Journal of Affective Disorders 350 (2024) 955–973

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

A systematic review and meta-analysis of the effectiveness of co-designed, in-person, mental health interventions for reducing anxiety and depression symptoms

Tamsin Greene Barker [a,b,*], Aoife O'Higgins [a,c], Peter Fonagy [b], Frances Gardner [a]

[a] Centre for Evidence-Based Intervention, Department of Social Policy and Intervention, University of Oxford, UK
[b] Research Department of Clinical, Educational and Health Psychology, University College London, UK
[c] Foundations What Works Centre for Children & Families, London, UK

ARTICLE INFO

Keywords:
Co-design
Mental health
Anxiety
Depression
Systematic review
Meta-analysis

ABSTRACT

Background: Co-design is recommended in mental health fields and has been associated with improved intervention efficacy. Despite its growing popularity, syntheses of evidence on the effectiveness of co-designed interventions are scarce, and little is known about their impact on anxiety and depression.
Methods: The purpose of this systematic review and meta-analysis was to consolidate evidence on the effectiveness of in-person, co-designed mental health interventions for reducing anxiety and depression symptoms. An exhaustive search was conducted across six electronic databases (PubMed, PsycINFO, Embase, CINAHL, CENTRAL, and ProQuest) and grey literature. Criteria for inclusion comprised studies utilizing randomized or quasi-randomized methods, implementing non-digital/in-person, co-designed interventions for mental health enhancement, and assessing anxiety and/or depression. Intervention impacts were evaluated using random-effects meta-analyses.
Results: The review identified 20 studies, with only three using the term 'co-design'. Other terminologies included 'co-developed' ($n = 2$), 'co-produced' (n = 2), and 'CBPR' ($n = 11$). Seventeen studies exhibited moderate risk of bias, while three demonstrated high risk. Meta-analyses demonstrated a moderate non-significant effect size of 0.5 (95 % CI: −0.8, 1.08; $p = 0.08$) on depression outcomes, and a small non-significant effect size of 0.12 (95 % CI: −0.1, 0.33; $p = 0.23$) on anxiety outcomes.
Limitations: The majority of studies lacked sufficient statistical power to detect between-group differences. Following GRADE criteria, confidence in estimates was low.
Conclusions: Notwithstanding widespread enthusiasm for co-design, the current evidence base is inadequate to confirm the impact of in-person, co-designed mental health interventions on anxiety and depression. More full-scale evaluation trials of higher quality are urgently needed, along with uniform terminology and measurement.

1. Background

Mental health issues, particularly anxiety and depression, represent a significant societal concern. One in six adults grapple with anxiety or depression each week (McManus et al., 2016); prevalence estimates are rising (WHO, 2017) and were especially elevated during the COVID-19 pandemic (Mahmud et al., 2023). As every individual has the right to the highest attainable standard of mental health (Asanbe et al., 2018), it is incumbent upon researchers, practitioners, and policymakers to devise effective strategies to enhance mental health, underpinned by contemporary, robust evidence.

Co-design, a distinctive approach to crafting interventions, is gaining traction in the field of mental health. Co-design is defined as a method of developing interventions in collaboration "with, not for, people" (McKercher, 2020, p.14), whereby stakeholders, such as individuals with lived experiences of the relevant issues or potential beneficiaries of the intervention, are actively involved in the intervention design process (Bevan-Jones et al., 2020; Burkett, 2016). A range of terms, such as co-production and co-development, are used to refer to such participatory approaches but despite terminological diversity, they largely share core attributes (Orygen, 2019).

Co-design's popularity in the realm of mental health is rising (Bevan-

* Corresponding author: Department of Social Policy and Intervention, University of Oxford, Barnett House, 32 Wellington Square, Oxford OX1 2ER, UK.
E-mail address: tamsin.greenebarker@spi.ox.ac.uk (T. Greene Barker).

https://doi.org/10.1016/j.jad.2023.12.080
Received 21 July 2023; Received in revised form 2 December 2023; Accepted 27 December 2023
Available online 9 January 2024

Jones et al., 2020; Marney and Elderton, 2021). Co-design is endorsed by a multitude of leading practitioners and researchers (e.g., Davies and Bergin, 2021; Halldorsson et al., 2021; NIHR, 2015; Oostermeijer et al., 2021; Pearce et al., 2021), and is earning recognition in policy domains (EIF, 2022; HSE, 2017; Roper et al., 2018; Tindall et al., 2021), with international calls for co-design within mental health services (CMHL, n.d.; Mental Health Commission of Canada, 2016; Palmer et al., 2021; Vanstone, 2021; WHO, 2013). Notably, co-design is recommended by the National Institute for Health and Care Excellence (2016) guidelines for community engagement in health and wellbeing interventions, and several leading mental health organizations, including the McPin Foundation, Mind, MQ and the Mental Health Foundation, advocate for co-design.

Co-design is said to create solutions with a keen understanding of local contexts, thereby ensuring that the outcomes of intervention design processes align with the needs of end-users (Jessup et al., 2018; Li et al., 2022a). It is believed to make interventions more relevant, engaging, and acceptable to end-users (Orlowski et al., 2015; Tindall et al., 2021). As a result, intervention uptake, adherence and sustainability may be improved by co-design (Bergin et al., 2020; Fleming et al., 2016; Halldorsson et al., 2021; Sanders and Stappers, 2008). Some suggest that the involvement of potential end-users in the co-design of interventions may enhance their effectiveness (Bevan-Jones et al., 2020; Burkett, 2016; Blomkamp, 2018; Centre for Coproduction, n.d.; Dekker and Williams, 2017; Eyles et al., 2016; Porche et al., 2022), and ultimately, lead to better outcomes (Craig et al., 2008; Evans et al., 2014; O'Brien et al., 2020; Williamson et al., 2021). However, there is a scarcity of explanations as to how co-design might boost intervention effectiveness.

The current state of evidence supporting co-design, particularly its impact on intervention effectiveness, and the efficacy of co-designed interventions for improving mental health, remains unclear. No studies were identified that directly compared co-design and non-co-design approaches in the development of mental health interventions, indicating a significant evidence gap. Given the uncertainty surrounding whether co-design improves intervention effectiveness, the following paragraphs delve into broader literature on co-design, with emphasis on reviews evaluating the effectiveness of co-designed interventions in enhancing mental health.

Bevan-Jones et al.'s (2020) exploratory review provides a useful summary of the principles, methods, and challenges of co-designing digital interventions. Similarly, Orlowski et al. (2015) examined the nature of consumer involvement in participatory processes for the design of technology-based interventions. Despite their intention to explore outcomes, lack of data prevented Orlowski et al. from determining the effect of participatory processes on intervention efficacy. Both reviews only considered co-design with youth and concluded that the benefits of participatory processes remain uncertain, necessitating further evaluation. Dekker and Williams (2017) searched two databases for games for reducing anxiety and depression which were developed using participatory processes. They identified three co-designed interventions but did not conduct meta-analyses due to included studies' designs. While reviews of digital interventions have helped to elucidate co-design processes, more evidence is needed to understand co-design's impacts.

Reviews focusing on non-digital, co-designed mental health interventions have also generally been limited in their examination of outcomes. Three such reviews were identified, with only one considering initiatives outside of healthcare. De Cotta et al. (2021) narratively synthesized the characteristics of 14 co-produced mental health initiatives (both within and outside of mental health services), limiting its scope to peer-reviewed studies conducted in rural Australia. Due to inconsistent evaluation methods and a lack of randomized or quasi-randomized trials, it was not possible to evaluate intervention effectiveness. Two systematic reviews looked at co-production and user involvement in child and adolescent mental health services (Norton, 2021; Viksveen et al., 2021). However, the low number of studies and nature of study designs (e.g., cross-sectional surveys) precluded meta-analyses.

This review seeks to build upon previous reviews by examining anxiety and depression outcomes and pooling quantitative results. Given some progress has been made in digital, co-design literature, while there is limited evidence synthesis in the realm of non-digital, co-designed interventions, our review focuses on in-person interventions. We also only include interventions implemented outside of formal healthcare given that health services represent a distinct type of intervention with their own literature base. For example, evidence for treatment of psychiatric diagnoses by healthcare professionals may substantially differ from evidence for community-based interventions aiming to prevent or improve (but not necessarily treat) anxiety and depression symptoms, with the former falling beyond the purview of this review.

To the authors' knowledge, the only review to pool data concerning the impact of in-person, co-designed interventions on mental health is Halvorsrud et al.'s (2021) review of reviews and meta-analysis, which examined the effectiveness of co-creation/co-production in health research. Their analysis suggested that the effect of co-creation was non-significant for mental health outcomes. However, mental health interventions comprised a small portion of their umbrella review. They pooled results of six mental health studies, which encompassed pre-post evaluations unsuitable for exploring cause-and-effect relationships. Moreover, they amalgamated a broad range of mental health outcomes in meta-analyses (PTSD, wellbeing and 'poor mental health') and excluded grey literature and interventions co-designed with youth. The present review, by contrast, focuses on randomized or quasi-randomized controlled trials which facilitate causal inferences. It includes grey literature and interventions co-designed with youth, and disaggregates intervention effects on anxiety and depression symptoms.

In summary, while co-design is strongly endorsed in mental health work and its popularity is on the rise, claims linking co-design to enhanced intervention effectiveness appear unsupported. Rigorous evaluations of the effectiveness of co-designed interventions in improving mental health outcomes are lacking. Narrowing the scope to non-healthcare interventions, this review found no systematic reviews of primary studies examining the impact of in-person, co-designed mental health interventions on mental health outcomes. Additionally, no meta-analyses have disaggregated effects of co-designed, in-person mental health interventions on anxiety and depression symptoms. Numerous calls have been made for a more explicit elucidation of the relationship between co-designed interventions and outcomes (Clarke et al., 2017; Eyles et al., 2016; Palmer et al., 2015, 2019; O'Brien et al., 2020).

This study has two aims, addressed through a systematic review and meta-analysis. Firstly, it aims to identify, appraise, and synthesize all studies assessing the impact of in-person, co-designed mental health interventions on anxiety and depression outcomes. Secondly, it seeks to quantitatively consolidate all evidence relating to the impact of in-person, co-designed mental health interventions on anxiety and depression outcomes. These outcomes are chosen because anxiety and depression are two of the most prevalent mental health conditions (WHO, 2017) and it is important to understand whether interventions are effective for improving anxiety and depression specifically.

2. Methods

Systematic review and meta-analysis methods were informed by the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al., 2022b) and reported in line with PRISMA guidelines (Page et al., 2022).

2.1. Eligibility criteria

Eligibility criteria were developed using the PICO framework

(McKenzie et al., 2022). Inclusion and exclusion criteria, and their rationale are detailed in Table 1. For a study to be included, all inclusion criteria had to be met.

2.2. Search strategy

Six electronic databases were searched from databases' inception to July 1st 2022: PubMed, PsycINFO, Embase, CINAHL, Cochrane Central Register of Controlled Trials (CENTRAL) and ProQuest.

Supplementary material displays the grey literature sources searched. Bibliographies of all included studies were searched for records not yet identified. Titles and abstracts of studies that cited included studies and reference lists of previous relevant reviews were screened.

2.3. Search strings

Search strings included terms pertaining to three areas: 1) co-design, 2) mental health, and 3) study design. Searches were designed to be sensitive and comprehensive as this is the first known review on this topic and preliminary scoping searches identified few relevant studies. Synonyms for each component were maximized, while the number of areas considered were minimized (Petticrew and Roberts, 2006).

2.4. Study selection

Titles and abstracts of identified citations were imported into reference manager software Covidence and duplicates removed. Titles and abstracts were then screened for relevance to research questions. Studies that met inclusion criteria and records whereby the abstract and title contained insufficient information to determine eligibility proceeded to full-text assessment.

Studies were excluded if authors did not explicitly mention that interventions were designed or implemented to improve mental health, anxiety, or depression. Studies were also excluded if they concerned interventions implemented as part of hospital, GP, or primary care, as these were beyond this review's scope. Whenever searches produced protocols, these were followed-up to determine whether planned studies had since been published. One reviewer performed screening and selection, with a second reviewer available to discuss difficult decisions.

2.5. Data extraction

Cochrane's (2017) EPOC data extraction form guided decisions on what data to retrieve.

2.6. Risk of bias (ROB) assessment

Included studies were assessed for ROB using the Cochrane ROB2 risk-of-bias tool for randomized trials (Sterne et al., 2019). As no quasi-randomized studies were included in the final list of studies, no other ROB tool was required.

2.7. Narrative synthesis

Following McKenzie and Brennan's (2022) guidelines, all studies included in the systematic review were synthesized narratively. This facilitated a rich description of relevant study information that was not captured in meta-analyses.

2.8. Meta-analyses

Statistical analyses were performed using R version 4.2.1 (R Core Team, 2022) and R Studio version 2022.07.1 + 554 (R Studio Team, 2022).

Table 1
Eligibility criteria.

	Inclusion criteria	Exclusion criteria	Rationale
Population	Children or adults of any age, including those with and without psychiatric diagnosis.	No population exclusions.	Both children and adults are well-represented in co-design literature.
Intervention	Any co-designed, in-person (i.e., non-digital) intervention designed, or implemented, to improve mental health, anxiety or depression. Co-designed interventions are defined as interventions whereby individuals who have lived experience of the issue(s) that the intervention is intended to address, and/or are potential intervention beneficiaries, were involved in the intervention's design process. The term 'co-design' did not have to be explicitly reported for studies to be included. More than one 'co-designer' was required for a study to be included. There was no minimum threshold for the degree of co-design.	Interventions not developed using co-design methods. Interventions co-designed with one person only e.g., one researcher had lived experience of depression. Studies in which authors did not explicitly report that the intervention was designed to improve mental health, anxiety, or depression. Digital/e-health interventions. Interventions combined with other interventions. Quality improvement interventions in hospital, GP, or primary care settings.	Many studies of co-designed interventions were not directly related to mental health yet measured anxiety and depression. These studies were excluded because their primary focus was not mental health.
Comparator	Studies comparing the effect of an intervention to a comparison condition including treatment as usual, wait-list control, inactive control interventions, or no intervention.	Studies with no comparison group.	Including a control group facilitates inferences that between-group differences were due to the intervention rather than other factors e.g., regression to the mean (Ruble, 2017).
Outcome	Studies that measured either depression and/or anxiety outcomes.	Studies that did not measure either depression or anxiety. Studies that only included general measures of mental health or psychiatric recovery.	Studies reporting only general measures of mental health were excluded as this review was exclusively interested in anxiety and depression outcomes. Studies that measured, but did not report, anxiety or depression outcomes were included to

(continued on next page)

Table 1 (*continued*)

	Inclusion criteria	Exclusion criteria	Rationale
			minimize bias from selective non-reporting.
Study design	Any randomized controlled trial (RCT) studies whereby participants were randomly assigned to either a treatment or comparison group. Quasi-experimental studies, defined as study designs that estimate causal effect sizes using exogenous variation in the exposure (or intervention) that is not directly controlled by the researcher (Rockers et al., 2015, p. 511). The following quasi-experimental study designs were eligible: natural experiments, instrumental variable analyses, regression discontinuity analyses, and difference-in-difference analyses due to their strength for causal inferences (Kim and Steiner, 2016).	Pre-post studies, qualitative studies, observational studies, case studies, secondary analyses, study protocols, commentaries, opinion papers, and reviews. Controlled studies where there was no randomization or quasi-randomization process for allocating participants to groups (e.g., matched subjects designs). Any quasi-experimental study design not specified in the inclusion criteria.	RCTs are deemed the 'gold standard' of evidence for examining cause-and-effect relationships between interventions and outcomes (Hariton and Locascio, 2018). Quasi-experimental studies were included as they are valuable for drawing causal inferences in situations where randomization is not feasible (Kim and Steiner, 2016).
Language	English language only.	Studies published in other languages.	
Location	Any geographical location.	No geographical restrictions.	
Years	Any year.	No restrictions for dates of publication.	
Publication type	Any trial, including published and unpublished studies (e.g., dissertations).		No restrictions on publication type to minimize risk of publication bias.

2.8.1. Criteria for inclusion in meta-analyses

Studies were included in meta-analyses if, for anxiety or depression outcomes, the following information was reported:

(1) mean, standard deviation, and sample size, measured pre- and post-intervention, for both the intervention and control group;

or

(2) change in mean value from baseline to post-intervention for both the intervention and control group, associated standard error for this change score, post-intervention sample sizes, and correlation coefficient r between pre- and post-intervention values.

2.8.2. Estimating Hedges' g effects

As studies employed different measurement instruments, a standardized mean difference (SMD) was calculated for each study (Andrade, 2020). Inputting means, standard deviations, and sample sizes into R package 'esc' (Lüdecke, 2019), Hedge's g effects and respective standard errors were calculated. Hedges' g was chosen over Cohen's d because it corrects for bias due to small sample sizes (Hedges and Olkin, 1985), which was applicable to many included studies. When post-intervention outcomes were reported at multiple time points, data from the first measurement post-intervention were used. Effect sizes were classified as small (0.2), moderate (0.5), or large (over 0.8), following Cohen (1988).

2.8.3. Performing meta-analyses

Using R package 'meta' (Schwarzer, 2022), random-effects models were used to pool individual studies' Hedges' g effects. Random-effects were chosen because interventions varied substantially; thus, between-study heterogeneity in true effects was expected. Effect estimates were combined using the inverse variance method (Deeks et al., 2022). Separate meta-analyses were conducted for depression and anxiety outcomes.

2.8.4. Heterogeneity assessment

Statistical heterogeneity was evaluated using Chi^2 tests and I^2 statistics. A p-value < 0.10 was chosen to establish statistical significance given the low number of studies and small sample sizes (Deeks et al., 2022).

2.9. Publication bias

Contour-enhanced funnel plots were employed to assess publication bias (Peters et al., 2008) but were only drawn if meta-analyses included over nine studies (Page et al., 2022).

2.10. Grade assessment

The GRADE framework was employed to evaluate the quality of evidence in each meta-analysis, i.e., the extent to which we can be confident that effect estimates correspond to true effects (Guyatt et al., 2011; Schünemann et al., 2022).

3. Results

3.1. Search results

Searches yielded 3820 results, with 2303 remaining after de-duplication, and 290 proceeding to full text review. Of these, 20 were included in the systematic review and 11 met criteria for meta-analysis. Fig. 1 shows the flow chart.

3.2. Characteristics of included studies

19 peer-reviewed journal articles and one PhD dissertation (Hedemann, 2019) met inclusion criteria. Studies were published between 2013 and 2021, with 70 % published since 2018. Table 2 summarizes study characteristics.

3.2.1. Settings

Studies were located in 11 countries, including Australia, Bangladesh, Canada, Gambia, Germany, Guatemala, Kenya, South Africa, Thailand, the United Kingdom ($n = 5$) and the United States ($n = 6$).

3.2.2. Participants

Across 20 studies, there were 2095 participants in total. Sample sizes ranged from $n = 33$ to $n = 448$ (M = 113.05; SD = 97.27). Participant ages ranged from 5 to 91 years. 11 studies considered adults only, five considered children only (including one which considered youth up to age 20), and four included both child and adult participants. A high proportion of participants were female, with eight studies including only females. Participants had a variety of ethnicities. Two studies' participants were predominantly Black, two predominantly Mexican, one

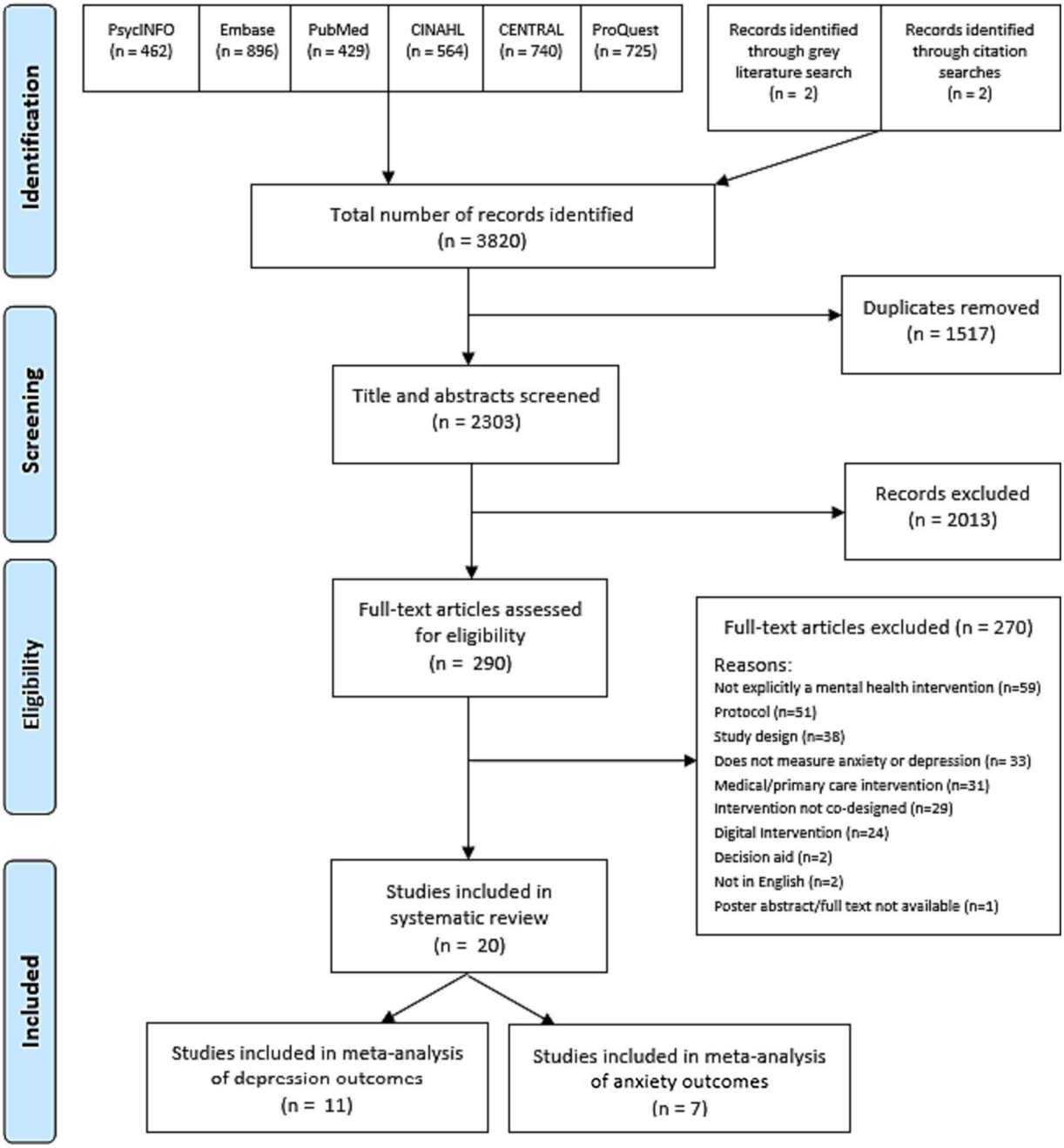

Fig. 1. Prisma flow chart.

predominantly Indigenous Peoples and six predominantly White. Eight studies did not report ethnicities.

3.2.3. *Mental health diagnoses and target populations*

Mental health diagnoses and clinical symptoms were mentioned in various forms in seven of the studies. Two implied that a subset of participants had psychiatric diagnoses; Pile et al. (2021) specified that individuals with mental health diagnoses were not excluded, while Henson et al. (2020) reported that 75 % of participants had never received a psychiatric diagnosis. To be included in Abel et al.'s (2020) study, families were required to have 'serious parental mental illness', defined as a severe psychiatric disorder requiring intervention, hospitalization or ongoing treatment. Other studies referred to participants having a history of clinical depression and/or current minor depression (Alsubaie et al., 2018) and a diagnosis of opioid use disorder (Carlyle et al., 2019). The interventions in the studies by Hassouneh et al. (2013) and Sanfilippo et al. (2020) were designed to reduce clinically significant depressive symptoms and reduce common mental disorder symptoms respectively. However, diagnostic criteria were absent from all studies.

There was a range of populations with respect to the severity of mental health problems. Five studies (Alsubaie et al., 2018; Hassouneh et al., 2013; Karasz et al., 2021, 2015; Rüsch et al., 2019) recruited participants on the basis of elevated psychological symptoms, e.g., Patient Health Questionnaire score ≥ 10.

Ten studies did not screen for symptoms but targeted specific populations at high risk for mental health problems, including perinatally HIV-infected adolescents, breast cancer survivors, refugee families, and young people in care self-reporting substance use (Abel et al., 2020; Alderson et al., 2020; Betancourt et al., 2020; Bhana et al., 2014; Carlyle

Table 2
Characteristics of included studies.

Study	Country	Design	Intervention(s)	Intervention duration	Sample sizes at baseline	Retention rates (from baseline to first post-intervention measurement)	Age of participants: mean (SD)	Sex of participants	Ethnicity of participants	Data collection time points
(Abel et al., 2020)	England	Feasibility RCT	Young Simplifying Mental Illness plus Life Enhancement Skills (Young SMILES)	8 sessions for children and 5 sessions for parents/caregivers delivered over 8 weeks	Intervention: $n = 20$ families/$n = 24$ children Control: $n = 15$ families/$n = 16$ children Total: $n = 35$ families/$n = 40$ children	84.48 %	Intervention children: 10.2 (2.3) Control children: 11.1 (2.7)	Intervention children: 50 % female; 50 % male Control children: 56 % female; 44 % male Intervention parents: 94 % female; 6 % male Control parents: 87 % female; 13 % male	Intervention parents: 89 % White British; 11 % Asian Control parents: 93 % White British; 7 % missing data	Baseline, 4 months and 6 months post-baseline
(Alderson et al., 2020)	England	Three-arm feasibility RCT	Motivational Enhancement Therapy (MET); Social Behavior and Network Therapy (SBNT) (both co-designed)	6 1-hour sessions offered weekly or fortnightly over 6–12 weeks	MET: $n = 38$ SBNT: n = 38 Usual care: n = 38 Total: $n = 114$	53.57 %	17.3 (2)	55 % female; 45 % male	95 % White British; 4 % other; 1 % missing data	Baseline and 12 months post-recruitment
(Alsubaie et al., 2018)	England	Three-arm RCT	MBCT-HELM (adapted version of mindfulness-based cognitive therapy)	Weekly 2.5-hour sessions delivered over 8 weeks	MBCT-HELM: n = 11 Alternative (non-co-designed) intervention: n = 11 Treatment as usual: n = 11 Total: $n = 33$	84.85 %	64.8 (10)	42 % female; 58 % male	94 % White British/ Irish; 3 % White other; 3 % Asian British	Baseline and post-intervention
(Ashton et al., 2017)	Australia	Two-arm pilot RCT	Harnessing Ehealth to enhance Young men's Mental health, Activity and Nutrition (HEYMAN)	1-hour weekly sessions (11 group-based and 1 individual) over 3 months	Intervention: $n = 26$ Control: n = 24 Total: $n = 50$	94 %	22.1 (2)	100 % male	NR	Baseline and 3 month follow-up
(Betancourt et al., 2020)	United States	Two-arm pilot feasibility and acceptability trial	Family Strengthening Intervention for Refugees (FSI-R)	10 90-minute sessions delivered over 10 weeks	Intervention: $n = 39$ families / $n = 130$ individuals Control: $n = 41$ families / $n = 132$ individuals Total:	67.18 %	Bhutenese children: 14.35 Somali Bhantu children: 14.6 Bhutenese caregivers:	Bhutenese children: 53.1 % female Somali Bhantu children: 59.2 % female Bhutenese caregivers: 52	NR	Baseline and post-intervention

(*continued on next page*)

Table 2 (*continued*)

Study	Country	Design	Intervention(s)	Intervention duration	Sample sizes at baseline	Retention rates (from baseline to first post-intervention measurement)	Age of participants: mean (SD)	Sex of participants	Ethnicity of participants	Data collection time points
					$n = 80$ families/$n = 262$ individuals		40.97 Somali Bhantu caregivers: 41.77 a b	% female Somali Bhantu caregivers: 79 % female b		
(Bhana et al., 2014)	South Africa	Two-arm pilot RCT	VUKA	6 sessions over 3 months	Intervention: n = 33 Wait-list control: $n = 32$ Total: $n = 65$	NR	NR	51 % female; 49 % male	All participants were Black South Africans of Zulu ethnicity	Baseline, post-intervention and 2 week follow-up
(Carlyle et al., 2019)	England	Mixed-methods design (three-arm)	Compassion focused therapy (CFT)	3 2-hour sessions over 3 weeks	CFT: n = 15 Non-co-designed intervention: $n = 12$ Control: n = 11 Total: n = 38	80.85 %	39.95 (10.44)	37 % female; 63 % male	NR	Baseline and post-intervention
(Chomat et al., 2019)	Guatemala	Parallel group pilot randomized study (two arm)	Women's Circles	2 sessions per week over 5 weeks	Intervention: $n = 81$ Control: $n = 71$ Total: $n = 152$	78.06 %	26.2 (6.4)	100 % female	Intervention: 66.2 % *Mam*; 21.1 % *K'iche'*; 12.7 % not indigenous Control: 78.2 % *Mam*; 14.5 % *K'iche'*; 7.3 % not indigenous	Baseline and post-intervention
(Hassouneh et al., 2013)	United States	Efficacy trial (two arm)	Healing Pathways	14 2.5-hour sessions over 14 weeks	Intervention: $n = 44$ Waitlist control: $n = 36$ Total: n = 80	81 %	51.81 (10.44)	100 % female	75.95 % White; 11.39 % Multiracial; 2.53 % African American/ Black; 2.53 % Native American; 2.53 % Hispanic; 1.27 % Asian; 3.8 % unknown	Baseline, post-intervention, 6 weeks and 3 month follow-up
(Hedemann, 2019)	United States	Cluster-RCT (two arm)	Emotion Regulation Skills Intervention (ERSI)	Multiple sessions over 7–12 weeks	Intervention: $n = 70$ Control: $n = 60$ Total: n = 130	NR	7.55 (1.66)	50 % female; 50 % male	61 % African-American, 22 % Hispanic, 7 % non-Hispanic White, 3 % Asian/Pacific Islander, 7 % other	Baseline and end of school year
(Henson et al., 2020)	Canada	2 × 2 crossover design	Self-Compassion Intervention (SCI)	One 75 min session per week for 8 weeks	Intervention: $n = 16$ (group A) Control: $n = 17$ (group B) Total: $n = 33$	87.88 %	Group A: 14.25 (0.45) Group B: 14.06 (0.24)	100 % female	79 % White; 21 % Other	Baseline and post-intervention
(Karasz et al., 2015)	United States	Preliminary pilot evaluation (two arm)	Action to Improve Self-esteem and Health through Asset building (ASHA)	2 sessions per month over 26 weeks	Intervention: n = 44 Control: $n = 22$ Total: $n = 66$	80 %	Intervention: 40.4 (6.41) Control: 42.1 (2.49)	100 % female	NR	Baseline and 6 months follow-up

(*continued on next page*)

Table 2 (*continued*)

Study	Country	Design	Intervention(s)	Intervention duration	Sample sizes at baseline	Retention rates (from baseline to first post-intervention measurement)	Age of participants: mean (SD)	Sex of participants	Ethnicity of participants	Data collection time points
(Karasz et al., 2021)	Bangladesh	Pilot RCT (two arm)	Action to Improve Self-esteem and Health through Asset building (ASHA)	2-hour sessions every other week for six months	Intervention: $n = 24$ Control: n = 24 Total: $n = 48$	NR	26.1 (4.6)	100 % female	NR	Baseline and 12 months post-baseline
(Nápoles et al., 2015)	United States	RCT (two arm)	Nuevo Amanecer, a cognitive-behavioral stress management program	90-minute session weekly over 8 weeks	Intervention: $n = 76$ Control: $n = 75$ Total: $n = 151$	95.36 %	50.5 (10.9)	100 % female	68 % Mexican; 23 % Central American; 9 % South American	Baseline, 3 month and 6 month follow up
(Nápoles et al., 2020)	United States	RCT (two arm)	Nuevo Amanecer (NA-II), a stress management program for rural, low literacy Latina breast cancer survivors	90-minute session weekly over 10 weeks	Intervention: n = 76 Control: $n = 77$ Total: $n = 153$	89.54 %	54.8 (10.5)	100 % female	97 % Mexican; 1 % Central American; 2 % other	Baseline, 3 month and 6 month follow up
(Nestadt et al., 2019)	Thailand	Pilot RCT (two arm)	CHAMP+, a multiple family group-based intervention to promote mental health and prevent sexual and drug use risk behavior	one weekend session per month over 6 months	Intervention: $n = 45$ child-caregiver dyads Control: $n = 43$ child-caregiver dyads Total: $n = 88$ child-caregiver dyads / 176 individuals	100 %	Children: 12.28 (1.41) Caregivers: 48.2 (12.5)	Children: 49 % female; 51 % male Caregivers: 89 % female; 11 % male	NR	Baseline, post-intervention, and 3 month follow-up
(Pile et al., 2021)	England	Feasibility RCT (two arm)	My Memory Forest	Families were asked to read and discuss the story and complete 6 exercises over 3 weeks	Intervention: $n = 29$ Active control (alternative, non-mental health intervention): $n = 27$ Total: $n = 56$	100 %	Intervention: 7.73 (1.22) Control: 7.79 (1.21)	63 % female; 27 % male	45 % White/Caucasian; 55 % NR	Baseline and post-intervention
(Puffer et al., 2016)	Kenya	Stepped-wedge cluster-RCT	READY	NR	448 individuals (4 churches) invited to receive the intervention, with order of roll-out randomized across churches	NR	Caregivers: 39.3 (SD NR) Youth: 12.3 (2)	Caregivers: 60 % female; 40 % male Youth: 51.9 % female; male NR	NR	Baseline, 1 month and 3 months post-intervention
(Rüsch et al., 2019)	Germany	Pilot parallel two-arm RCT	Peer-led group program that supports unemployed people with mental health problems in terms of help-seeking, job search and recovery	4 2-hour sessions over 6 weeks	Intervention: $n = 23$ Treatment as usual: $n = 19$ Total: $n = 42$	90.48 %	46.1 (9.9)	52 % female; 48 % male	100 % Caucasian	Baseline, 3 weeks, 6 weeks, and 6 months post-baseline
(Sanfilippo et al., 2020)	Gambia	Stepped-wedge cluster-RCT	Community Health Intervention through Musical Engagement (CHIME)	60-minute session once per week for 6 weeks	Intervention: $n = 50$ Control: $n = 74$ Total: $n = 124$	80 %	26.95 (5.72)	100 % female	NR	Baseline and 4 weeks post-intervention

et al., 2019; Chomat et al., 2019; Nápoles et al., 2015, 2020; Nestadt et al., 2019; Sanfilippo et al., 2020).

The remaining five studies targeted specific populations (e.g., children, men aged 18–25 years, individuals assigned female at birth, and churchgoers in rural Kenya) and emphasized the importance of taking action to prevent mental health problems in these populations, but participants were not necessarily described as at elevated risk for anxiety or depression (Ashton et al., 2017; Hedemann, 2019; Henson et al., 2020; Pile et al., 2021; Puffer et al., 2016).

3.2.4. Study designs

All studies involved randomization to an intervention or comparison group, but reported study designs varied, as evident in Table 2. 14 were explicitly RCTs while others were described as two- or three-arm trials. 11 employed the term 'pilot' or 'feasibility' to describe their design. None compared co-designed with non-co-designed interventions.

3.2.5. Comparison groups

All studies contained a comparison condition whereby participants did not receive the intervention. 11 studies had a wait-list control group. In seven studies, participants in the comparison condition did not receive the intervention but had access to their usual services.

3.2.6. Interventions

The scope of this review encompasses 20 unique co-designed interventions, which featured in the 20 studies analyzed. Of these, two studies pertained to the same intervention (ASHA), while one study evaluated two separate co-designed interventions.

The specific objectives and approaches of interventions were markedly heterogeneous. Eight interventions were centered exclusively on ameliorating mental health outcomes. The strategies employed to achieve this ranged from facilitating reflection and sharing of personal experiences in group settings, implementing mindfulness-based cognitive therapy, utilizing music, promoting self-compassion, developing emotion regulation skills and delivering autobiographical memory interventions. In contrast, the remaining twelve interventions had a broader focus, aiming to enhance mental health in conjunction with other objectives. These objectives spanned a range of domains, including but not limited to, reducing HIV risk and substance use, mitigating unemployment, cultivating communication and stress management skills, fostering better family relationships, promoting economic empowerment, and enhancing self-esteem and overall quality of life. The techniques used to achieve these multifaceted objectives included, among others, educational sessions, guided group discussions, physical exercise activities and home-visiting programs.

The duration of interventions varied widely, from as short as three weeks to as long as 26 weeks. The personnel facilitating these interventions also differed across studies. Four interventions were administered by professional psychological practitioners, two by lay counsellors, and three by researchers. Nine interventions were facilitated by specially trained individuals such as peer facilitators, interventionists or community leaders. One intervention by Pile et al. (2021) was administered by the parents or caregivers of the child participants, while the intervention in the study by Rüsch et al. (2019) was co-facilitated by a clinical psychologist in tandem with an individual who had personal experience with mental health issues.

The review's eligibility criteria excluded interventions that, for the sake of the study, were explicitly combined with other interventions or healthcare treatments (e.g., psychotropic drugs). Accordingly, none of the 20 included studies provided details on other treatments being carried out alongside the co-designed interventions. Participants in some studies may have been receiving their usual healthcare treatments (e.g., mental health medication), but with the exception of Carlyle et al. (2019), information on usual treatment was not reported.

NR = not reported.
a = standard deviations not reported.
b = intervention and control group numbers not reported.

3.2.7. Co-design terminology and processes

While only three studies explicitly referenced the term 'co-design', all encompassed some form of participatory involvement with potential beneficiaries during the development phase of their interventions. Related terminology such as 'co-develop' ($n = 2$), 'co-produce' (n = 2), 'co-create' ($n = 1$), and 'collaboratively designed' (n = 2) were also noted. However, none of these studies provided a formal definition or theoretical framework for 'co-design' or its equivalent terminology. In 11 studies, the approach was characterized under the umbrella of community-based participatory research (CBPR). A total of 13 studies detailed the active involvement of potential beneficiaries in the design or development of the interventions. Conversely, seven studies reported the adaptation of pre-existing interventions to meet the specific needs of a particular group or context through participatory processes.

As for the methodologies adopted in co-design processes, a diverse array was reported including focus groups, workshops, community consultations, and qualitative interviews. The specific nature of co-design activities was explicitly stated in only eight studies (40 %).

The degree of co-design implementation varied considerably among the studies. The extent to which the interventions were co-designed was clearly evident in eight studies. For instance, it was elucidated whether potential beneficiaries merely functioned as informants after the preliminary design of the intervention, assuming roles subordinate to researchers or practitioners, or whether they were engaged in the design process from the onset, playing a pivotal role in shaping the intervention. The level of co-design remained vague in twelve studies. Nevertheless, ten studies (six of the aforementioned twelve) cited other publications by the same research group that offered supplementary details regarding co-design processes.

A total of seven studies provided justification for their choice to use co-design. Motivations ranged from aspirations to 'enhance program effectiveness' (Ashton et al., 2017, p.3), to improve 'cultural safety, acceptability and feasibility' (Chomat et al., 2019, p.2), and to ensure the interventions were 'contextually appropriate' (Sanfilippo et al., 2020, p.3).

3.3. Anxiety and depression outcomes

Depression outcomes were reported in 17 studies using 11 distinct measures, including the Patient Health Questionnaire-8 (PHQ-8; Kroenke et al., 2009), Patient Health Questionnaire-9 (PHQ-9; Spitzer et al., 1999), Center for Epidemiologic Studies-Depression Scale (CES-D; Radloff, 1977), Center for Epidemiologic Studies Depression Scale for Children (CES-DC; Faulstich et al., 1986), Children's Depression Inventory (CDI; Kovacs, 1992), Edinburgh Postnatal Depression Scale (EPDS; Cox et al., 2014), Brief Symptom Inventory (BSI; Derogatis and Melisaratos, 1983), Revised Child Anxiety and Depression Scale (RCADS; Chorpita et al., 2000), Depression Anxiety Stress Scale (DASS; Lovibond and Lovibond, 1995), Beck Youth Inventory (BYI-II; Beck et al., 2005), and the Hopkins Symptom Checklist-25 (HSCL-25; Derogatis et al., 1974).

Anxiety outcomes were presented in 11 studies utilizing nine different measures: Generalized Anxiety Disorder-7 (GAD-7; Spitzer et al., 2006), Multi-Dimensional Anxiety Scale for Children (MASC; March et al., 1997), Spence Children's Anxiety Scale (SCAS; Spence, 1998), BYI-II, BSI, RCADS, CES-DC, DASS, and the HSCL-25.

Combined anxiety and depression outcomes were reported in seven studies, employing five measures: the Depression Anxiety Stress Scale-21 (DASS-21; Antony et al., 1998), EQ-5D-5L (Janssen et al., 2013), SRQ-20 (Beusenberg and Orley, 1994), HSCL-25, and RCADS.

All of the aforementioned outcome measurements were self-reported by study participants. Data were collected immediately post-intervention in nine studies. The remaining studies gathered data at various time points after baseline, ranging from three weeks post-intervention (Rüsch et al., 2019) to 12 months post-recruitment (Alderson et al., 2020). Data collection at multiple time points post-intervention was conducted in eight studies.

3.4. Statistical analyses and results of individual studies

The reported statistical methods and effects varied among the studies, as detailed in the supplementary material. The findings across these studies were heterogeneous. Three studies did not perform any formal statistical analyses. Eight studies reported statistically non-significant results, suggesting no discernable difference in outcomes between the intervention and control conditions. Nine studies reported statistically significant results ($p < 0.05$), indicating that co-designed interventions led to improvements in anxiety or depression outcomes relative to control conditions. In twelve studies, the authors explicitly stated that their analyses did not possess sufficient statistical power to detect differences between groups.

3.5. Risk of bias

Figs. 2 and 3 summarize studies' risk of bias (ROB). 17 studies were judged to have moderate, and three high ROB. High ROB stemmed from deviations from the intended intervention due to the trial context (Hedemann, 2019), group differences in missing data (Karasz et al., 2015), and performing analyses not specified in the protocol (Sanfilippo et al., 2020).

3.6. Meta-analysis results

3.6.1. Depression

There was no effect of the co-designed interventions on depression. A pooled standardized mean difference of 0.5 (95 % CI:-0.08, 1.08; $p = 0.08$) indicated a medium positive effect that was non-significant. An I^2 statistic of 56 % suggested moderate heterogeneity ($Tau^2 = 0.37$, $Chi^2 = 22.56$, df = 10, $p = 0.01$). Fig. 4 displays a forest plot for depression outcomes.

3.6.2. Anxiety

There was no effect of the co-designed interventions on anxiety. A pooled standardized mean difference of 0.12 (95 % CI:-0.1, 0.33; $p = 0.23$) indicated a small effect that was non-significant. An I^2 statistic of 0 % implied no heterogeneity ($Tau^2 = 0.00$, $Chi^2 = 1.5$, df = 6, $p = 0.96$). However, this should be interpreted with caution due to the low number of trials (von Hippel, 2015). Fig. 5 displays a forest plot for anxiety outcomes.

3.7. Sensitivity analyses

Sensitivity analyses indicated the choice of time point for post-intervention data and excluding studies with high risk of bias did not markedly affect results (see supplementary material).

3.8. Publication bias

Fig. 6 presents a funnel plot evaluating potential publication bias in the meta-analysis of depression outcomes. The majority of effect estimates were of similar magnitude and centered around zero, but three estimates skewed the pooled-effect estimate upward, without corresponding effects on the other side of the funnel to establish balance. Underrepresentation of negative effects, particularly in the statistically significant shaded regions where results would be unfavorable to interventions, indicates potential publication bias (Peters et al., 2008).

Publication bias in which researchers primarily report statistically significant findings seems to be largely absent. The majority of effects were situated within the white pyramid, suggesting they were statistically non-significant ($p > 0.1$). However, other forms of publication bias that were not captured by the plot might also be present.

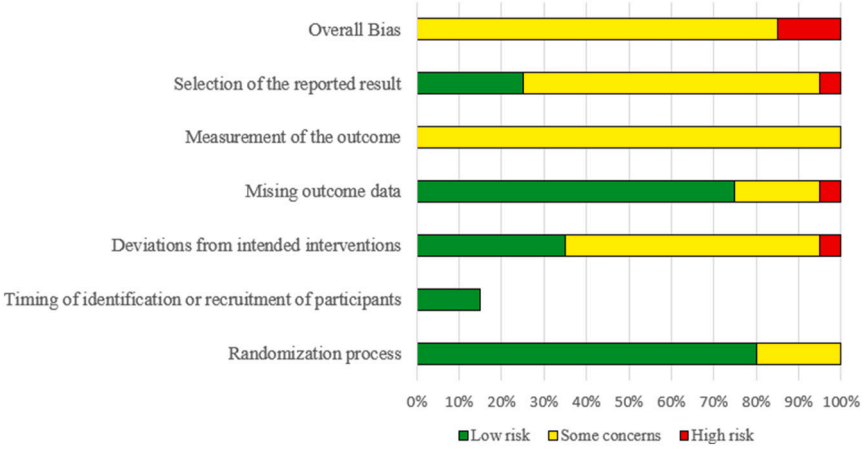

Fig. 2. Overall risk of bias for studies included in systematic review.

4. Discussion

4.1. Systematic review

This review provides a comprehensive summary of current literature assessing the impact of in-person, co-designed mental health interventions on anxiety and depression symptoms. A total of 20 studies met the inclusion criteria. Primary findings are as follows:

Firstly, interest in co-designing mental health interventions is burgeoning. A significant proportion of relevant studies were published within the past five years, and our search unveiled numerous protocols or studies slated for future publication. The array of interventions investigated in the included studies demonstrates that the appeal of co-design is not confined to a particular type of intervention.

The review also reveals a lack of standard terminology in the literature concerning co-designed mental health interventions. Moreover, studies tend to lack detailed information on the degree of co-design in the interventions, the stages at which co-design occurred during the development of interventions, and theory or empirical evidence that informed co-design approaches. Our findings suggest that 'co-design' and its related terms encompass manifold approaches and processes, including focus groups, workshops, community consultations and qualitative interviews. However, beyond naming such activities, information about co-design processes is often scant, relegated to other publications, or entirely missing.

Instruments used to measure anxiety and depression outcomes vary significantly across the co-design literature, highlighting a need for more consistency in measurement methods. The results were mixed, creating an ambiguous picture of the interventions' impact on anxiety and depression. Importantly, more than half of the included studies were small pilot or feasibility trials which, individually, were insufficiently powered to detect between-group differences.

4.2. Meta-analyses

The results indicate that there is currently no evidence to suggest that in-person, co-designed mental health interventions have an impact on anxiety and depression outcomes. In accordance with the GRADE criteria, confidence in the effect estimates is low. Considering the risk of bias, wide confidence intervals for effect estimates, and the limited number of studies that reported anxiety and depression measures in a manner that could be converted into standardized mean differences, these meta-analyses cannot conclusively determine whether in-person, co-designed mental health interventions have any effect on anxiety and depression. More high-quality studies are necessary before definitive conclusions can be made. Additionally, given the diversity of interventions included in this review, exploring heterogeneity of effects among different types of co-designed mental health interventions would be valuable.

The magnitude of change in anxiety and depression symptoms is certainly not the only important outcome measure for assessing the impact of co-design. The inherent value of participation in co-design could improve wellbeing (Renedo Illarregi, 2022), potentially justifying the approach. It's possible that co-design results in interventions that are more acceptable to end-users (Tindall et al., 2021), even if they are not more effective than non-co-designed interventions or treatments. Greater acceptability of co-designed interventions could lead to improved participant engagement, reduced dropout rates, and consequently better long-term outcomes. It should be noted that none of the studies included in this review measured outcomes beyond 12 months post-recruitment. If longer-term outcome data were available, a different pattern of results may have emerged.

Therefore, future research should not only explore the direct impact of co-designed interventions on symptom severity, but also consider parameters such as acceptability, engagement, dropout rates, sustainability, and long-term outcomes. Similarly, further investigation into potential psychological and wellbeing benefits derived from the co-design process itself could offer additional insights into its overall effectiveness and value. This comprehensive approach to evaluation would more accurately reflect the multifaceted nature of co-design and its potential impacts.

4.3. Overall completeness and applicability of evidence

A strength of this review is its comprehensive search strategy. Nonetheless, searches may have missed relevant evidence where studies reported co-design in methods, but not in their title, keywords, or abstract. This review excluded general health or wellbeing interventions and any study whereby authors did not report that the intervention was designed, or implemented, to improve mental health. In some cases, differences between mental health and wellbeing interventions were negligible. These exclusions may have meant that relevant evidence was missed.

Internal validity of evidence was relatively strong given all studies were randomized trials. Studies included participants of varying ages and ethnicities, which aids the external validity of the review's findings. Similarly, interventions were implemented across 11 countries and 5 continents, including high, middle, and low-income regions, which strengthens the generalizability of findings (Cook and Campbell, 1976). That said, for reasons outlined below, results should be interpreted with caution.

Fig. 3. Risk of bias assessments for individual studies.

4.4. Certainty of evidence

Table 3 summarizes GRADE evaluations. For both meta-analyses, evidence was rated as low quality, suggesting limited confidence in estimates and that true effects may differ considerably from estimated effects.

4.5. Limitations

Following Cochrane and PRISMA guidelines, robust methods were employed to identify, appraise, and synthesize studies. Nonetheless, several limitations of this review should be noted.

Firstly, no protocol was published. Searches were restricted to studies reported in English, which may have engendered language bias (Lefebvre et al., 2022). 60 % of included studies lacked statistical power to detect effects, hampering their contribution to the review (Turner et al., 2013). High attrition was also common, with only five studies having retention rates over 90 %. Attrition can engender bias in study results, e.g., overestimation of intervention effects, which subsequently biases review findings (Deke et al., 2015). Data for all outcomes considered came from self-report measures, which can be unreliable or impacted by social desirability bias (Althubaiti, 2016).

45 % of studies did not report the necessary statistics to calculate effects and therefore were omitted from meta-analyses. Although random-effects models incorporated the assumption that different trials estimated distinct, yet related, intervention effects (DerSimonian and Laird, 1986; Borenstein et al., 2010), the presence of heterogeneity widened confidence intervals around pooled-effect estimates and

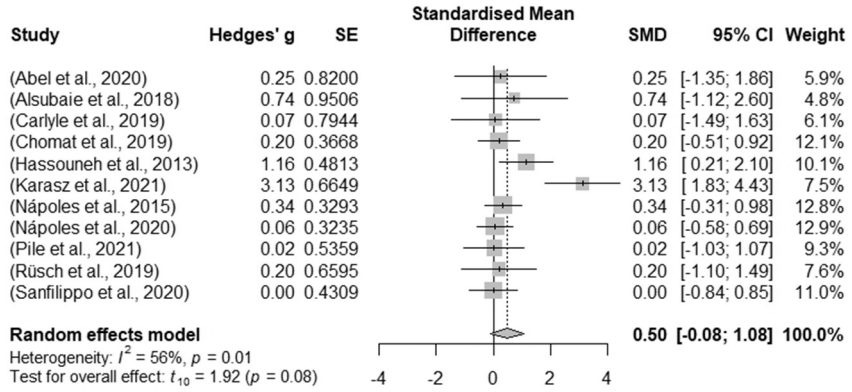

Fig. 4. Forest plot of intervention effects on depression outcomes.

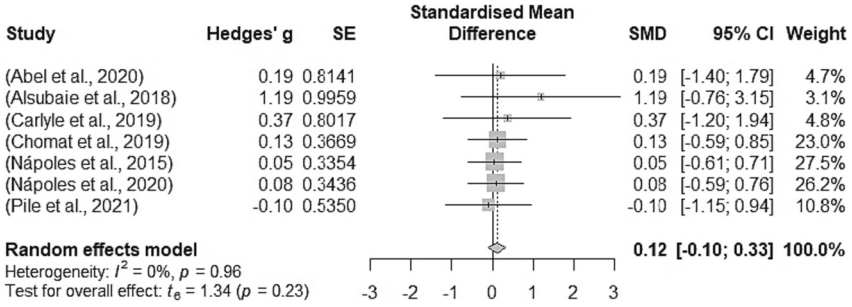

Fig. 5. Forest plot of intervention effects on anxiety outcomes.

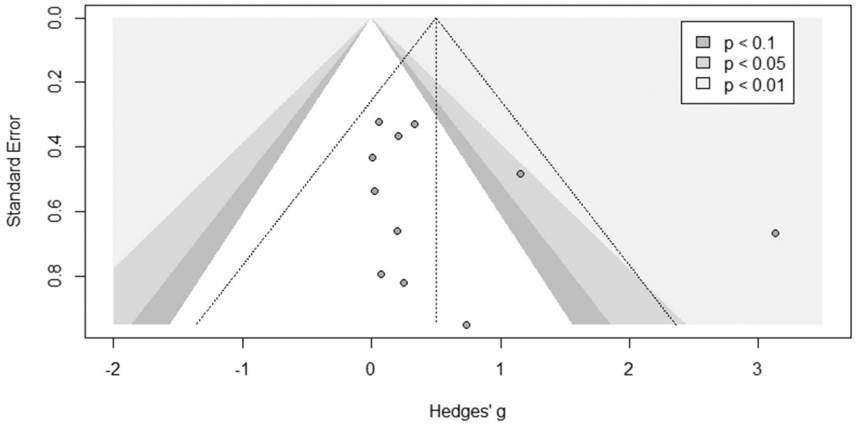

Fig. 6. Contour-enhanced funnel plot of Hedges' g effects included in meta-analysis of depression outcomes.

compromised studies' comparability (Deeks et al., 2022).

4.6. Agreements and disagreements with other literature

This review offers a distinctive contribution to the field. To our knowledge, it is the first systematic review and/or meta-analysis to synthesize evidence on the impact of in-person, co-designed mental health interventions on anxiety and depression symptoms. Previous reviews largely focused on digital and healthcare interventions and did not quantitatively pool evidence on anxiety and depression outcomes. Interestingly, despite the differences between this review and previous literature, their conclusions largely align.

Firstly, the increasing interest in co-design resonates with previous research (Bevan-Jones et al., 2020; Grindell et al., 2022). The wide range of approaches to co-design also agrees with earlier findings (Bevan-Jones et al., 2020; Orlowski et al., 2015). Additionally, the lack of clarity in the literature regarding what co-design entails aligns with the views of Dekker and Williams (2017) and O'Brien et al. (2020).

The observed lack of standardized terminology or consistent reporting of co-design activities corroborates the findings of prior research (Clarke et al., 2017; Dekker and Williams, 2017; Halvorsrud et al., 2021; Orlowski et al., 2015; Slattery et al., 2020). Although included studies' definitions of co-design, co-production, co-development, etc., align with the co-design definition employed in this

Table 3
Summary of GRADE evaluations.

Question: What is the effect of in-person, co-designed mental health interventions on anxiety and depression outcomes?

Population: Any
Intervention: In-person, co-designed mental health interventions
Comparison: No intervention/services as usual/wait-list control

Outcomes	№ of participants (№ of studies) Follow-up	Certainty of the evidence (GRADE)	Anticipated absolute effects: In-person, co-designed mental health interventions
Depression	809 (11 RCTs)	⨁⨁◯◯ Low[a,b]	SMD 0.5 SD higher (0.08 lower to 1.08 higher)
Anxiety	527 (7 RCTs)	⨁⨁◯◯ Low[b,c]	SMD 0.12 SD higher (0.1 lower to 0.33 higher)

CI: confidence interval; **SMD:** standardized mean difference

GRADE Working Group grades of evidence

High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.

Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.

Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Explanations

a. 10 studies were evaluated as having some risk of bias concerns, while one study (Sanfilippo et al., 2020) was evaluated as having high risk of bias due reporting analyses not specified in the protocol.

b. Studies had small sample sizes and effect sizes had wide confidence intervals.

c. All studies evaluated as having some risk of bias concerns e.g., due to self-report measurements.

review, there is not yet a universally agreed-upon operationalization of co-design. Inconsistent terminology presents challenges for reviews and meta-analyses of co-designed intervention studies. Thus, like Dekker and Williams (2017), this review underscores the need for consistency in terminology describing co-design and recommends that authors publish more details regarding co-design activities.

Our meta-analysis results are consistent with Halvorsrud et al.'s (2021) finding that the effect of co-creation on mental health outcomes was non-significant. This contrasts sharply with claims that co-design has the potential to make interventions more effective (Bevan-Jones et al., 2020; Burkett, 2016; Blomkamp, 2018; Centre for Coproduction, n.d.), and raises questions about the widespread endorsement of co-design in mental health fields.

Lastly, during the screening process, several studies concerning interventions explicitly labeled as 'co-designed' were excluded as their evaluation methods (e.g., pre-post designs) did not meet eligibility criteria. Furthermore, the quality of included studies ranged from low to moderate. Together, these findings highlight the need for more high-quality research utilizing rigorous evaluation methods in co-design literature, a conclusion that corresponds with those of previous reviews (Bevan-Jones et al., 2020; De Cotta et al., 2021; Norton, 2021; Orlowski et al., 2015).

4.7. Implications for policy and practice

Based on the low-quality evidence identified by this systematic review and the non-significant results of our meta-analyses, the current evidence does not support the recommendation of in-person, co-designed mental health interventions for reducing anxiety and depression symptoms.

To our knowledge, this is the first systematic review and meta-analysis of its kind, providing evidence on the effectiveness of in-person, co-designed interventions for reducing anxiety and depression. To date, no studies have conclusively demonstrated that co-design enhances intervention effectiveness compared to non-co-designed approaches. Therefore, those advocating for the use of co-design in mental health work based on the assumption that co-design engenders effective, in-person interventions, or on the basis that co-design bolsters intervention effectiveness, may need to reevaluate their positions. Co-design is currently a popular concept within mental health fields, but more discourse is needed regarding its true value, particularly its relation to intervention effectiveness.

However, this does not mean that co-design approaches should be completely discarded. Practitioners should not halt ongoing co-designed mental health interventions, but rather, they should ensure these interventions are rigorously evaluated to contribute to much-needed evidence in this area. If future reviews, armed with higher-quality evidence and greater confidence in effect estimates, confirm the findings of this review, then the use of co-design approaches for this purpose should be reconsidered. Until then, practitioners and policy-makers should be cautious of any claims about co-design and intervention effectiveness with regards to reducing anxiety and depression symptoms.

4.8. Future research

This review highlights the need for higher-quality evidence in this area, particularly fully-fledged evaluation trials of co-designed mental health interventions, as opposed to preliminary feasibility and pilot trials. Studies with sample sizes sufficiently powered to detect differences between groups are required.

There's a clear need for standardized terminology for referring to co-design and its sub-processes. Establishing recommended reporting standards for describing co-design activities would be beneficial. Consistent terminology would aid future reviews in identifying all relevant evidence and facilitate comparisons between studies. Uniformity in tools for measuring anxiety and depression would also be advantageous.

Despite burgeoning co-design literature, details of how co-design unfolds in practice and the theories underpinning such participatory approaches remain largely unclear. Future studies should detail the co-design processes used, the extent to which interventions are co-designed, and any theory or evidence supporting co-design methods. Researchers should also clearly define the roles of each stakeholder (or group) involved in the co-design process, as well as the stages at which they were involved. While journal word count limits may make this challenging, additional materials or other co-design forums can facilitate this type of information sharing (Davies and Bergin, 2021). Detailed descriptions of co-design would allow for meta-regressions that could provide insights into whether the degree of co-design, the type of co-design processes, and the stages at which co-design occurs, are associated with intervention outcomes.

To truly understand the added value of co-design (or lack thereof), comparative effectiveness trials are required. These should include three study arms: (1) a standard mental health intervention, (2) the same intervention, but adapted using co-design, and (3) no intervention or standard care.

Although this review focused on effectiveness, it's important to explore the functions of co-design beyond intervention effectiveness, particularly if future research indicates that co-design does not enhance effectiveness. Qualitative research on co-design exists (e.g., Grindell et al., 2022; O'Brien et al., 2020; Reed et al., 2020), but more is welcome, particularly comprehensive analyses of the mechanisms and outcomes of co-design, including potential benefits and harms in different contexts.

Finally, searches identified many relevant protocols and soon-to-be published studies. This is an emerging research area and it will be important to update this review as more, and higher quality, evidence becomes available.

5. Conclusion

Despite widespread enthusiasm and endorsement for co-design, research in this field remains in the early stages. High-quality evaluations of co-designed mental health interventions are relatively rare. Addressing a significant gap in the literature, this review has revealed that, as of now, there is insufficient evidence to conclude that in-person, co-designed mental health interventions have a significant impact on anxiety and depression. The field needs more, and higher quality, full-scale evaluation trials, along with uniformity in terminology and measurement instruments. Until such evidence becomes available, practitioners should approach claims regarding the effectiveness of co-designed mental health interventions for improving anxiety and depression with caution. Despite its popularity in mental health work, the value of co-design, particularly its relation to intervention effectiveness, warrants further debate and exploration.

Role of the funding source

This work was supported by funding from the Economic and Social Research Council Grand Union Doctoral Training Partnership. This funding source had no role in the study design, data collection, analysis, manuscript writing or decision to submit the article for publication.

CRediT authorship contribution statement

Tamsin Greene Barker: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing. **Aoife O'Higgins:** Conceptualization, Supervision, Writing – review & editing. **Peter Fonagy:** Supervision, Writing – review & editing. **Frances Gardner:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Tamsin Greene Barker reports financial support was provided by The Economic and Social Research Council Grand Union Doctoral Training Partnership.

Acknowledgements

We would like to thank Dr. Bridget Steele for her helpful advice regarding meta-analysis methods.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jad.2023.12.080.

References

Studies Included in Systematic Review

Abel, K.M., Bee, P., Gega, L., Gellatly, J., Kolade, A., Hunter, D., Callender, C., Carter, L. A., Meacock, R., Bower, P., Stanley, N., Calam, R., Wolpert, M., Stewart, P., Emsley, R., Holt, K., Linklater, H., Douglas, S., Stokes-Crossley, B., Green, J., 2020. An intervention to improve the quality of life in children of parents with serious mental illness: the Young SMILES feasibility RCT. Health Technol. Assess. 24 (59), 1–136. https://doi.org/10.3310/hta24590.

Alderson, H., Kaner, E., McColl, E., Howel, D., Fouweather, T., McGovern, R., Copello, A., Brown, H., McArdle, P., Smart, D., Brown, R., Lingam, R., 2020. A pilot feasibility randomised controlled trial of two behaviour change interventions compared to usual care to reduce substance misuse in looked after children and care leavers aged 12–20 years: the SOLID study. PloS One 15 (9), e0238286. https://doi.org/10.1371/journal.pone.0238286.

Alsubaie, M., Dickens, C., Dunn, B.D., Gibson, A., Ukoumunne, O.C., Evans, A., Vicary, R., Gandhi, M., Kuyken, W., 2018. Feasibility and acceptability of mindfulness-based cognitive therapy compared with mindfulness-based stress reduction and treatment as usual in people with depression and cardiovascular disorders: a three-arm randomised controlled trial. Mindfulness 11 (1), 30–50. https://doi.org/10.1007/s12671-018-0999-8.

Ashton, L.M., Morgan, P.J., Hutchesson, M.J., Rollo, M.E., Collins, C.E., 2017. Feasibility and preliminary efficacy of the 'HEYMAN' healthy lifestyle program for young men: a pilot randomised controlled trial. Nutr. J. 16 (2) https://doi.org/10.1186/s12937-017-0227-8.

Betancourt, T.S., Berent, J.M., Freeman, J., Frounfelker, R.L., Brennan, R.T., Abdi, S., Maalim, A., Abdi, A., Mishra, T., Gautam, B., Creswell, J.W., Beardslee, W.R., 2020. Family-based mental health promotion for Somali bantu and Bhutanese refugees: feasibility and acceptability trial. J. Adolesc. Health 66 (3), 336–344. https://doi.org/10.1016/j.jadohealth.2019.08.023.

Bhana, A., Mellins, C.A., Petersen, I., Alicea, S., Myeza, N., Holst, H., Abrams, E., John, S., Chhagan, M., Nestadt, D.F., Leu, C.S., McKay, M., 2014. The VUKA family program: piloting a family-based psychosocial intervention to promote health and mental health among HIV infected early adolescents in South Africa. AIDS Care 26 (1), 1–11. https://doi.org/10.1080/09540121.2013.806770.

Carlyle, M., Rockliff, H., Edwards, R., Ene, C., Karl, A., Marsh, B., Hartley, L., Morgan, C. J., 2019. Investigating the feasibility of brief compassion focused therapy in individuals in treatment for opioid use disorder. Subst. Abuse Res. Treat. 13, 1178221819836726 https://doi.org/10.1177/1178221819836726.

Chomat, A.M., Menchú, A.I., Andersson, N., Ramirez-Zea, M., Pedersen, D., Bleile, A., Letona, P., Araya, R., 2019. Women's circles as a culturally safe psychosocial intervention in Guatemalan indigenous communities: a community-led pilot randomised trial. BMC Women's Health 19 (53). https://doi.org/10.1186/s12905-019-0744-z.

Hassouneh, D., Nguyen, T., Chen, Z., McNeff, E., 2013. Healing pathways: a program for women with physical disabilities and depression. Rehabil. Res. Pract. 2013, 1–15. https://doi.org/10.1155/2013/649875.

Hedemann, E.R., 2019. Reducing Risks for Anxiety and Depression Among Urban Youth: Leveraging After School Programs to Promote Emotion Regulation (No. 4209). FIU Electronic Theses and Dissertations. https://digitalcommons.fiu.edu/etd/4209/?utm_source=digitalcommons.fiu.edu%2Fetd%2F4209&utm_medium=PDF&utm_campaign=PDFCoverPages.

Henson, A., Ruglis, J., Sinacore, A., Fitzpatrick, M., Lanteigne, D., 2020. Self-compassion for youth in small city centres: a school-based pilot project. Couns. Psychother. Res. 21 (3), 719–728. https://doi.org/10.1002/capr.12374.

Karasz, A., Raghavan, S., Patel, V., Zaman, M., Akhter, L., Kabita, M., 2015. ASHA: using participatory methods to develop an asset-building mental health intervention for Bangladeshi immigrant women. Prog. Community Health Partnersh. 9 (4), 469–470. https://doi.org/10.1353/cpr.2015.0069.

Karasz, A., Anne, S., Hamadani, J.D., Tofail, F., 2021. The ASHA (Hope) project: testing an integrated depression treatment and economic strengthening intervention in rural Bangladesh: a pilot randomized controlled trial. Int. J. Environ. Res. Public Health 18 (1), 279. https://doi.org/10.3390/ijerph18010279.

Nápoles, A.M., Ortíz, C., Santoyo-Olsson, J., Stewart, A.L., Gregorich, S., Lee, H.E., Durón, Y., McGuire, P., Luce, J., 2015. Nuevo Amanecer: results of a randomized controlled trial of a community-based, peer-delivered stress management intervention to improve quality of life in Latinas with breast cancer. Am. J. Public Health 105 (S3), e55–e63. https://doi.org/10.2105/ajph.2015.302598.

Nápoles, A.M., Santoyo-Olsson, J., Stewart, A.L., Ortiz, C., Samayoa, C., Torres-Nguyen, A., Palomino, H., Coleman, L., Urias, A., Gonzalez, N., Cervantes, S.A., Totten, V.Y., 2020. Nuevo Amanecer-II: results of a randomized controlled trial of a community-based participatory, peer-delivered stress management intervention for rural Latina breast cancer survivors. Psycho-Oncology 29 (11), 1802–1814. https://doi.org/10.1002/pon.5481.

Nestadt, D.F., Saisaengjan, C., McKay, M.M., Bunupuradah, T., Pardo, G., Lakhonpon, S., Gopalan, P., Leu, C.S., Petdachai, W., Kosalaraksa, P., Srirompotong, U., Ananworanich, J., Mellins, C.A., 2019. CHAMP+ Thailand: pilot randomized control trial of a family-based psychosocial intervention for perinatally HIV-infected early adolescents. AIDS Patient Care STDS 33 (5), 227–236. https://doi.org/10.1089/apc.2019.0021.

Pile, V., Winstanley, A., Oliver, A., Bennett, E., Lau, J.Y., 2021. Targeting image-based autobiographical memory in childhood to prevent emotional disorders: intervention development and a feasibility randomised controlled trial. Behav. Res. Ther. 144, 103913 https://doi.org/10.1016/j.brat.2021.103913.

Puffer, E.S., Green, E.P., Sikkema, K.J., Broverman, S.A., Ogwang-Odhiambo, R.A., Pian, J., 2016. A church-based intervention for families to promote mental health and prevent HIV among adolescents in rural Kenya: results of a randomized trial. J. Consult. Clin. Psychol. 84 (6), 511–525. https://doi.org/10.1037/ccp0000076.

Rüsch, N., Staiger, T., Waldmann, T., Dekoj, M.C., Brosch, T., Gabriel, L., Bahemann, A., Oexle, N., Klein, T., Nehf, L., Becker, T., 2019. Efficacy of a peer-led group program for unemployed people with mental health problems: pilot randomized controlled trial. Int. J. Soc. Psychiatry 65 (4), 333–337. https://doi.org/10.1177/0020764019846171.

Sanfilippo, K.R.M., McConnell, B., Cornelius, V., Darboe, B., Huma, H.B., Gaye, M., Ceesay, H., Ramchandani, P., Cross, I., Glover, V., Stewart, L., 2020. Community psychosocial music intervention (CHIME) to reduce antenatal common mental disorder symptoms in the Gambia: a feasibility trial. BMJ Open 10 (11), e040287. https://doi.org/10.1136/bmjopen-2020-040287.

Other references

Adams, J., Hillier-Brown, F.C., Moore, H.J., Lake, A.A., Araujo-Soares, V., White, M., Summerbell, C., 2016. Searching and synthesising 'grey literature' and 'grey information' in public health: critical reflections on three case studies. Syst. Rev. 5 (1) https://doi.org/10.1186/s13643-016-0337-y.

Akobeng, A.K., 2005. Understanding randomised controlled trials. Arch. Dis. Child. 90 (8), 840–844. https://doi.org/10.1136/adc.2004.058222.

Alford, J., 2013. The multiple facets of co-production: building on the work of Elinor Ostrom. Public Manag. Rev. 16 (3), 299–316. https://doi.org/10.1080/14719037.2013.806578.

Althubaiti, A., 2016. Information bias in health research: definition, pitfalls, and adjustment methods. J. Multidiscip. Healthc. 211 https://doi.org/10.2147/jmdh.s104807.

American Psychological Association, n.d. External validity. In APA Dictionary of Psychology. Retrieved April 10, 2022, from https://dictionary.apa.org/external-validity.

Andrade, C., 2018. Internal, external, and ecological validity in research design, conduct, and evaluation. Indian J. Psychol. Med. 40 (5), 498–499. https://doi.org/10.4103/ijpsym.ijpsym_334_18.

Andrade, C., 2020. Mean difference, standardized mean difference (SMD), and their use in meta-analysis. J. Clin. Psychiatry 81 (5). https://doi.org/10.4088/jcp.20f13681.

Antony, M.M., Bieling, P.J., Cox, B.J., Enns, M.W., Swinson, R.P., 1998. Psychometric properties of the 42-item and 21-item versions of the depression anxiety stress scales in clinical groups and a community sample. Psychol. Assess. 10 (2), 176–181. https://doi.org/10.1037/1040-3590.10.2.176.

Asanbe, C., Gaba, A., Yang, J., 2018. Mental health is a human right. Psychology International. https://www.apa.org/international/pi/2018/12/mental-health-rights.

Balshem, H., Helfand, M., Schünemann, H.J., Oxman, A.D., Kunz, R., Brozek, J., Vist, G. E., Falck-Ytter, Y., Meerpohl, J., Norris, S., 2011. GRADE guidelines: 3. Rating the quality of evidence. J. Clin. Epidemiol. 64 (4), 401–406. https://doi.org/10.1016/j.jclinepi.2010.07.015.

Beck, A., Beck, J., Jolly, J., Steer, R., 2005. Beck Youth Inventories, 2nd ed. Harcourt Assessment.

Bergin, A.D., Vallejos, E.P., Davies, E.B., Daley, D., Ford, T., Harold, G., Hetrick, S., Kidner, M., Long, Y., Merry, S., Morriss, R., Sayal, K., Sonuga-Barke, E., Robinson, J., Torous, J., Hollis, C., 2020. Preventive digital mental health interventions for children and young people: a review of the design and reporting of research. Npj Digit. Med. 3 (1) https://doi.org/10.1038/s41746-020-00339-7.

Beusenberg, M., Orley, J.H., 1994. A User's Guide to the Self Reporting Questionnaire (SRQ). World Health Organization, Division of Mental Health. https://apps.who.int/iris/handle/10665/61113.

Bevan-Jones, R., Stallard, P., Agha, S.S., Rice, S., Werner-Seidler, A., Stasiak, K., Kahn, J., Simpson, S.A., Alvarez-Jimenez, M., Rice, F., Evans, R., Merry, S., 2020. Practitioner review: co-design of digital mental health technologies with children and young people. J. Child Psychol. Psychiatry 61 (8), 928–940. https://doi.org/10.1111/jcpp.13258.

Blomkamp, E., 2018. The promise of co-design for public policy. Aust. J. Public Adm. 77 (4), 729–743. https://doi.org/10.1111/1467-8500.12310.

Borenstein, M., Hedges, L.V., Higgins, J.P., Rothstein, H.R., 2010. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res. Synth. Methods 1 (2), 97–111. https://doi.org/10.1002/jrsm.12.

Boutron, I., Page, M.J., Higgins, J.P.T., Altman, D.G., Lundh, A., Hróbjartsson, A., 2022. Chapter 7: considering bias and conflicts of interest among the included studies. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V. A. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions version 6.3. Cochrane. http://www.training.cochrane.org/handbook.

Burkett, I., 2016. An introduction to co-design. Centre for Social Impact. http://ingridburkett.com/wp-content/uploads/2017/09/Introduction-to-Codesign-2.pdf.

Centre for Coproduction, n.d. What is co-production? Centre for Coproduction in Mental Health and Social Care. http://centreforcoproduction.com/about-us/ (accessed 24 August 2022).

Chan, D., 2008. So why ask me? Are self-report data really that bad? In: Lance, C.E., Vandenberg, R.J. (Eds.), Statistical and Methodological Myths and Urban Legends: Doctrine, Verity, and Fable in the Organizational and Social Sciences. Routledge, pp. 309–336.

Chandler, J., Cumpston, M., Thomas, J., Higgins, J.P.T., Deeks, J.J., Clarke, M.J., 2022. Chapter I: introduction. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions version 6.3. Cochrane. http://www.training.cochrane.org/handbook.

Chorpita, B.F., Yim, L., Moffitt, C., Umemoto, L.A., Francis, S.E., 2000. Assessment of symptoms of DSM-IV anxiety and depression in children: a revised child anxiety and depression scale. Behav. Res. Ther. 38 (8), 835–855. https://doi.org/10.1016/s0005-7967(99)00130-8.

Clarke, D., Jones, F., Harris, R., Robert, G., 2017. What outcomes are associated with developing and implementing co-produced interventions in acute healthcare settings? A rapid evidence synthesis. BMJ Open. 7 (7), e014650 https://doi.org/10.1136/bmjopen-2016-014650.

CMHL, n.d. CMHL co-design program. Centre for Mental Health Learning (CMHL). https://cmhl.org.au/work/cmhl-co-design-program (accessed 22 August 2022).

Cochrane Effective Practice and Organisation of Care (EPOC), 2017. Data Collection Form. EPOC Resources for review authors, Cochrane Effective Practice and Organisation of Care (EPOC). https://epoc.cochrane.org/resources/epoc-resources-review-authors.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. Routledge.

Cohn, L.D., Becker, B.J., 2003. How meta-analysis increases statistical power. Psychol. Methods 8 (3), 243–253. https://doi.org/10.1037/1082-989x.8.3.243.

Cook, T.D., Campbell, D.T., 1976. The design and conduct of quasi experiments and true experiments in field settings. In: Dunette, M.D. (Ed.), Handbook of Industrial and Organizational Psychology. Rand McNally, pp. 115–136.

Veritas Health Innovation, 2022. Covidence systematic review software, Melbourne Australia. www.covidence.org.

Cox, J.L., Holden, J., Henshaw, C., 2014. Perinatal Mental Health: the Edinburgh Postnatal Depression Scale (EPDS) Manual. RCPsych publications, pp. 17–26.

Craane, B., Dijkstra, P.U., Stappaerts, K., de Laat, A., 2012. Methodological quality of a systematic review on physical therapy for temporomandibular disorders: influence of hand search and quality scales. Clin. Oral Investig. 16 (1), 295–303. https://doi.org/10.1007/s00784-010-0490-y.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., Petticrew, M., 2008. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ a1655. https://doi.org/10.1136/bmj.a1655.

Davies, E.B., Bergin, A.D., 2021. Commentary: Let's get digital: a commentary on Halldorsson et al.'s call for more rigorous development and evaluation of immersive digital interventions for children and young people's mental health. J. Child Psychol. Psychiatry 62 (5), 606–609. https://doi.org/10.1111/jcpp.13423.

de Cotta, T., Knox, J., Farmer, J., White, C., Davis, H., 2021. Community co-produced mental health initiatives in rural Australia: a scoping review. Aust. J. Rural Health 29 (6), 865–878. https://doi.org/10.1111/ajr.12793.

Deeks, J.J., Higgins, J.P.T., Altman, D.G., 2022. Chapter 10: analysing data and undertaking meta-analyses. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions version 6.3. Cochrane. http://www.training.cochrane.org/handbook.

Deke, J., Sama-Miller, E., Hershey, A., 2015. Addressing Attrition Bias in Randomized Controlled Trials: Considerations for Systematic Evidence Reviews. Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. https://homvee.acf.hhs.gov/sites/default/files/2019-06/HomVEE-Attrition-White_Paper-7-2015.pdf.

Dekker, M.R., Williams, A.D., 2017. The use of user-centered participatory design in serious games for anxiety and depression. Games Health J. 6 (6), 327–333. https://doi.org/10.1089/g4h.2017.0058.

Derogatis, L.R., Lipman, R.S., Rickels, K., Uhlenhuth, E.H., Covi, L., 1974. The Hopkins symptom checklist (HSCL): a self-report symptom inventory. Behav. Sci. 19 (1), 1–15. https://doi.org/10.1002/bs.3830190102.

Derogatis, L.R., Melisaratos, N., 1983. The brief symptom inventory: an introductory report. Psychol. Med. 13 (3), 595–605. https://doi.org/10.1017/S0033291700048017.

DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. Control. Clin. Trials 7 (3), 177–188. https://doi.org/10.1016/0197-2456(86)90046-2.

Dickersin, K., Min, Y.I., 1993. Publication bias: the problem that won't go away. Ann. N. Y. Acad. Sci. 703, 135–148. https://doi.org/10.1111/j.1749-6632.1993.tb26343.x.

Dimopoulos-Bick, T.L., O'Connor, C., Montgomery, J., Szanto, T., Fisher, M., Sutherland, V., Baines, H., Orcher, P., Stubbs, J., Maher, L., Verma, R., Palmer, V.J., 2019. "Anyone can co-design?": a case study synthesis of six experience-based co-design (EBCD) projects for healthcare systems improvement in New South Wales, Australia. Patient Experience J. 6 (2), 93–104. https://doi.org/10.35680/2372-0247.1365.

Early Intervention Foundation (EIF), 2022. Early intervention foundation and what works for children's social care announce decision to merge [Press release]. https://www.eif.org.uk/press-release/early-intervention-foundation-and-what-works-for-children-s-social-care-announce-decision-to-merge.

Edwards, P., Clarke, M., DiGuiseppi, C., Pratap, S., Roberts, I., Wentz, R., 2002. Identification of randomized controlled trials in systematic reviews: accuracy and reliability of screening records. Stat. Med. 21 (11), 1635–1640. https://doi.org/10.1002/sim.1190.

Egger, M., Smith, G.D., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. BMJ 315 (7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629.

Eldridge, S., Campbell, M.K., Campbell, M.J., Drahota, A.K., Giraudeau, B., Reeves, B.C., Siegfried, N., Higgins, J.P.T., 2021. Revised Cochrane risk of bias tool for randomized trials (RoB 2) Additional considerations for cluster-randomized trials (RoB 2 CRT). Cochrane Bias Methods Group (BMG). https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool/rob-2-for-cluster-randomized-trials?authuser=0.

Evans, L., Green, S., Howe, C., Sharma, K., Marinho, F., Bell, D., Thomas, P., 2014. Improving patient and project outcomes using interorganisational innovation, collaboration and co-design. London J. Prim. Care 6 (2), 29–34. https://doi.org/10.1080/17571472.2014.11493410.

Eyles, H., Jull, A., Dobson, R., Firestone, R., Whittaker, R., te Morenga, L., Goodwin, D., Mhurchu, C.N., 2016. Co-design of mHealth delivered interventions: a systematic review to assess key methods and processes. Curr. Nutr. Rep. 5 (3), 160–167. https://doi.org/10.1007/s13668-016-0165-7.

Faulstich, M.E., Carey, M.P., Ruggiero, L., Enyart, P., Gresham, F., 1986. Assessment of depression in childhood and adolescence: an evaluation of the Center for Epidemiological Studies Depression Scale for children (CES-DC). Am. J. Psychiatry 143 (8), 1024–1027. https://doi.org/10.1176/ajp.143.8.1024.

Fleming, T.M., de Beurs, D., Khazaal, Y., Gaggioli, A., Riva, G., Botella, C., Baños, R.M., Aschieri, F., Bavin, L.M., Kleiboer, A., Merry, S., Lau, H.M., Riper, H., 2016. Maximizing the impact of e-therapy and serious gaming: time for a paradigm shift. Front. Psych. 7 https://doi.org/10.3389/fpsyt.2016.00065.

Forbes, D., 2013. Blinding: an essential component in decreasing risk of bias in experimental designs. Evid Based Nurs 16 (3), 70–71. https://doi.org/10.1136/eb-2013-101382.
Fox, C., Smith, A., Traynor, P., Harrison, J., 2018. Co-creation and Co-Production in the United Kingdom – A Rapid Evidence Assessment. Policy Evaluation and Research Unit. Manchester Metropolitan University. https://mmuperu.co.uk/wp-content/uploads/2020/09/Co-Creation_and_Co-Production_in_the_United_Kingdom_-_A_Rapid_Evidence_Assessment_-_March_2018.pdf.
Gartlehner, G., Affengruber, L., Titscher, V., Noel-Storr, A., Dooley, G., Ballarini, N., König, F., 2020. Single-reviewer abstract screening missed 13 percent of relevant studies: a crowd-based, randomized controlled trial. J. Clin. Epidemiol. 121, 20–28. https://doi.org/10.1016/j.jclinepi.2020.01.005.
Granholm, A., Alhazzani, W., Møller, M.H., 2019. Use of the GRADE approach in systematic reviews and guidelines. Br. J. Anaesth. 123 (5), 554–559. https://doi.org/10.1016/j.bja.2019.08.015.
Grindell, C., Coates, E., Croot, L., O'Cathain, A., 2022. The use of co-production, co-design and co-creation to mobilise knowledge in the management of health conditions: a systematic review. BMC Health Serv. Res. 22 (1) https://doi.org/10.1186/s12913-022-08079-y.
Guyatt, G., Oxman, A.D., Akl, E.A., Kunz, R., Vist, G., Brozek, J., Norris, S., Falck-Ytter, Y., Glasziou, P., deBeer, H., 2011. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. J. Clin. Epidemiol. 64 (4), 383–394. https://doi.org/10.1016/j.jclinepi.2010.04.026.
Halldorsson, B., Hill, C., Waite, P., Partridge, K., Freeman, D., Creswell, C., 2021. Annual research review: immersive virtual reality and digital applied gaming interventions for the treatment of mental health problems in children and young people: the need for rigorous treatment development and clinical evaluation. J. Child Psychol. Psychiatry 62 (5), 584–605. https://doi.org/10.1111/jcpp.13400.
Halvorsrud, K., Kucharska, J., Adlington, K., Rüdell, K., Brown Hajdukova, E., Nazroo, J., Haarmans, M., Rhodes, J., Bhui, K., 2021. Identifying evidence of effectiveness in the co-creation of research: a systematic review and meta-analysis of the international healthcare literature. J. Public Health 43 (1), 197–208. https://doi-org.ezproxy-prd.bodleian.ox.ac.uk/10.1093/pubmed/fdz126.
Harrer, M., Cuijpers, P., Furukawa, T.A., Ebert, D.D., 2021. Doing Meta-analysis with R: A Hands-on Guide. Chapman & Hall/CRC Press, Boca Raton, FL and London. ISBN 978-0-367-61007-4.
Hariton, E., Locascio, J.J., 2018. Randomised controlled trials - the gold standard for effectiveness research. BJOG 125 (13), 1716. https://doi.org/10.1111/1471-0528.15199.
Hedges, L.V., Olkin, I., 1985. Statistical Methods for Meta-analysis. Academic Press.
Higgins, J.P.T., Savović, J., Page, M.J., Elbers, R.G., Sterne, J.A.C., 2022a. Chapter 8: assessing risk of bias in a randomized trial. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions version 6.3. Cochrane. http://www.training.cochrane.org/handbook.
Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), 2022b. Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022). Cochrane www.training.cochrane.org/handbook.
Hohls, J.K., König, H.H., Quirke, E., Hajek, A., 2021. Anxiety, depression and quality of life—a systematic review of evidence from longitudinal observational studies. Int. J. Environ. Res. Public Health 18 (22), 12022. https://doi.org/10.3390/ijerph182212022.
HSE, 2017. Co-Production in Practice Guidance Document: 2018–2020. Health Service Executive (HSE). https://www.hse.ie/eng/services/list/4/mental-health-services/advancingrecoveryireland/national-framework-for-recovery-in-mental-health/co-production-in-practice-guidance-document-2018-to-2020.pdf.
IntHout, J., Ioannidis, J.P., Borm, G.F., 2014. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med. Res. Methodol. 14 (1) https://doi.org/10.1186/1471-2288-14-25.
Janssen, M.F., Pickard, A.S., Golicki, D., Gudex, C., Niewada, M., Scalone, L., Swinburn, P., Busschbach, J., 2013. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care Rehab. 22 (7), 1717–1727. https://doi.org/10.1007/s11136-012-0322-4.
Jessup, R.L., Osborne, R.H., Buchbinder, R., Beauchamp, A., 2018. Using co-design to develop interventions to address health literacy needs in a hospitalised population. BMC Health Serv. Res. 18 (1) https://doi.org/10.1186/s12913-018-3801-7.
Kim, Y., Steiner, P., 2016. Quasi-experimental designs for causal inference. Educ. Psychol. 51 (3–4), 395–405. https://doi.org/10.1080/00461520.2016.1207177.
King, P.T., Cormack, D., Edwards, R., Harris, R., Paine, S.J., 2022. Co-design for indigenous and other children and young people from priority social groups: a systematic review. SSM-Popul. Health. 18, 101077 https://doi.org/10.1016/j.ssmph.2022.101077.
Kovacs, M., 1992. Manual for the Children's Depression Inventory. Multi-Health Systems.
Kroenke, K., Strine, T.W., Spitzer, R.L., Williams, J.B., Berry, J.T., Mokdad, A.H., 2009. The PHQ-8 as a measure of current depression in the general population. J. Affect. Disord. 114 (1–3), 163–173. https://doi.org/10.1016/j.jad.2008.06.026.
Langan, D., Higgins, J.P., Jackson, D., Bowden, J., Veroniki, A.A., Kontopantelis, E., Viechtbauer, W., Simmonds, M., 2019. A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. Res. Synth. Methods 10 (1), 83–98. https://doi.org/10.1002/jrsm.1316.
Lasserson, T.J., Thomas, J., Higgins, J.P.T., 2022. Chapter 1: starting a review. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions version 6.3. Cochrane. http://www.training.cochrane.org/handbook.
Lefebvre, C., Glanville, J., Briscoe, S., Featherstone, R., Littlewood, A., Marshall, C., Metzendorf, M.I., Noel-Storr, A., Paynter, R., Rader, T., Thomas, J., Wieland, L.S., 2022. Chapter 4: searching for and selecting studies. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions version 6.3. Cochrane. http://www.training.cochrane.org/handbook.
Li, S., Achilles, M., Spanos, S., Habak, S., Werner-Seidler, A., O'Dea, B., 2022. A cognitive behavioural therapy smartphone app for adolescent depression and anxiety: co-design of ClearlyMe. Cogn. Behav. Ther. 15 https://doi.org/10.1017/s1754470x22000095.
Li, T., Higgins, J.P.T., Deeks, J.J., 2022b. Chapter 5: collecting data. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions version 6.3. Cochrane. http://www.training.cochrane.org/handbook.
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. Behav. Res. Ther. 33 (3), 335–343. https://doi.org/10.1016/0005-7967(94)00075-u.
Lüdecke, D., 2019. Package 'esc': effect size computation for meta analysis. R package version 0.5.1. CRAN repository. https://cran.r-project.org/web/packages/esc/esc.pdf.
Mahmud, S., Mohsin, M., Dewan, M.N., Muyeed, A., 2023. The global prevalence of depression, anxiety, stress, and insomnia among general population during COVID-19 pandemic: a systematic review and meta-analysis. Trends Psychol. 31 (1), 143–170. https://doi.org/10.1007/s43076-021-00116-9.
March, J.S., Parker, J.D., Sullivan, K., Stallings, P., Conners, C.K., 1997. The multidimensional anxiety scale for children (MASC): factor structure, reliability, and validity. J. Am. Acad. Child Adolesc. Psychiatry 36 (4), 554–565. https://doi.org/10.1097/00004583-199704000-00019.
Marney, T., Elderton, K., 2021. Five factors to consider when doing co-design in mental health. Nous Group. https://nousgroup.com/insights/co-design-in-mental-health/.
McGill, B., Corbett, L., Grunseit, A.C., Irving, M., O'Hara, B.J., 2022. Co-produce, co-design, co-create, or co-construct—who does it and how is it done in chronic disease prevention? A scoping review. Healthcare 10 (4), 647. https://doi.org/10.3390/healthcare10040647.
McKenzie, J.E., Brennan, S.E., 2022. Chapter 12: synthesizing and presenting findings using other methods. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions Version 6.3. Cochrane. http://www.training.cochrane.org/handbook.
McKenzie, J.E., Brennan, S.E., Ryan, R.E., Thomson, H.J., Johnston, R.V., 2022. Chapter 9: summarizing study characteristics and preparing for synthesis. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions version 6.3. Cochrane. http://www.training.cochrane.org/handbook.
McKercher, K.A., 2020. Beyond Sticky Notes: Co-design for Real: Mindsets, Methods and Movements. Beyond Sticky Notes.
McManus, R.J., Wilson, S., Delaney, B.C., Fitzmaurice, D.A., Hyde, C.J., Tobias, R.S., Jowett, S., Hobbs, F.D.R., 1998. Review of the usefulness of contacting other experts when conducting a literature search for systematic reviews. BMJ 317 (7172), 1562–1563. https://doi.org/10.1136/bmj.317.7172.1562.
McManus, S., Bebbington, P., Jenkins, R., Brugha, T., 2016. Adult psychiatric morbidity survey: survey of mental health and wellbeing, 2014. NHS Digital. https://digital.nhs.uk/data-and-information/publications/statistical/adult-psychiatric-morbidity-survey/adult-psychiatric-morbidity-survey-survey-of-mental-health-and-wellbeing-england-2014.
McPin Foundation, n.d. https://mcpin.org.
Mental Health Commission of Canada, 2016. Advancing the mental health strategy for Canada: a framework for action (2017–2022). https://www.mentalhealthcommission.ca/wp-content/uploads/drupal/2016-08/advancing_the_mental_health_strategy_for_canada_a_framework_for_action.pdf.
Mental Health Foundation, 2020. Mental Health Foundation strategy 2020–2025: making prevention happen. https://www.mentalhealth.org.uk/about-us/strategy.
Mind, 2020. Case study: mind membership co-design group. https://www.mind.org.uk/workplace/influence-and-participation-toolkit/how/methods/remote-influence-and-participation/mind-membership-focus-group-case-study/.
Doughty, B., 2021. What is co-production? MQ Mental health research. https://www.mqmentalhealth.org/co-production/.
Nederhof, A.J., 1985. Methods of coping with social desirability bias: a review. Eur. J. Soc. Psychol. 15 (3), 263–280. https://doi.org/10.1002/ejsp.2420150303.
NHS, n.d. Adult and older adult mental health. NHS England. https://www.england.nhs.uk/mental-health/adults/ (accessed 26 August 2022).
National Institute for Health and Care Excellence, 2016. Community engagement: improving health and wellbeing and reducing health inequalities. National Institute for Health and Care Excellence. https://www.nice.org.uk/guidance/ng44.
NIHR, 2015. Going the extra mile: improving the nation's health and wellbeing through public involvement in research. National Institute for Health Research. https://www.nihr.ac.uk/documents/about-us/our-contribution-to-research/how-we-involve-patients-carers-and-the-public/Going-the-Extra-Mile.pdf.
Norton, M.J., 2021. Co-production within child and adolescent mental health: a systematic review. Int. J. Environ. Res. Public Health 18 (22), 11897. https://doi.org/10.3390/ijerph182211897.
O'Brien, J., Fossey, E., Palmer, V.J., 2020. A scoping review of the use of co-design methods with culturally and linguistically diverse communities to improve or adapt

mental health services. Health Soc. Care Community 29 (1), 1–17. https://doi.org/10.1111/hsc.13105.
Oostermeijer, S., Brasier, C., Harvey, C., Hamilton, B., Roper, C., Martel, A., Fletcher, J., Brophy, L., 2021. Design features that reduce the use of seclusion and restraint in mental health facilities: a rapid systematic review. BMJ Open 11 (7), e046647. https://doi.org/10.1136/bmjopen-2020-046647.
Orlowski, S.K., Lawn, S., Venning, A., Winsall, M., Jones, G.M., Wyld, K., Damarell, R.A., Antezana, G., Schrader, G., Smith, D., Collin, P., Bidargaddi, N., 2015. Participatory research as one piece of the puzzle: a systematic review of consumer involvement in design of technology-based youth mental health and well-being interventions. JMIR Hum. Factors 2 (2), e12. https://doi.org/10.2196/humanfactors.4361.
Orygen, 2019. Co-designing with young people: the fundamentals. The National Centre of Excellence in Youth Mental Health. https://www.orygen.org.au/Training/Resources/Service-knowledge-and-development/Guidelines/Co-designing-with-young-people-The-fundamentals/Orygen-Co-designing-with-YP-the-fundamentals?ext=.
Ospina-Pinillos, L., Davenport, T., Mendoza Diaz, A., Navarro-Mancilla, A., Scott, E.M., Hickie, I.B., 2019. Using participatory design methodologies to co-design and culturally adapt the Spanish version of the mental health eClinic: qualitative study. J. Med. Internet Res. 21 (8), e14127 https://doi.org/10.2196/14127.
Ostrom, E., 1996. Crossing the great divide: coproduction, synergy, and development. World Dev. 24 (6), 1073–1087. https://doi.org/10.1016/0305-750x(96)00023-x.
Paez, A., 2017. Gray literature: an important resource in systematic reviews. J. Evid. Based Med. 10 (3), 233–240. https://doi.org/10.1111/jebm.12266.
Page, M.J., Higgins, J.P.T., Sterne, J.A.C., 2022. Chapter 13: assessing risk of bias due to missing results in a synthesis. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions Version 6.3. Cochrane. http://www.training.cochrane.org/handbook.
Palmer, V.J., Chondros, P., Furler, J., Herrman, H., Pierce, D., Godbee, K., Densley, K., Gunn, J.M., 2021. The CORE study—an adapted mental health experience codesign intervention to improve psychosocial recovery for people with severe mental illness: a stepped wedge cluster randomized-controlled trial. Health Expect. 24 (6), 1948–1961. https://doi.org/10.1111/hex.13334.
Palmer, V.J., Chondros, P., Piper, D., Callander, R., Weavell, W., Godbee, K., Potiriadis, M., Richard, L., Densely, K., Herrman, H., Furler, J., Pierce, D., Schuster, T., Iedema, R., Gunn, J., 2015. The CORE study protocol: a stepped wedge cluster randomised controlled trial to test a co-design technique to optimise psychosocial recovery outcomes for people affected by mental illness in the community mental health setting. BMJ Open 5 (3), e006688. https://doi.org/10.1136/bmjopen-2014-006688.
Palmer, V.J., Weavell, W., Callander, R., Piper, D., Richard, L., Maher, L., Boyd, H., Herrman, H., Furler, J., Gunn, J., Iedema, R., Robert, G., 2019. The participatory zeitgeist: an explanatory theoretical model of change in an era of coproduction and codesign in healthcare improvement. Med. Humanit. 45 (3), 247–257. https://doi.org/10.1136/medhum-2017-011398.
Pearce, E., Myles-Hooton, P., Johnson, S., Hards, E., Olsen, S., Clisu, D., Pais, S.M.A., Chesters, H.A., Shah, S., Jerwood, G., Politis, M., Melwani, J., Andersson, G., Shafran, R., 2021. Loneliness as an active ingredient in preventing or alleviating youth anxiety and depression: a critical interpretative synthesis incorporating principles from rapid realist reviews. Transl. Psychiatry 11 (1). https://doi.org/10.1038/s41398-021-01740-w.
Peters, J.L., Sutton, A.J., Jones, D.R., Abrams, K.R., Rushton, L., 2008. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. J. Clin. Epidemiol. 61 (10), 991–996. https://doi.org/10.1016/j.jclinepi.2007.11.010.
Petticrew, M., Roberts, H., 2006. Systematic Reviews in the Social Sciences: A Practical Guide, 1st ed. Blackwell Publishing. https://doi.org/10.1002/9780470754887.
Porche, M.V., Folk, J.B., Tolou-Shams, M., Fortuna, L.R., 2022. Researchers' perspectives on digital mental health intervention co-design with marginalized community stakeholder youth and families. Front. Psych. 13 https://doi.org/10.3389/fpsyt.2022.867460.
R Core Team, 2022. R: a language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/.
R Studio Team, 2022. RStudio: integrated development for R. RStudio, Inc. http://www.rstudio.com/.
Radloff, L.S., 1977. The CES-D scale: a self-report depression scale for research in the general population. Appl. Psychol. Measur. 1 (3), 385–401. https://doi.org/10.1177/014662167700100306.
Reed, H., Couturiaux, D., Davis, M., Edwards, A., Janes, E., Kim, H.S., Melendez-Torres, G.J., Murphy, S., Rotevatn, T.A., Smith, J., Evans, R., 2020. Co-production as an emerging methodology for developing school-based health interventions with students aged 11–16: systematic review of intervention types, theories and processes and thematic synthesis of Stakeholders' experiences. Prev. Sci. 22 (4), 475–491. https://doi.org/10.1007/s11121-020-01182-8.
Renedo Illarregi, E., 2022. Co-design as Healing: Exploring the Experiences of Participants Facing Mental Health Problems. PhD thesis.. The Open University https://oro.open.ac.uk/82218/1/CoDesignAsHealing-DoctoralDissertationErikaRenedoIllarregi.pdf.
Rockers, P.C., Røttingen, J.A., Shemilt, I., Tugwell, P., Bärnighausen, T., 2015. Inclusion of quasi-experimental studies in systematic reviews of health systems research. Health Policy 119 (4), 511–521. https://doi.org/10.1016/j.healthpol.2014.10.006.
Roper, C., Grey, F., Cadogan, E., 2018. Co-production: Putting principles into practice in mental health contexts. Melbourne School of Health Sciences, University of Melbourne. https://healthsciences.unimelb.edu.au/_data/assets/pdf_file/0007/3392215/Coproduction_putting-principles-into-practice.pdf.
Ruble, R., 2017. Control groups. In: Allen, M. (Ed.), The Sage Encyclopedia of Communication Research Methods. SAGE Publications, pp. 253–254. https://doi.org/10.4135/9781483381411.n94.
Sanders, E.B.N., Stappers, P.J., 2008. Co-creation and the new landscapes of design. CoDesign 4 (1), 5–18. https://doi.org/10.1080/15710880701875068.
Schulz, K.F., Altman, D.G., Moher, D., 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. Trials 11 (1). https://doi.org/10.1186/1745-6215-11-32.
Schünemann, H., Brożek, J., Guyatt, G., Oxman, A. (Eds.), 2013. GRADE Handbook for Grading Quality of Evidence and Strength of Recommendations. The GRADE Working Group https://gdt.gradepro.org/app/handbook/handbook.html.
Schünemann, H.J., Higgins, J.P.T., Vist, G.E., Glasziou, P., Akl, E.A., Skoetz, N., Guyatt, G.H., 2022. Chapter 14: completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions version 6.3. Cochrane. http://www.training.cochrane.org/handbook.
Schwarzer, G., 2022. Package 'meta': general package for meta-analysis. R package version 5.5–0. CRAN repository. https://cran.r-project.org/web/packages/meta/meta.pdf.
Shadish, W.R., Cook, T.D., Campbell, D.T., Chelimsky, E., Learning, Cengage, 2002. Experimental and Quasi-experimental Designs for Generalized Causal Inference. Houghton Mifflin.
Sin, J., Henderson, C., Woodham, L.A., Sesé Hernández, A., Gillard, S., 2019. A multicomponent eHealth intervention for family carers for people affected by psychosis: a coproduced design and build study. J. Med. Internet Res. 21 (8), e14374 https://doi.org/10.2196/14374.
Slattery, P., Saeri, A.K., Bragge, P., 2020. Research co-design in health: a rapid overview of reviews. Health Res. Policy Syst. 18 (1) https://doi.org/10.1186/s12961-020-0528-9.
Spence, S.H., 1998. A measure of anxiety symptoms among children. Behav. Res. Ther. 36 (5), 545–566. https://doi.org/10.1016/s0005-7967(98)00034-5.
Spieth, P.M., Kubasch, A.S., Penzlin, A.I., Illigens, B.M.W., Barlinn, K., Siepmann, T., 2016. Randomized controlled trials – a matter of design. Neuropsychiatr. Dis. Treat. 12, 1341–1349. https://doi.org/10.2147/ndt.s101938.
Spitzer, R.L., Kroenke, K., Williams, J.B., 1999. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. JAMA 282 (18), 1737–1744. https://doi.org/10.1001/jama.282.18.1737.
Spitzer, R.L., Kroenke, K., Williams, J.B., Löwe, B., 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch. Intern. Med. 166 (10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092.
Stapelberg, N.J.C., Bowman, C., Woerwag-Mehta, S., Walker, S., Davies, A., Hughes, I., Michel, K., Pisani, A.R., van Engelen, H., Delos, M., Hageman, T., Fullerton-Smith, K., Krishnaiah, R., McDowell, S., Cameron, A., Scales, T.L., Dillon, C., Gigante, T., Heddle, C., Turner, K., 2021. A lived experience co-designed study protocol for a randomised control trial: the attempted suicide short intervention program (ASSIP) or brief cognitive behavioural therapy as additional interventions after a suicide attempt compared to a standard suicide prevention pathway (SPP). Trials 22 (1). https://doi.org/10.1186/s13063-021-05658-y.
Stephens, L., 2013. Co-production in mental health - why everybody wins. Mind. https://www.mind.org.uk/information-support/your-stories/co-production-in-mental-health-why-everybody-wins/.
Sterne, J.A.C., Harbord, R.M., 2004. Funnel plots in meta-analysis. Stata J. Promot. Commun. Stat. Stata. 4 (2), 127–141. https://doi.org/10.1177/1536867x0400400204.
Sterne, J.A.C., Savović, J., Page, M.J., Elbers, R.G., Blencowe, N.S., Boutron, I., Cates, C. J., Cheng, H.Y., Corbett, M.S., Eldridge, S.M., Emberson, J.R., Hernán, M.A., Hopewell, S., Hróbjartsson, A., Junqueira, D.R., Jüni, P., Kirkham, J.J., Lasserson, T., Li, T., Higgins, J.P.T., 2019. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ, l4898. https://doi.org/10.1136/bmj.l4898.
Sterne, J.A.C., Sutton, A.J., Ioannidis, J.P.A., Terrin, N., Jones, D.R., Lau, J., Carpenter, J., Rucker, G., Harbord, R.M., Schmid, C.H., Tetzlaff, J., Deeks, J.J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D.G., Moher, D., Higgins, J.P.T., 2011. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 343 (jul22 1), d4002. https://doi.org/10.1136/bmj.d4002.
Sterne, J.A., Egger, M., 2001. Funnel plots for detecting bias in meta-analysis. J. Clin. Epidemiol. 54 (10), 1046–1055. https://doi.org/10.1016/s0895-4356(01)00377-8.
Thabrew, H., Fleming, T., Hetrick, S., Merry, S., 2018. Co-design of eHealth interventions with children and young people. Front. Psych. 9 https://doi.org/10.3389/fpsyt.2018.00481.
Thomas, J., Kneale, D., McKenzie, J.E., Brennan, S.E., Bhaumik, S., 2022. Chapter 2: determining the scope of the review and the questions it will address. In: Higgins, J. P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions version 6.3. Cochrane. http://www.training.cochrane.org/handbook.
Tindall, R.M., Ferris, M., Townsend, M., Boschert, G., Moylan, S., 2021. A first-hand experience of co-design in mental health service design: opportunities, challenges, and lessons. Int. J. Ment. Health Nurs. 30 (6), 1693–1702. https://doi.org/10.1111/inm.12925.
Turner, R.M., Bird, S.M., Higgins, J.P.T., 2013. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. PloS One 8 (3), e59202. https://doi.org/10.1371/journal.pone.0059202.

van Aert, R.C.M., Wicherts, J.M., van Assen, M.A.L.M., 2019. Publication bias examined in meta-analyses from psychology and medicine: a meta-meta-analysis. PloS One 14 (4), e0215052. https://doi.org/10.1371/journal.pone.0215052.

Vanstone, C., 2021. The realities of co-designing a better mental health system. The Australian Centre for Social Innovation (TACSI). https://www.tacsi.org.au/news-ideas/realities-of-co-designing-mental-health-system.

Vargas, C., Whelan, J., Brimblecombe, J., Allender, S., 2022. Co-creation, co-design, co-production for public health – a perspective on definition and distinctions. Public Health Res. Pract. 32 (2) https://doi.org/10.17061/phrp3222211.

VicHealth, 2022. How to co-design with young Victorians. https://www.vichealth.vic.gov.au/media-and-resources/publications/co-design. (Accessed 28 February 2022).

Viksveen, P., Bjønness, S.E., Cardenas, N.E., Game, J.R., Berg, S.H., Salamonsen, A., Storm, M., Aase, K., 2021. User involvement in adolescents' mental healthcare: a systematic review. Eur. Child Adolesc. Psychiatry. https://doi.org/10.1007/s00787-021-01818-2.

von Hippel, P.T., 2015. The heterogeneity statistic I2 can be biased in small meta-analyses. BMC Med. Res. Methodol. 15 (1) https://doi.org/10.1186/s12874-015-0024-z.

Voorberg, W.H., Bekkers, V.J.J.M., Tummers, L.G., 2014. A systematic review of co-creation and co-production: embarking on the social innovation journey. Public Manag. Rev. 17 (9), 1333–1357. https://doi.org/10.1080/14719037.2014.930505.

Waffenschmidt, S., Knelangen, M., Sieben, W., Bühn, S., Pieper, D., 2019. Single screening versus conventional double screening for study selection in systematic reviews: a methodological systematic review. BMC Med. Res. Methodol. 19 (1) https://doi.org/10.1186/s12874-019-0782-0.

WHO, 2013. Mental health action plan 2013–2020. World Health Organization. https://www.who.int/publications/i/item/9789241506021.

WHO, 2017. Depression and Other Common Mental Disorders: Global Health Estimates. World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf.

WHO, 2022a. Mental health: strengthening our response. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response.

WHO, 2022b. COVID-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide: wake-up call to all countries to step up mental health services and support. World Health Organization. https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide.

Williamson, V., Larkin, M., Reardon, T., Pearcey, S., Hill, C., Stallard, P., Spence, S.H., Breen, M., Macdonald, I., Ukoumunne, O., Ford, T., Violato, M., Sniehotta, F., Stainer, J., Gray, A., Brown, P., Sancho, M., Creswell, C., 2021. Codesign and development of a primary school based pathway for child anxiety screening and intervention delivery: a protocol, mixed-methods feasibility study. BMJ Open 11 (4), e044852. https://doi.org/10.1136/bmjopen-2020-044852.

World Bank, 2022. Population, female (% of total population). https://data.worldbank.org/indicator/SP.POP.TOTL.FE.ZS.

Zwetsloot, P.P., van der Naald, M., Sena, E.S., Howells, D.W., IntHout, J., de Groot, J.A., Chamuleau, S.A., MacLeod, M.R., Wever, K.E., 2017. Standardized mean differences cause funnel plot distortion in publication bias assessments. eLife 6. https://doi.org/10.7554/elife.24260.