positive functioning following treatment. Recently, a study investigated how a specific cognitive-behavioral intervention (Identity Intervention Program) increased positive self-schemas and psychological well-being in a sample of ED patients [20]. No studies have instead explored if such changes in psychological well-being are observable following first-line standard CBT-based and nutritional rehabilitation treatment for EDs [21], and whether such changes reach healthy levels. Therefore, the aim of this study is to examine whether dimensions of psychological well-being in ED out-patients change after CBT-based and nutritional rehabilitation treatment and whether any observed gains in remitted patients reach optimal levels found in matched healthy controls.

2. Materials and methods

2.1. Sample

A convenience sample of consecutively screened female out-patients (n = 195) who met DSM IV-TR diagnostic criteria for EDs [22], began an integrated treatment in an ED specialized out-patient clinic (Bologna, Italy). ED diagnoses were established at intake by the consensus of a psychiatrist and a clinical psychologist independently using the Structured Clinical Interview for DSM-IV (SCID) [23]. Patients who completed treatment were evaluated for remission from EDs status through face-to-face interviews by the consensus of a psychiatrist and the treating psychotherapist independently. Remission was defined as diagnostic, behavioral, BMI and psychological criteria, following Ackard et al.'s [24] suggestions, specifically in our study as: no current eating disorder according to DSM-IV-TR criteria using the SCID; body mass index (BMI) (kg/m^2) improvement; full absence of binge-eating, purging or fasting (in the past three months); a total score below the cut-off of 30 by the end of treatment in the Eating Attitudes Test [25].

Fifty-five controls matched for socio-demographic characteristics (sex and age) were recruited by advertisements from the general population and included in the study after excluding the presence of current psychiatric disorders with SCID face-to-face interviews.

2.2. Treatment

Treatment consisted in cognitive-behavioral therapy for EDs [26] and a nutritional rehabilitation program, as recommended by the practice guidelines for the treatment of ED patients [21]. Treatment was provided in a multidisciplinary clinical outpatient setting and consisted in individual weekly sessions composed of one hour psychotherapy sessions, provided by PsyD-level psychotherapists trained in CBT, either preceded or followed by one-hour sessions with a nutritional physician specialized in EDs. Average duration of treatment is one year, roughly 48 sessions. Main elements of the treatment include cognitive-behavioral techniques such as cognitive restructuring, assertiveness training, self-monitoring through the use of the diary, behavioral homework, exposure to avoided foods, and nutritional rehabilitation elements, such as psychoeducation on nutrition, on weight restoration, on health consequences of the illness. Integration is facilitated by daily case discussions between psychotherapists and nutritional physicians of clinically useful information in order to tailor sessions to emerging themes and needs.

Treatment integrity was checked by submitting eight randomly selected recorded sessions to two independent assessors, who correctly identified presence of treatment elements (see above) in the course of the provided sessions, by using a checklist created ad-hoc by the multidisciplinary équipe.

The study, conducted between January 2012 and December 2014, was approved by an institutional review board and informed consent was obtained from all participants.

2.3. Assessment

All ED patients were evaluated using the following self-report instruments at intake, at session number 16, 32,

and finally at session 48, considered the post-treatment assessment session:

The Eating Attitude Test (EAT-40) is a 40-item screening measure designed to identify behaviors and cognitive patterns of EDs. Items are constructed on a 0–3 four-point Likert scale, yielding three subscale scores for dimensions of dieting, body, and food preoccupation, oral control, and a total score ranging from 0 to 120. Higher scores indicate higher eating-related pathology and greater risk of developing an ED in the general population. Cronbach's alpha coefficient ranged from .79 in female patients with AN to .94. Test–retest reliability was .84, and the validity coefficient was .87 (25). In this study, the Italian validation of the EAT-40 was applied, which reports subscale Cronbach alphas of .80 for dieting, .70 for bulimic preoccupations, and .83 for oral control [27].

The General Health Questionnaire (GHQ-30) [28] is a 30-item instrument with a four-point Likert scale with item scores ranging from 0 to 3, aimed at assessing depressive and anxiety symptoms, sleeping problems, social functioning, well-being, and coping abilities. A total score is used ranging from 0 to 90, with higher scores indicating greater mental distress. Cronbach's alpha coefficients tested in various empirical studies in community samples ranged from approximately .82–.93. Test–retest reliability coefficients varied from .50 to .90. In this study, the Italian version, with Cronbach alpha of .81, was applied [29].

The Psychological Well-being Scales (PWB) 84 items version [30] measures the six dimensions of PWB according to Ryff's model: autonomy, environmental mastery, personal growth, positive relationships with others, purpose in life, and self-acceptance. Items are constructed on a six-point 1–6 Likert scale, yielding six subscale scores ranging from 14 to 84. Cronbach's alpha coefficients in a sample of 321 individuals from the general population for the six scales ranged from .85 to .91. Test–retest reliability varied between .81 and .88, whereas validity correlations extended between .25 and.73 (30). PWB scales were also administered to controls matched for socio-demographic characteristics. The Italian version was used [31] with the following Cronbach alphas reported by Gremigni and Stuart-Brown [32]: autonomy – .86, environmental mastery – .78, personal growth – .75, positive relations – .84, purpose in life – .73, and self-acceptance – .71.

Socio-demographic characteristics (age, marital status, and education level) were also collected for all participants while clinical variables (BMI, illness duration) were also collected for patients. BMI was calculated at the intake visit by the nutritional physician and monthly during treatment.

2.4. Analyses

Descriptive analyses of the whole sample were run. The Kolmogorov–Smirnov normality test was performed and determined that the samples' scores in psychometric measures did not conform to a normal distribution. Therefore, changes between pre-treatment and post-treatment in PWB scales and eating disorder-related and general psychopathology in terms of EAT and GHQ scores were analyzed by Wilcoxon signed rank tests with last-observation carried forward (LOCF) for missing data. In order to address any baseline differences that may constitute sampling biases, χ^2 test and the independent t test were used to compare any treatment completers and drop-outs, test completers and non-completers.

Descriptive analyses were run to verify the clinical characteristics of remitted patients according to the applied remission criteria. Independent t test were run to compare remitted patients and controls in socio-demographic characteristics (age, education levels, and married status). To evaluate whether psychological well-being reached normative levels in remitted patients, Mann–Whitney U tests were conducted to compare their post-treatment PWB scale scores to those of healthy controls.

In order to better understand in clinical terms the resulting PWB scores after treatment in remitted patients, we first computed the quartiles of the scores for each of the six PWB dimensions of healthy controls to utilize as reference for optimal functioning. Secondly, we calculated how many remitted patients (reported in percentages) scored below and above the median score of the control group, corresponding to the 50th percentile.

All analyses were performed using SPSS version 20.0 with statistical significance set at a p level of .05. Effect sizes were calculated using Rosenthal's [33] formula for effect size correlations: $r = Z/\sqrt{N}$. Multiple testing was corrected statistically by adjusting for the number of hypothesis tests performed using Bonferroni single-step adjusted p values.

3. Results

Out of the 195 consecutively screened patients entering treatment, 76.9% completed treatment ($N = 150$) while 23.1% ($N = 45$) dropped out. Ten drop-out patients did not have any baseline data and were excluded from the analyses.

No statistically significant differences emerged in age, illness duration, BMI, and psychometric measures (PWB scale scores, GHQ and EAT total scores) between treatment completers and drop-outs and between the subsample that completed all assessment measures at post-treatment with those who did not. For descriptive data on controls, ED patients and, χ^2 and t test results see Table 1.

The patient sample entered in LOCF analyses ($n = 185$) was composed of 150 treatment completers and 35 drop-outs. In terms of ED diagnoses, 57 patients had AN, 32 BN, 62 BED and 34 had EDNOS with mean duration of

Table 1
Descriptive data of whole sample, treatment completers and drop-outs, test completers and non-completers.

Variables	Total control sample (N = 55)	Total ED sample M ± SD (N = 185)	Treatment completers M ± SD (n = 150)	Drop outs M ± SD (n = 35)	p[a]	Test completers M ± SD (n = 75)	Test non-completers M ± SD (n = 75)	p[b]
Age (y)	29.27 ± 10.42	28.89 ± 10.25	28.95 ± 10.30	27.09 ± 7.86	.979	27.97 ± 9.38	29.85 ± 11.09	.277
Educational level	5.4%middle school 53.6% high school 39.3% college	18.9% middle school 41.1% high school 34.1% college	17.9% middle school 46.4% high school 35.7% college	26.2% middle school 35.7% high school 38.1% college	.363	18.6% middle school 48.6% high school 32.9% college	17.7% middle school 43.5% high school 38.7% college	.776
ED diagnosis	–	AN: 17.3% BN: 30.8% BED:33.5% EDNOS:18.4%	AN:16.7% BN: 32.7% BED: 31.3% EDNOS:19.3%	AN:17.7% BN: 22.2% BED:40.0% EDNOS 13.3%	.523	AN: 17.3% BN: 34.7% BED: 32.0% EDNOS:16.0%	AN: 17.3% AN BN: 29.3% BED: 34.6% EDNOS: 18.6%	.697
Illness duration (y)	–	9.03 ± 8.24	9.12 ± 8.37	8.65 ± 7.13	.583	7.97 ± 7.38	10.75 ± 9.38	.060
BMI (kg/m^2)	21.24 ± 3.21	AN: 17.62 ± 4.74 BN: 22.11 ± 4.95 BED:33.42 ± 6.24 EDNOS: 22.28 ± 5.46	AN: 16.23 ± 1.76 BN: 22.47 ± 5.44 BED:33.20 ± 6.11 EDNOS: 22.89 ± 5.67	AN: 21.76 ± 7.95 BN: 20.66 ± 1.38 BED:33.97 ± 6.91 EDNOS: 19.41 ± 3.37	.291	AN:15.82 ± 1.72 BN:23.11 ± 6.21 BED: 33.95 ± 5.75 EDNOS: 22.850 ± 6.04	AN: 16.72 ± 1.77 BN: 21.67 ± 4.29 BED:32.30 ± 6.54 EDNOS: 22.95 ± 5.58	.738
EAT total score	7.46 ± 3.92	32.19 ± 20.61	32.41 ± 20.74	29.17 ± 19.67	.374	34.53 ± 21.54	31.48 ± 20.14	.398
GHQ total score	4.29 ± 5.49	12.98 ± 9.04	12.73 ± 9.39	12.81 ± 7.71	.960	13.15 ± 9.51	12.91 ± 9.40	.881
PWB autonomy	61.58 ± 12.10	51.89 ± 14.16	52.41 ± 14.02	50.49 ± 14.69	.686	52.41 ± 14.49	51.61 ± 13.88	.738
PWB environmental mastery	59.76 ± 10.25	47.32 ± 13.54	47.33 ± 13.76	44.09 ± 12.31	.657	45.48 ± 14.44	48.70 ± 12.77	.163
PWB personal growth	66.11 ± 8.92	60.72 ± 11.49	60.17 ± 11.67	62.20 ± 9.97	.142	59.61 ± 12.12	60.45 ± 11.52	.676
PWB positive relations	67.80 ± 11.62	59.34 ± 14.58	58.58 ± 13.34	61.56 ± 18.71	.060	57.12 ± 13.37	59.27 ± 13.30	.340
PWB life purpose	63.82 ± 11.30	55.36 ± 13.63	54.87 ± 13.45	54.69 ± 14.71	.181	53.75 ± 14.55	55.46 ± 12.39	.453
PWB self-acceptance	62.38 ± 12.60	46.88 ± 16.01	47.59 ± 16.41	40.73 ± 14.27	.614	44.79 ± 17.00	49.66 ± 15.54	.078

Abbreviations – AN, anorexia nervosa; BED, binge-eating disorder; BN, bulimia nervosa; BMI, body mass index; EAT, Eating Attitudes Test; ED, eating disorder; EDNOS, eating disorder not otherwise specified; GHQ, General Health Questionnaire; M, mean; PWB, Psychological Well-being Scales; SD, standard deviation; p, 2-tailed significance p value.

[a] Comparison between treatment completers and drop-outs.

[b] Comparison between test completers and non-completers.

Table 2
EAT, GHQ, and PWB scales changes between pre-treatment and post-treatment in ED outpatients (N = 185).

Outcome variable	Before treatment M ± SD	After treatment M ± SD	Before treatment median	After treatment median	Standardized test statistic (Z)	p	Effect size (r)
EAT total score	32.19 ± 20.61	25.28 ± 19.41	27.00	20.00	−5.835	<.0001*	−.429
GHQ total score	12.98 ± 9.04	9.80 ± 8.57	13.00	8.00	−5.779	<.0001*	−.425
PWB autonomy	51.89 ± 14.16	53.94 ± 14.40	53.00	55.00	−3.654	.0002*	−.269
PWB environmental mastery	47.32 ± 13.54	51.42 ± 13.70	46.00	51.00	−5.691	<.0001*	−.418
PWB personal growth	60.72 ± 11.49	63.62 ± 10.92	63.00	65.00	−4.769	<.0001*	−.351
PWB positive relations	59.34 ± 14.58	60.70 ± 12.71	59.00	63.00	−3.063	.002*	−.225
PWB life purpose	55.36 ± 13.63	57.85 ± 12.97	55.00	60.00	−3.893	<.0001*	−.286
PWB self-acceptance	46.88 ± 16.01	50.69 ± 14.98	47.00	51.00	−4.632	<.0001*	−.341

Abbreviations: EAT, Eating Attitudes Test; GHQ, General Health Questionnaire; M, mean; PWB, Psychological Well-being Scales; r, Pearson's r; SD, standard deviation; p, 2-tailed significance p value; Z, standardized Wilcoxon signed rank value.
* Statistical significance reached and maintained after alpha adjustments for multiple testing.

illness 9.03 ± 8.241 years. The mean age of the patient group was 28.89 ± 10.25 (age range 13–59 years). Mean body mass index (BMI) by diagnostic subgroup was 17.62 ± 4.74 kg/m^2 for AN, 22.10 ± 4.95 kg/m^2 for BN, 33.42 ± 6.24 kg/m^2 for BED, and 22.28 ± 8.19 kg/m^2 for EDNOS. About a fifth of patients were married (20.5%) and two-thirds (73%) were single. A third had a college degree (34.1%), 41.1% had a high school diploma, while 18.9% had completed middle school. Mean treatment length measured in number of sessions was 52.09 ± 24.69, while the median number of sessions was 46.

With regards to changes in psychological well-being over time (see Table 2), LOCF Wilcoxon signed rank tests showed significant improvements in ED outpatients in all psychological well-being scales with moderate effect sizes for environmental mastery (Z = −5.691; p < .0001; r = −.418), personal growth (Z = −4.769; p < .0001; r = −.351), and self-acceptance (Z = −4.632; p < .0001; r = −.341). Psychopathology in terms of EAT total score (Z = −5.835; p < .0001; r = −.429) and GHQ total score (Z = −5.779; p < .0001; r = −.425) significantly decreased with moderate effect sizes between pre- and post-treatment. All changes remained statistically significant when compared to the α critical level adjusted for multiple testing (p = .00625).

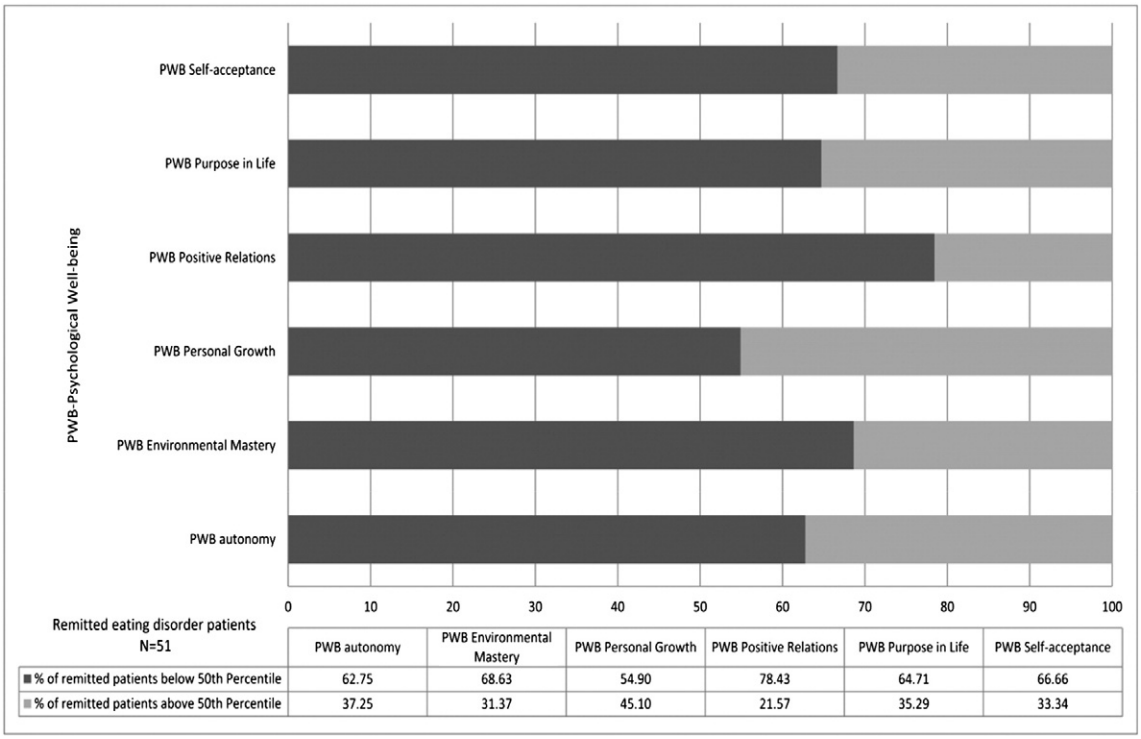
Fifty-one ED patients (34% of treatment completers) were defined as remitted following strict remission criteria [24]: 8 with AN, 20 with BN, 8 with BED, and 15 with EDNOS. Mean BMI at end of treatment was 21.49 ± 2.23 kg/m^2 for BN, 19.86 ± 1.31 kg/m^2 for AN, 23.37 ± 1.06 kg/m^2 for BED, and 21.76 ± 2.36 kg/m^2 for EDNOS. Remitted patients had a mean EAT total score of 14.13 ± 8.12 at post-treatment. Mean age was 27.71 ± 8.56 years with mean illness duration of 8.16 ± 6.97 years.

Remitted completers of treatment and healthy controls did not differ significantly in socio-demographic characteristics such as age, education levels, and married status. Mann–Whitney U tests between remitted completers of treatment (N = 51) and healthy controls (N = 55) (see Table 3) showed that post-treatment PWB median scores of patients in remission reached those of healthy controls with the exception of three scales that remain impaired and significantly lower with moderate and small effect sizes: PWB-autonomy (Z = −2.064, p = .039, r = −.200), PWB-positive relations (Z = −3.710, p = .002, r = −.360)

Table 3
Comparison between ED patients in remission (N = 51) and healthy controls (N = 55).

Outcome variable	Groups	n	Median	IR	Minimum	Maximum	Standardized test statistic (Z)	p	Effect size (r)
PWB autonomy	ED remitted patients	51	55.50	24	26	81	−2.064	.039	−.200
	Controls	55	62.00	16	23	81			
PWB environmental mastery	ED remitted patients	51	54.00	25	22	80	−1.863	.062	−.181
	Controls	55	61.00	16	28	77			
PWB personal growth	ED remitted patients	51	67.50	13	38	83	−.378	.706	−.037
	Controls	55	69.00	14	46	80			
PWB positive relations	ED remitted patients	51	62.00	20	35	83	−3.710	.002*	−.360
	Controls	55	71.00	13	25	84			
PWB purpose in life	ED remitted patients	51	61.00	22	25	77	−1.766	.077	−.172
	Controls	55	67.00	17	32	81			
PWB self-acceptance	ED remitted patients	51	55.00	24	22	79	−2.970	.003*	−.288
	Controls	55	62.00	17	35	102			

Abbreviations: ED, eating disorder; IR, interquartile range; M, mean; PWB, Psychological Well-being Scales; r, Pearson's r; SD, standard deviation; p, 2-tailed significance p value; Z, standardized Mann–Whitney U value.
* Statistical significance reached and maintained after alpha adjustments for multiple testing.



Remitted eating disorder patients N=51	PWB autonomy	PWB Environmental Mastery	PWB Personal Growth	PWB Positive Relations	PWB Purpose in Life	PWB Self-acceptance
% of remitted patients below 50th Percentile	62.75	68.63	54.90	78.43	64.71	66.66
% of remitted patients above 50th Percentile	37.25	31.37	45.10	21.57	35.29	33.34

Fig. 1. Percentages of remitted eating disorder patients below and above 50th percentile of Psychological Well-Being (PWB) control group scores.

and PWB-self-acceptance ($Z = -2.970$, $p = .003$, $r = -.288$). Applying adjustments to α levels for multiple testing (lowered critical $p = .008$), differences remain statistically significant ($p < .008$) in PWB-positive relations and PWB-self-acceptance.

While statistical analyses show that only two PWB scales remained impaired in remitted patients compared to controls, Fig. 1 illustrates that more than 50% of ED outpatients in remission have PWB scores that fall below the 50th percentile of healthy controls in all psychological well-being dimensions, despite significant treatment response.

4. Discussion

While it is known that CBT reduces symptomatology in EDs [1], less is known about how changes in positive functioning may ensue. Previous studies have investigated well-being impairment among ED patients and compared such scores to general population controls [4], as well as investigated psychological well-being after a psychotherapeutic intervention [20]. However, no studies had yet evaluated psychological well-being changes in ED patients following standard CBT-based treatment.

The current exploratory study, documenting remission rates similar to those reported by Ackard et al. [24], provides evidence of an improvement in psychological well-being in domains concerning autonomy, mastery over one's environment, positive evaluation of one's self, a sense of continued growth and development, the belief that life is purposeful and meaningful, and the possession of quality relationships with others following first-line ED treatment. Remitted ED patients, however, still exhibit impaired positive relations with others and self-acceptance after treatment compared to controls. Remitted patients did not greatly differ from controls in psychological well-being after treatment, but remained below healthy optimal levels. Such results are in line with previous studies exploring other positive gains in EDs, such as quality of life and social functioning, which improve with treatment but do not reach general population levels [2,34].

Several mechanisms of change following treatment in psychological well-being domains observed in the current study may be hypothesized based on the available literature on EDs. The increase in a sense of personal growth may be a consequence of CBT's efficacy in the reduction of symptoms [1] (such as food and body preoccupations, calorie counting, binging episodes, and obsessive thinking about food), giving patients a chance to invest their time and energy in themselves and their continued self-realization rather than in their disorder. Indeed, in AN patients, eating disorder symptoms were found to correlate significantly with a lack of existential well-being, encompassing also a lack of satisfaction in life progress [35].

The observed significant gains in environmental mastery in the current sample are in line with previous results in the literature in which a similar construct of self-efficacy improves in ED patients following cognitive behavioral-based treatment [36]. CBT may enhance environmental mastery by reducing

negative attributional styles and cognitions frequently concentrated in eating disorder patients on thoughts of self-blame and feelings of inadequacy and subjective incompetence. AN patients have been found to focus their ability to control and master their environment on their bodies and weight [38,39] which may be related to their perceived feelings of inadequacy and low ability to master their external lives and environment [4,37-39].

Autonomy was also found to significantly improve. In the individual suffering from an eating disorder, both psychodynamic and cognitive models underline the etiological role of the struggle for independence and autonomy, in the onset and maintenance of the disorder, respectively in terms of family enmeshment and problematic beliefs about the self and lack of control [22,40]. CBT through behavioral techniques such as self-monitoring, weekly weighing and regular eating may enhance autonomous behaviors and personal responsibility in the recovery process. Additionally, CBT may promote a sense of autonomy through cognitive strategies which reduce negative self-talk. However, the improvements in both environmental mastery and autonomy compared to controls were marginal. Such dimensions may require more specific and targeted intervention strategies than those offered in standard treatment in EDs for ratings to consistently reach and maintain healthy levels.

Purpose in life was found to improve in our sample. Women with AN symptoms commonly report low existential meaning in their lives compared to their unaffected peers [35] and may attribute purpose in life to their oral control [41]. Moreover, former ED patients identify reaching a healthy sense of purpose unrelated to eating behaviors and body issues as an important aspect in their quality of life post-treatment [42]. The ability to excessively self-regulate in terms of oral control and dietary restraints, is responsible for weight loss and may be linked with the patients' sense of accomplishment and self-worth, contributing to a greater distorted sense of purpose in life [38,41,43]. Through behavioral exercises aimed at exploring other qualities and abilities in life, cognitive restructuring, and nutritional rehabilitation, CBT integrated treatment may be able to break the dysfunctional association between meaning and purpose and dietary behaviors.

Wilson [44] had argued that self-acceptance, defined as holding a positive regard towards oneself, constitutes a pivotal aspect in EDs and should be considered a critical target for effective interventions. Self-acceptance improved in the current study but did not reach healthy levels in our remitted sample. CBT in general, in addition to primarily focusing on changing maladaptive behavior, includes several strategies which may promote acceptance of unchangeable aspects, such as body shape, in a healthy way. In our treatment, such strategies included education about body weight and eating and cognitive restructuring which modifies the patient's tendency of defining their self-worth in terms of body shape and weight [45], a hallmark feature of AN and BN [46].

Interpersonal maintenance models of eating disorders have recently been proposed according to which deficits in social skills, social anxiety, low assertiveness, and high aggressiveness play a role in the maintenance of EDs [47]. In our remitted patients, scores in positive quality relationships did not reach those of controls. This is consistent with previous findings according to which social functioning has been found to improve with treatment, but remain significantly below levels of the general population in spite of improvements in depressive and eating disorder symptomatology [48]. The observed improvements may be attributable to behavioral exposure to social situations and cognitive restructuring concerning dysfunctional thoughts regarding others.

From a clinical standpoint, the current exploratory study underscores an important finding, that is, the majority of patients considered remitted by strict criteria still exhibit impaired positive functioning. When psychological well-being remains compromised after standard pharmacological or psychotherapeutic treatment, much like the persistence of psychological symptoms [6,49] patients may be at higher risk of relapse [6,48-50] as reported in other remitted psychiatric populations [51–53]. Such considerations are important in EDs that are particularly hard to treat and in which drop-out and relapse are common phenomena [54,55]. More specifically, a higher dropout risk has been found to be associated with a lack of positive psychological characteristics, such as a low ability to pursue life goals [55].

The study has several limitations such as the lack of a control group to test whether gains in PWB may spontaneously happen in patients awaiting treatment. However, Ryff's theorized dimensions do not necessarily present fluctuations in the short-term, representing both state and trait elements [6]. Other limits include applying LOFC analyses for missing post-treatment data in psychometric measures in a large subset of treatment completers, and not controlling for psychopathology. While the small sample size and subsequent lack of adequate statistical power may be a limit, this lowers the risk for type I error possibly indicating replicable findings.

5. Conclusions

The assessment of treatment outcome in EDs may benefit from considering changes in positive functioning [2,34,40] such as psychological well-being, in addition to the standard measurement of BMI, symptomatology and behavioral parameters. CBT-based treatments for EDs may be strengthened by promoting the development of optimal domains particularly in the interpersonal realm, such as building of quality and warm relationships and focusing on enhancing self-acceptance.

As the development of novel approaches to improve recovery rates is needed in eating disorders [1], the integration of new psychotherapeutic strategies in standard treatment aimed at the promotion of balanced levels of psychological well-being may aid ED patients in their

recovery process. Future studies are warranted to investigate whether optimal functioning gains are maintained in the long-term. Moreover, future studies may explore the possible predictive role of psychological well-being in treatment outcomes in eating disorders.

Acknowledgment

The authors thank the therapists and the collaborators of the Eating Disorder Clinic "Centro Gruber" and the Fondazione Gruber.

References

[1] Hay P. A systematic review of evidence for psychological treatments in eating disorders: 2005–2012. Eat Disord 2013;46:462-9, http://dx.doi.org/10.1002/eat.22103.
[2] Ackard DM, Richter S, Egan A, Engel S, Cronemeyer CL. The meaning of (quality of) life in patients with eating disorders: a comparison of generic and disease-specific measures across diagnosis and outcome. Eat Disord 2014;47:259-67, http://dx.doi.org/10.1002/eat.22193.
[3] Fava GA, Bech P. The concept of euthymia. Psychother Psychosom 2015;85:1-5, http://dx.doi.org/10.1159/000441244.
[4] Tomba E, Offidani E, Tecuta L, Schumann R, Ballardini D. Psychological well-being in out-patients with eating disorders: a controlled study. Eat Disord 2014;47:252-8, http://dx.doi.org/10.1002/eat.22197.
[5] Fava GA. Well-being therapy. Treatment manual and clinical applications. Basel, Karger; 2016.
[6] Ryff C. Psychological well-being revisited: advances in the science and practice of eudaimonia. Psychother Psychosom 2014;83:10-28, http://dx.doi.org/10.1159/000353263.
[7] Hasler G. Well-being: an important concept for psychotherapy and psychiatric neuroscience. Psychother Psychosom 2016;85:255-61, http://dx.doi.org/10.1159/000447268.
[8] Snyder CR, & Lopez SJ, editors. Handbook of positive psychology. New York, New York: Oxford University Press; 2002.
[9] Fava GA, Tomba E. Increasing psychological well-being and resilience by psychotherapeutic methods. J Pers 2009;77:1903-34, http://dx.doi.org/10.1111/j.1467-6494.2009.00604.x.
[10] Helliwell J, Layard R, Sachs J. World happiness report 2016 update. New York, sustainable development solutions network, 1; 2016.
[11] Ryff C. Happiness is everything, or is it? Explorations on the meaning of psychological well-being. J Pers Soc Psychol 1989;57:1069-81.
[12] Jahoda M. Current concepts of positive mental health. New York, NY: Basic Books; 1958.
[13] Kennard BD, Emslie GJ, Mayes TL, Nakonezny PA, Jones JM, Foxwell AA, et al. Sequential treatment with fluoxetine and relapse-prevention CBT to improve outcomes in pediatric depression. Psychiatry 2014;171:1083-90.
[14] Stangier U, Hilling C, Heidenreich T, Risch AK, Barocka A, Schlösser R, et al. Maintenance cognitive-behavioral therapy and manualized psychoeducation in the treatment of recurrent depression: a multicenter prospective randomized controlled study. Psychiatry 2013;170:624-32.
[15] Tomba E, Tecuta L. Well-being therapy in a patient with anorexia nervosa. Psychother Psychosom 2016;85:369-70, http://dx.doi.org/10.1159/000448052.
[16] Doll HA, Petersen SE, Stewart-Brown SL. Eating disorders and emotional and physical well-being: associations between student self-reports of eating disorders and quality of life as measured by the SF-36. Qual Life Res 2005;14:705-17.
[17] Mond JM, Hay PJ, Rodgers B, Owen C, Beumont PJ. Assessing quality of life in eating disorder patients. Qual Life Res 2005;14:171-8.
[18] Jenkins PE, Hoste RR, Meyer C, Blissett JM. Eating disorders and quality of life: a review of the literature. Clin Psychol Rev 2001;31:113-21, http://dx.doi.org/10.1016/j.cpr.2010.08.003.
[19] Engel SG, Adair CE, Las Hayas C, Abraham S. Health related quality of life and eating disorders: a review and update. Eat Disord 2009;42:179-87, http://dx.doi.org/10.1002/eat.20602.
[20] Stein KF, Corte C, Chen DG, Nuliyalu U, Wing J. A randomized clinical trial of an identity intervention programme for women with eating disorders. Eur Eat Disord Rev 2013;21(2):130-42, http://dx.doi.org/10.1002/erv.2195.
[21] American Psychiatric Association. Practice guideline for the treatment of patients with eating disorders. 3rd ed. Washington: American Psychiatric Association; 2010.
[22] American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. Washington: American Psychiatric Association; 2000.
[23] First MB, Spitzer RL, Gibbon M, Williams JBW. Structured clinical interview for DSM-IV Axis I disorders, clinician version (SCID-CV). Washington, DC: American Psychiatric Press, Inc; 1996.
[24] Ackard DM, Richter SA, Egan AM, Cronemeyer CL. What does remission tell us about women with eating disorders? Investigating applications of various remission definitions and their associations with quality of life. J Psychosom Res 2014;76:12-8, http://dx.doi.org/10.1016/j.jpsychores.2013.10.002.
[25] Garner DM, Garfinkel PE. The eating attitude test: an index of the symptoms of anorexia nervosa. Psychol Med 1979;9:273-9.
[26] Fairburn CG. Eating disorders. In: Clark DM, & Fairburn CG, editors. Science and practice of cognitive behavior therapy. New York: Oxford University Press; 1997. p. 209-41.
[27] Cuzzolaro M, Petrilli A. Validazione della versione italiana dell'EAT-40. Psichiatria dell'Infanzia e dell'Adolescenza 1988;55:209-17.
[28] Goldberg D. The detection of psychiatric illness by questionnaire: a technique for the identification and assessment of non-psychotic psychiatric illness. London: Oxford University Press; 1972.
[29] Bellantuono C, Fiorio R, Zanotelli R, Tansella M. Psychiatric screening in general practice in Italy. A validity study of GHQ. Soc Psychiatry 1987;22:113-7.
[30] Ryff C, Singer B. Psychological well-being: meaning, measurements and implications of psychotherapy research. Psychother Psychosom 1996;65:14-23.
[31] Ruini C, Ottolini F, Rafanelli C, Tossani E, Ryff CD, Fava GA. The relationship of psychological well-being to distress and personality. Psychother Psychosom 2003;72:268-75.
[32] Gremigni P, Stuart-Brown S. Una misura del benessere mentale: validazione italiana della Warwick–Edinburgh mental well-being scale (WEMWBS). G Ital Psichiatr 2011;XXXVIII(2):543-63.
[33] Rosenthal R. Parametric measures of effect size. In: Cooper H, & Hedges LV, editors. The handbook of research synthesis. New York: Russell Sage Foundation; 1994. p. 231-44.
[34] Padierna A, Quintana JM, Arostegui I, Gonzalez N, Horcajo MJ. Changes in health related quality of life among patients treated for eating disorders. Qual Life Res 2002;11:545-52.
[35] Fox AP, Leung N. Existential well-being in younger and older people with anorexia nervosa – a preliminary investigation. Eur Eat Disord Rev 2009;17:24-30.
[36] Ohmann S, Popow C, Wurzer M, Karwautz A, Sackl-Pammer P, Schuch B. Emotional aspects of anorexia nervosa: results of prospective naturalistic cognitive behavioral group therapy. Neuropsychiatr 2013;27:119-28, http://dx.doi.org/10.1007/s40211-013-0065-7.
[37] Vitousek K, Watson S, Wilson GT. Enhancing motivation for change in treatment-resistant eating disorders. Clin Psychol Rev 1998;18:391-420.
[38] Horesh N, Zalsman G, Apter A. Internalized anger, self-control. Mastery experience in inpatient anorexic adolescents. J Psychosom Res 2000;49:247-53.

[39] Kazdin AE. Understanding how and why psychotherapy leads to change. Psychother Res 2009;19(4–5):418-28, http://dx.doi.org/10.1080/10503300802448899.
[40] Bruch H. Anorexia nervosa: therapy and theory. Psychiatry 1982;139:1531-8.
[41] Mond JM, Hay PJ, Rodgers B, Owen C, Beumont PJ. Assessing quality of life in eating disorder patients. Life Res 2005;14:171-8.
[42] de la Rie S, Noordenbos G, Donker M, van Furth E. The patient's view on quality of life and eating disorders. Eat Disord 2007;40:13-20.
[43] Nordbo RH, Espeset EM, Gulliksen KS, Skarderud F, Holte A. The meaning of self-starvation: qualitative study of patient's perception of anorexia nervosa. Eat Disord 2006;39:556-64.
[44] Wilson GT. Acceptance and change in the treatment of eating disorders and obesity. Behav Ther 1996;27:417-39.
[45] Fairburn CG, Marcus MD, Wilson GT. In: Fairburn CG, & Wilson GT, editors. Binge eating: nature, assessment and treatment cognitive behaviour therapy for binge eating and bulimia nervosa: a comprehensive treatment manual. New York: Guilford Press; 1993. p. 361-404.
[46] McFarlane T, McCabe RE, Jarry J, Olmsted MP, Polivy J. Weight-related and shape-related self-evaluation in eating-disordered and non-eating-disordered women. Eat Disord 2001;29:328-35.
[47] Arcelus J, Haslam M, Farrow C, Meyer C. The role of interpersonal functioning in the maintenance of eating psychopathology: a systematic review and testable model. Clin Psychol Rev 2012;33:156-67, http://dx.doi.org/10.1016/j.cpr.2012.10.009.
[48] Norman DK, Herzog DB. A 3-year outcome study of normal-weight bulimia: assessment of psychosocial functioning and eating attitudes. Psychiatry Res 1986;19:199-205.
[49] Lo Sauro C, Castellini G, Lelli L, Faravelli C, Ricca V. Psychopathological and clinical features of remitted anorexia nervosa patients: a six-year follow-up study. Eur Eat Disord Rev 2013;21:78-83.
[50] Guidi J, Tomba E, Fava GA. The sequential integration of pharmacotherapy and psychotherapy in the treatment of major depressive disorder: a meta-analysis of the sequential model and a critical review of the literature. Psychiatry 2016;173:128-37, http://dx.doi.org/10.1176/appi.ajp.2015.15040476.
[51] Rafanelli C, Park SK, Ruini C, Ottolini F, Cazzaro M, Fava GA. Rating well-being and distress. Stress Med 2000;16:55-61.
[52] Wood AM, Joseph S. The absence of positive psychological (eudemonic) well-being as a risk factor for depression: a ten year cohort study. J Affect Disord 2010;122:213-7, http://dx.doi.org/10.1016/j.jad.2009.06.032.
[53] Fava GA. The clinical role of psychological well-being. World Psychiatry 2012;11:102-3.
[54] Guarda AS. Treatment of anorexia nervosa: insights and obstacles. Physiol Behav 2008;94:113-20.
[55] Fassino S, Pierò A, Tomba E, Abbate-Daga G. Factors associated with dropout from treatment for eating disorders: a comprehensive literature review. BMC Psychiatry 2009;9:1-9, http://dx.doi.org/10.1186/1471-244X-9-67.