Comprehensive Psychiatry 122 (2023) 152372

Contents lists available at ScienceDirect

Comprehensive Psychiatry

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

Psychological therapies for people with intellectual disabilities: An updated systematic review and meta-analysis

Katherine Tapp [a], Leen Vereenooghe [b], Olivia Hewitt [a,c], Emma Scripps [a], Kylie M. Gray [a,f,g], Peter E. Langdon [a,d,e,g,*]

[a] Centre for Educational Development, Appraisal and Research (CEDAR), University of Warwick, Coventry CV4 7AL, United Kingdom
[b] v. Bodelschwinghsche Stiftungen Bethel, Psychologischer Dienst, proWerk, Nazarethweg 4, 33617 Bielefeld, Germany
[c] Berkshire Healthcare NHS Foundation Trust, Learning Disabilities Service, Erlegh House, Earley Gate, Whiteknights Road, Reading, Berkshire RG6 6BZ, United Kingdom
[d] Brooklands Hospital, Coventry and Warwickshire Partnership NHS Trust, Birmingham B37 5RY, United Kingdom
[e] Herefordshire and Worcestershire Health and Care NHS Trust, 2 Kings Way, Charles Hastings Way, Worcester WE5 1JR, United Kingdom
[f] Department of Psychiatry, School of Clinical Health Sciences at Monash Health, Monash University, Clayton, VIC 3168, Australia
[g] Centre for Mental Health and Wellbeing Research, University of Warwick, Coventry CV4 7AL, United Kingdom

ARTICLE INFO

Keywords:
Learning disabilities
Neurodevelopmental disorders
Psychotherapy
Effect size
Effectiveness

ABSTRACT

Objective: The aim of this systematic review and meta-analysis (PROSPERO 2020 CRD42020169323) was to evaluate the efficacy of psychological therapy for people with intellectual disabilities.
Method: A comprehensive literature search yielded 22,444 studies which were screened for eligibility. Studies were eligible for inclusion if a psychological therapy was delivered to people with intellectual disabilities compared to a group who did not receive the therapy. Thirty-three controlled trials were eligible for inclusion in the review, with 19 included within a DerSimonian-Laird random effects meta-analysis. Subgroup analysis was completed by clinical presentation, and by comparing randomised trials to non-randomised trials, and group-based to individually delivered psychotherapy.
Results: Following the removal of outliers, psychological therapy for a range of mental health problems was associated with a small and significant effect size, g = 0.43, 95% CI [0.20, 0.67], $N = 698$. There was evidence of heterogeneity and bias due to studies with small sample sizes and a lack of randomisation. Non-randomised studies were associated with a large effect size, g = 0.90, 95% CI [0.47, 1.32], $N = 174$, while randomised studies were associated with a small effect size, g = 0.36, 95% CI [0.17, 0.55], $N = 438$, excluding outliers. Individually delivered psychological therapy was associated with a small and non-significant effect size, g = 0.32, 95% CI [−0.01, 0.65], $N = 146$, while group-based interventions were associated with a small and significant effect size, g = 0.37, 95% CI [0.05, 0.68], $N = 361$, again, excluding outliers. Psychological therapy for anger was associated with a moderate effect size, g = 0.60, 95% CI [0.26, 0.93], $N = 324$, while treatment for depression and anxiety was associated with a small and non-significant effect size, g = 0.38, 95% CI [−0.10, 0.85], $N = 216$, after outliers were removed.
Conclusions: Studies are fraught with methodological weaknesses limiting the ability to make firm conclusions about the effectiveness of psychological therapy for people with intellectual disabilities. Improved reporting standards, appropriately powered and well-designed trials, and greater consideration of the nature and degree of adaptations to therapy are needed to minimise bias and increase the certainty of conclusions.

1. Introduction

Mental illness occurs more frequently amongst people with intellectual disabilities relative to the general population [13], which can be attributed to a variety of biological, psychological, and social factors. These include stressful life events [11,28,65], genetics [7], degree of intellectual disability, primary care utilisation, and factors relating to physical health [11]. Systemic issues within services for people with intellectual disabilities, including staff changes and reduced care associated with financial strain, are also relevant [4], along with barriers

* Corresponding author at: Centre for Educational Development, Appraisal and Research (CEDAR), University of Warwick, Coventry CV4 7AL, United Kingdom.
E-mail address: Peter.Langdon@warwick.ac.uk (P.E. Langdon).

https://doi.org/10.1016/j.comppsych.2023.152372

Available online 26 January 2023

inherent to accessing mental health care [70].

Evidence suggests that psychotherapy is an effective intervention for mental illness in the general population [48], but there is often reliance upon psychotropic medicine with people with intellectual disabilities, which has been criticised [46]. Accessibility of psychotherapy for people with intellectual disabilities is affected by barriers such as diagnostic overshadowing, therapeutic disdain, and concerns regarding cognitive impairments limiting beneficial effects [25].

Diagnostic overshadowing, which refers to the attribution of symptoms of mental illness to symptoms of intellectual disabilities ([53]; [79]), can lead to a decreased likelihood of referral to therapeutic services. Also, staff in mental health services have cited feelings of inadequacy or intimidation in working with those with intellectual disabilities [57], creating therapeutic disdain. This relates to concerns around the perception that those with cognitive impairments experience decreased efficacy of psychotherapy. However, evidence suggests that people with intellectual disabilities do have some of the necessary skills for therapies such as cognitive behavioural therapy (CBT), or these skills can be taught [12,31,66,68].

The National Institute for Health and Social Care Excellence [50] have recommended that CBT, relaxation therapy, and graded exposure therapy should be considered for use with people with intellectual disabilities. NICE acknowledged that evidence supporting such interventions was rated as low or very low quality. Many of the studies designed to explore the use of psychotherapy with people with intellectual disabilities were not powered to detect an effect and had a large disparity in sampling strategies and methods, including the degree of intellectual disabilities amongst participants and length of treatment provided. A lack of reporting details about the intervention delivered, including the nature and degree of adaptations, may have contributed to these disparities [67].

Systematic reviews and meta-analyses have attempted to determine the effectiveness of psychotherapies for people with intellectual disabilities. Hamers, Festen and Hermans [22] reviewed interventions for depression and noted that CBT was more widely studied than behavioural therapy, exercise interventions, social problem-solving skills training, and bright-light therapy. The authors acknowledged that many of the included studies were of low quality. Koslowski et al. [34] completed a meta-analysis about interventions for mental illness amongst people with intellectual disabilities, including medication, psychotherapy, and system-level interventions. Their inclusion criteria stated that only studies reporting the ranges of participants' general intellectual functioning were included. They reported that interventions had no significant effect upon behaviour problems, depression, anxiety, quality of life, or functioning.

In a recent systematic review on psychotherapy for people with intellectual disabilities by Witwer et al. [75], randomised controlled trials were the minority study design despite being of higher quality than other study designs, highlighting methodological weaknesses in this area. The authors acknowledged that more RCTs had been conducted since 2016, which is promising. Additionally, Graser et al. [20] recently completed a meta-analysis focusing solely upon CBT with people with intellectual disabilities, or studies using components of CBT. The authors concluded that CBT was associated with a moderate effect size for the treatment of depression and anger but had a non-significant effect upon anxiety.

Recent systematic reviews and meta-analyses have limited the inclusion of studies by year of publication or by reporting of intellectual functioning or have restricted focus to a specific type of psychotherapy. To date, the only meta-analysis that appears to examine psychotherapy broadly with people with intellectual disabilities was conducted by Vereenooghe and Langdon [67].

Consequently, the current systematic review and meta-analysis aimed to update the work of Vereenooghe and Langdon [67]. This updated review will follow the same aims as the original review, which were to "*(a) identify and evaluate controlled outcome studies of psychological therapies with people with intellectual disabilities, excluding approaches such as applied behavioural analysis, (b) conduct meta-analyses to determine overall efficacy of treatment, as well as the efficacy of various psychotherapies for different mental health problems where possible, and (c) identify areas with limited available evidence to suggest directions for future research.*" ([67], p. 4087). No exclusions based on year of publication, or reporting of general intellectual functioning, or type of psychotherapy used were made.

2. Methods

2.1. Eligibility criteria

2.1.1. Inclusion criteria

Studies were included if: (a) study participants had an intellectual disability evidenced by data to indicate Full-Scale IQ <70, or other evidence to indicate an intellectual disability (i.e., an assessment of adaptive behaviour focused upon conceptual, social, and practical daily living skills). Studies that also included participants with borderline intellectual disabilities (IQ score between 69 and 79) alongside people with mild to profound intellectual disabilities were not excluded, (b) the participants were older than 5 years, (c) contained at least one intervention that was a form of psychotherapy, defined by Norcross [51] as "*the systematic application of interventions based on well-established psychological principles and techniques aimed at the prevention or treatment of emotional, behavioural or mental health problems*", and (d) at least two independent groups were included, where one group received a psychotherapy which was compared to a control group who did not receive a psychotherapy and, (e) they were published in English within a peer reviewed journal.

2.1.2. Exclusion criteria

Studies were excluded if they used strict behavioural interventions, or non-traditional psychotherapies. Behavioural interventions were not included as they are arguably not examples of psychotherapy in this context [62].

The protocol was registered with the international prospective register of systematic reviews (PROSPERO 2020 CRD42020169323).

2.2. Information sources

2.2.1. Databases

PsycINFO and MEDLINE were searched on 07/01/2022 through Ovid. CINAHL was searched on the same date through EBSCOhost. Google Scholar, Cochrane Library and British ETHOS searches were also run on the same day.

2.2.2. Citation searching

Backward searching of references for eligible papers identified from previous information sources were screened and a further 10 papers were identified. Forwards searching was conducted by reviewing papers citing eligible studies on Google Scholar, but no new papers were identified. Studies cited in NICE [50] guideline number 54 were screened and a further 2 eligible studies were identified. References from literature reviews examining similar topics were also examined and one eligible study was found within the review by Witwer et al. [75].

All searches were repeated on 26/06/2022 and no new studies were identified.

2.3. Search strategy

The search terms for Ovid, EBSCOhost, and Web of Science covered different terms for intellectual disabilities, specific conditions related to intellectual disability, and different terms for psychotherapy.

The search strategy was produced by the research team and was

more exhaustive compared to the original review [67] to ensure no manuscripts were missing. No restrictions were placed on searches. The full list of search terms can be found in the Supplementary Materials.

For Google Scholar, Cochrane Library and British ETHOS searches the search terms were "intellectual disability" OR "learning disability" AND "psychotherapy".

2.4. Selection process

Results from database searches were initially exported into EndNote and then exported into Rayyan [52], a software for collaborative systematic literature reviews. Overall, 23,119 papers were identified with 683 duplicates found using EndNote and 1305 found using Rayyan. After removal of duplicates 22,444 papers remained.

Title and abstract screening of 100% of the results from the database searches was conducted by KT with 10% of the results independently reviewed by a co-rater (ES). This process was conducted in a blinded collaborative project using Rayyan. The project was then unblinded to identify conflicts. Agreement was near perfect, $k = 0.99$. Disagreements were discussed and full agreement was achieved.

Full text review was conducted by KT on 100% of the selected references and 78% were independently reviewed by a co-rater (OH). Inter-rater reliability was 90.01%, $k = 0.82$. All conflicts were resolved through discussion between KT and OH.

2.5. Data collection process

Data extraction was conducted by KT, with 40% of the papers randomly selected for review by an independent second rater (OH). Any missed extracted information was discussed, and full agreement was reached.

2.6. Data items

Data extraction forms were based on the Cochrane template for RCTs, with changes relevant to the review and meta-analysis. Details on the following were collected:

2.6.1. Study eligibility

Confirmation that participants had an intellectual disability diagnosis, were aged >5 years old, at least one intervention could be defined as psychotherapy, there were two or more independent groups.

2.6.2. Study design

Number of intervention/control groups, type of intervention, type of control condition.

2.6.3. Participants

Overall number of participants, number by allocation to intervention group, average age, sex, average IQ, level of intellectual disability, measure of intellectual disability used, behavioural/emotional/mental health problem targeted, country of study.

2.6.4. Intervention

Description of intervention, length of intervention, length of sessions, number of participants per group, location/setting of intervention.

2.6.5. Outcome

Measurement used, method of data analysis, results of main outcome, means and standard deviations of main outcome at pre-treatment and post-treatment.

2.6.6. Follow-up

Timepoint of follow-up, observed differences from post-treatment, means and standard deviations of main outcome measures.

As a large proportion of studies contained multiple outcome measures, the selected data for meta-analysis input was based on the identified primary outcome or the most appropriate measure for the treatment target. This was usually stated by the authors but was not clear in every case and judgments had to be made by KT which were discussed with PEL until consensus was reached. Only results at post-treatment were included in the meta-analysis.

2.7. Risk of bias

The NICE quality appraisal checklist for quantitative studies was used for the quality appraisal process, which contains critical appraisal items [49]. Adaptations were made for suitability for the present review. Reporting of a measure of intellectual functioning was considered an important variable that needed to be described. Outcome measures were only considered reliable if there was evidence of effective use with people with intellectual disabilities. Quality appraisal was conducted separately by KT and LV on 100% of the included papers, disagreements not resolved through discussion were brought to PEL and KMG until 100% agreement on all papers was achieved.

2.7.1. External validity

Source populations were considered well described when characteristics essential to the study were reported. This typically included a measure of intellectual disability alongside a measure of the targeted emotional, behavioural, or mental health problem. The eligible population and selected participants were considered representative if participant characteristics were balanced, different degrees of intellectual disability were considered, and there was low risk of selection bias. It was also important that the eligible population was selected across multiple centres to avoid overrepresentation from a single centre.

2.7.2. Internal validity

Studies were considered internally valid if they: (a) used appropriate randomisation, (b) described interventions well, (c) concealed allocation, (d) ensured blinding of participants and assessors, (e) allowed for adequate exposure of intervention based on a fidelity check, (f) contamination between experimental and control groups was low, (g) showed a similarity in baseline characteristics and exposure to other interventions between groups, (8) had no missing data, (9) used outcome measures appropriate for intellectual disabilities and assessed all important outcomes and potential harms, (10) conducted follow-up at a meaningful time, (11) used intention to treat analysis, particularly where there was high drop-out, (12) were sufficiently powered, and (13) used appropriate analytical methods reported in enough detail to calculate effect size estimates and precision of intervention effects.

2.8. Effect measures

The standardised mean difference between intervention and control groups (Hedge's g) was calculated using RevMan 5.4.1 [54] using the DerSimonian-Laird random effects method as some studies had small sample sizes. Effect sizes were determined as small ≥0.20, medium ≥0.50, or large ≥0.80 according to Cohen's [9] interpretation of effect size thresholds.

2.9. Synthesis methods

Studies were included in the meta-analysis if they had at least two groups where one received an active intervention and the other acted as a control group (e.g., waiting list or treatment-as-usual, but not another psychotherapy as far as could be determined) and reported data that could be extracted and included in the meta-analysis (i.e., means and standard deviations of the data).

The meta-analysis was completed in RevMan 5.4.1 with results visualised through Forest plots, ordered by year of publication. A

random effects model was selected as studies examined a range of behaviours and disorders and used different outcome measures.

Planned subgroup analyses to further investigate heterogeneity included randomised vs. non-randomised studies, individual vs. group-based interventions, and clinical presentation (anger, depression, interpersonal functioning, and trauma). Each study was only included in one subgroup per comparison, to remain consistent with Vereenooghe and Langdon's [67] original meta-analysis. We pooled studies that treated both anxiety and depression as some of the included studies attempted to treat both anxiety and depression (e.g., [38]).

The data for Rose, Dodd and Rose [55] and Rose, O'Brien, and Rose [58] used the same control data, making the data non-independent [8]. In such cases, aggregation of the data has been identified as the optimal solution to avoid duplication bias and exclusion bias [76]. Therefore, the Rose studies were combined using the control group results reported by Rose, O'brien, and Rose [58] and the average of the aggregated means and standard deviations of both intervention groups.

2.10. Additional analyses

2.10.1. Heterogeneity

The I^2 statistic was reported for every analysis to examine the proportion of variance in observed effects that can be attributed to variation in true effects relative to sampling error. This was interpreted with the assumption that any results over 40% are problematic [14]. Confidence intervals based on chi-squared distribution (x^2) for I^2 was also reported.

To avoid complete reliance on I^2, prediction intervals were also calculated to estimate heterogeneity in our estimate of true effect sizes.

A Galbraith plot was constructed using SPSS Version 28.0.1.1 to examine which studies may be unduly contributing to our estimate of effect size which can be found in our Supplementary Material. Studies presented as outliers were removed and the analysis re-run.

2.10.2. Reporting bias

To examine the impact of potential publication bias, funnel plots were generated and inspected for asymmetry using SPSS Version 28.0.1.1 for our main analysis. The Egger, Davey Smith, Schneider and Minder [18] regression was calculated to test the significance of any asymmetry. A Trim and Fill analysis [17] was completed to estimate the number of missing studies and determine the impact these may have upon the effect size estimate. File drawer analysis using the Rosenthal method [59] was conducted to determine the number of new studies with null results that would have to be published to generate a non-significant effect size estimate.

3. Results

3.1. Study selection

Sixty studies were eligible for full text review and 30 were selected for inclusion in the review. Fig. 1 describes this process in further detail. One of the main reasons for excluding studies at this stage was that participants did not have intellectual disabilities, and instead had learning difficulties (such as specific reading and writing disorders) (e.g., [60]). Further, not all interventions fit our definition of psychotherapy, for example, Lindsay and Baty [36] focused exclusively on relaxation techniques. One thesis was identified as using psychoanalytic psychotherapy, but targeted improvement in intellectual functioning rather than emotional, behavioural, or mental health problems [3]. Other reasons included identification of study protocols or conference abstracts as opposed to full studies, no inclusion of a control or comparison group, duplicate publications, and targeting challenging behaviour only.

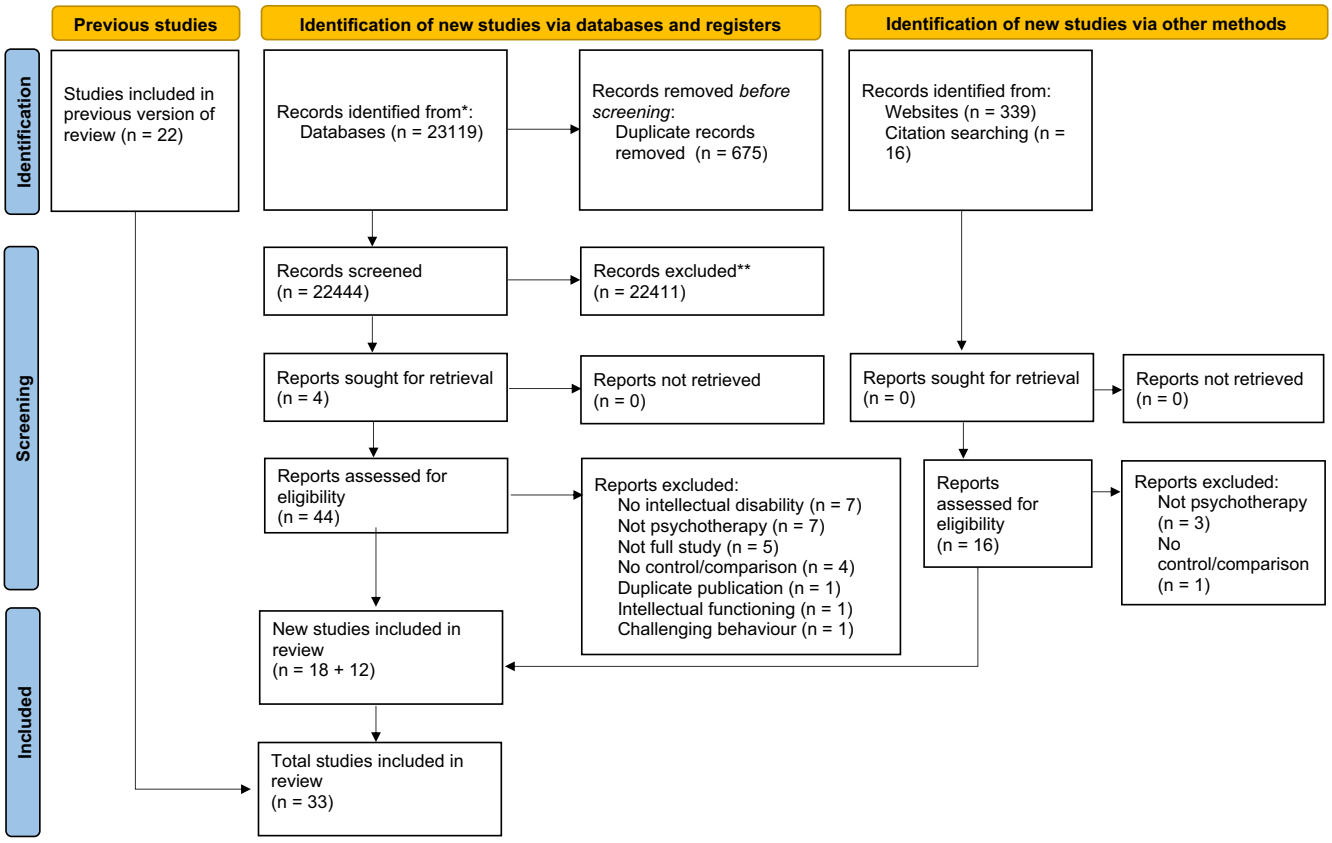

Fig. 1. Flow diagram of search and selection process.

3.2. Study characteristics

Data extraction allowed for observations of consistencies and inconsistencies between studies. Details from the data extraction process are presented in Table 1.

3.2.1. Population

There was an inconsistency in the reporting of the degree of intellectual disabilities amongst participants, with some studies providing a measure of intellectual functioning and others only reporting degree of intellectual disability (e.g., mild or moderate). One study included IQ scores from the recruitment centre's records but relied on the diagnosis of an intellectual disability where this was unavailable [10]. Most studies focused on participants with mild intellectual disabilities, with only two having reported using interventions with people with severe intellectual disabilities [16,21]. Some studies included participants with borderline intellectual disabilities, which can include those with an IQ score between 69 and 79 [21,33,39,43,55,61]. Within a selection of studies, receptive vocabulary ([2,16,21,55,56,72]; [58]; [23,24]), non-verbal reasoning [21], or adaptive behaviour [23] was measured, or authors used an abbreviated assessment of IQ, such as the abbreviated battery IQ from the Stanford-Binet Intelligence Scales [23].

A few studies reported participant race or ethnicity [2,23,29,39], mostly reporting that most of their sample was white. Only three studies delivered interventions to children or young adolescents ([39]; [77]; [33]).

3.2.2. Outcomes

The Glasgow Anxiety Scale – Intellectual Disabilities, and the Glasgow Depression Scale – Learning Disabilities were frequently used [10,23,29,32,38]. The Beck Depression Inventory-II was also used often [42,47], with one study using the Beck Depression Inventory – Youth scale [24], and has been validated for use with participants with intellectual disabilities [37].

3.2.3. Adaptations

The most common method for making interventions accessible for people with intellectual disabilities was to have a support worker, psychologist, or caregiver present ([55,73]; [58]; [23]; [10,29,38,45]). Use of visual aids were also widely reported [10,21,43,55,74]. Some authors reported provision of training to those delivering the interventions in working with people with intellectual disabilities [16,24,44,47].

3.3. Quality appraisal

3.3.1. Internal validity

The main reasons for lower internal validity were: (1) recruitment from <3 centres ([21,43,55,61]; [58]; [77]; [10]), (2) recruitment through staff referral [41,72,73], and (3) an overrepresentation of participant characteristics not related to the study, such as male only samples [63] or a high prevalence of unrelated mental health disorders [40].

3.3.2. External validity

The most frequent reasons for lower external validity included: (1) no reporting or use of allocation concealment, with a minority of studies reporting this process [16,21,24,29,32,47,63,74], (2) no reporting/use of blinding, with a minority of studies reporting this process ([24,41,47,63]; [77]; [29,32,44,74]) or partially implementing blinding ([2,61]; [38]), (3) study was underpowered, with only 11 studies appearing to be sufficiently powered ([24,55,56,58]; [77]; [10,29,33,44,45,74]), (4) randomisation not used ([2,40,43,61]; [38]; [55,56,64,72]; [58,77]; [23,38]), (5) no fidelity check, with only some studies including one [23,24,29,32,40,45,61,72,74] or provided supervision during the intervention [55,64] and, (6) harms not discussed, with only 4 studies mentioning consideration of harms [10,29,43,74].

The overall quality ratings and summary of discussions for each study can be found in Table 2.

3.4. Results of synthesis

Many studies utilised more than one outcome measure. The primary outcome used within our meta-analysis for each study was that identified by the authors of the respective studies. Where authors did not identify their primary outcome, a decision was made to include the outcome measure that was most closely related to the target of the intervention. In cases where this was not clear, we chose to use the outcome which most closely reflected outcome measures used by other studies included within our analysis. For example, depression measures were selected in cases where studies measured anxiety and depression separately as primary outcomes [10,24].

3.4.1. All psychotherapy studies

An initial meta-analysis of all included studies, displayed in Fig. 2, resulted in a moderate effect size of psychotherapy, g = 0.69, 95% CI [0.37, 1.00], $N = 837$, Fig. 2. There was evidence of inconsistency in findings across studies, $I^2 = 76\%$, $x^2\ (17) = 71.42$, $p < 0.001$, while the distribution of true effect sizes varied widely, as indicated by the prediction interval [−0.58, 1.95].

Inspection of the funnel plot revealed significant asymmetry, $t = 3.22$, $p = .005$, indicating that small studies had an undue impact upon the estimate of effect size, Fig. 3. However, the trim and fill analysis did not generate any missing studies, leading to a change in the estimate of effect size. Rosenthal file-drawer analysis found that 358 studies with null results would have to be reported for results of the meta-analysis to become non-significant.

A Galbraith plot, Supplementary Materials: Appendix 2, was constructed to visualise possible outliers and Hartley, et al. [23] was initially removed, which decreased the overall effect size to g = 0.57, 95% CI [0.31, 0.84], $N = 813$, $I^2 = 67\%$, $x^2\ (16) = 48.00$, $p < 0.001$, and the prediction interval [−0.41, 1.56] decreased but continued to cross zero. Following this, McCabe et al. [42] was removed which decreased the overall effect size to g = 0.51, 95% CI [0.25, 0.76], $N = 764$, $I^2 = 61\%$, $x^2\ (15) = 38.38$, $p < 0.001$, and the prediction interval [−0.37, 1.38] again decreased but continued to cross zero. Rose et al. [55] and Rose et al. [58] were then removed which decreased the overall effect size further to small, g = 0.43, 95% CI [0.20, 0.67], $N = 698$, $I^2 = 50\%$, $x^2\ (14) = 28.04$, $p = 0.01$, and the prediction interval [−0.29, 1.16] further decreased but continued to cross zero. Visual inspection of the funnel plot following the removal of studies suggested asymmetry, while Egger's regression test was non-significant, $t = 1.39$, $p = .19$. The trim and fill analysis led to the generation of 2 missing studies, and a reduction in the estimate of effect size which remained small, g = 0.34, 95% CI [0.09, 0.59], Fig. 4.

3.4.2. Sub-group analysis: Randomised vs. non-randomised studies

Examining randomised trials alone revealed a small effect size, Fig. 5, g = 0.43, 95% CI [0.14, 0.72], $N = 551$, with an I^2 of 58%, $x^2\ (10) = 24.03$, $p = .008$, and wide a prediction interval, [−0.45, 1.31], Fig. 5. A small majority of participants overall fell into this subgroup with a weight of 64.3%. Visualisation using a Galbraith plot, Supplementary Materials: Appendix 3, identified McCabe, McGillivray, and Newton [42] as an outlier with removal reducing effect size, g = 0.30, 95% CI [0.08, 0.52], $N = 502$, and I^2 to 22%, $x^2\ (9) = 11.54$, $p = 0.24$, and the prediction interval [−0.15, 0.75]. Klein et al. [33] was then identified as an outlier and removed resulting in an increase in effect size, g = 0.36, 95% CI [0.17, 0.55], $N = 438$, $I^2 = 0\%$, $x^2\ (8) = 7.51$, $p = 0.48$, and a prediction interval that did not overlap with zero [0.13, 0.59].

Non-randomised studies resulted in a large effect size, g = 1.20, 95% CI [0.50, 1.89], $N = 286$, which was just significantly larger than the randomised subgroup, $x^2\ (1) = 3.96$, $p = .05$. I^2 was 83%, $x^2\ (6) = 35.20$, $p < 0.001$, and the prediction interval was [−1.11, 3.50]. Visualisation

Table 1
Characteristics of studies eligible for systematic review.

First author (year)	Study design	Participants	Intervention	Outcome	Follow-up
Benson, Rice, and Miranti [2] [a]	Four groups EG1: relaxation training EG2: self-instruction EG3: problem solving EG4: combined anger management	*N* = 54 Avg. age = 32 37 M/17F IQ not reported Receptive vocabulary (PPVT) Mild/moderate intellectual disability Target: Anger/self-control Location: USA	12 weeks 90 min sessions EG1: tension release EG2: discrimination between coping statements and trouble statements EG3: 4-step plan EG4: all 3 interventions Group size: 5–9 per group Setting: vocational centres	Reduction in aggressive gestures in role-play maintained at post-hoc. Improvements in voice loudness, tone, and conflict situations test not maintained at post-hoc	4–5 weeks Effects maintained
Cooney et al. [10]	Two groups EG: computerised CBT CG: TAU	EG: n = 24, avg. age = 42, 8 M/16F, avg. IQ (only available for 10 participants) 54 CG: *n* = 25, avg. age = 39.24, 11 M/14F, avg. IQ (only available for 8 participants) 55.63 Target: Anxiety/depression Location: Ireland	7 weeks 1 h sessions CBT via social stories on computer with between-therapy skills workbook Group size: Individual Setting: house or day-centre	**Glasgow Depression Scale for people with a Learning Disability** Medium effect size on anxiety only No significant differences for depression	3 months Large effect size on anxiety
Dowling et al. [16] [a]	Two groups EG1: traditional counselling EG2: integrated intervention	EG1: N = 24, 11 M/12F EG2: *N* = 32, 3 M/8F (not all data available) Age 18+ Mild/moderate/severe ID Receptive vocabulary (PPVT) Target: Bereavement Location: UK	EG1 around 15 weekly/fortnightly sessions, 1 h. Intervention not described. EG2 duration not reported. Loss-oriented grief work and restoration-oriented work. Group size: Individual Location: at home, care home, or day-centre	Significant improvement in EG1 No significant improvement in EG2, high rates of non-completion	None
Hagiliassis et al. [21]	Two groups EG: anger management CG: wait list	EG: N = 14, avg. age = 44.93, 7F/7M, avg. PPVT score = 60 CG: *N* = 15, avg. age = 43.57, 9 M/6F, avg. PPVT score = 56.77 Target: Anger Location: Australia	12 weeks 2 h Relaxation techniques, cognitive restructuring, problem solving and assertiveness skills training 7 per group	**Novaco Anger Scale (NAS)** EG reduced anger on NAS, significant at treatment condition x time of assessment interaction Non-significant improvement on outcome rating scale for EG	4 months Effects maintained
Hartley, et al. [23]	Two groups EG: CBT CG: TAU	EG: *N* = 16, avg. age = 38.81, 8 M/8F, avg. IQ 62.38 CG: *N* = 8, avg. age = 40.25, 5 M/3F, avg. IQ 61.13 Target: Depressive disorder Location: USA	10 weeks 1.4 h sessions CBT intervention not detailed Group size: 5–6 per group Setting: group home	**Self-Report Depression Questionnaire** EG lower depressive scores No significant differences for social skills	3 months Effects maintained
Hassiotis, et al. [24]	Two groups EG: manualised-individualised CBT CG: TAU	N = 32 (EG = 16; CG = 16) Avg. age: EG = 33.7; CG = 38 EG: 5 M/11F; CG = 7 M/9F IQ not reported Mild intellectual disabilities Target: Common mood disorders Location: UK	16 weeks 1 h sessions Initial preparatory phase of 5 sessions, middle phase of therapy and end phase summary of techniques. Group size: Individual Setting: Community ID services sites	**Beck Depression Inventory - Youth** Non-significant improvements in favour of CG. Non-significant improvement in EG in depressive symptoms	6 months Effects maintained
[77] [a]	Two groups EG: trauma-focused CBT CG: TAU	*N* = 87 (EG: 42; CG: 45) Age range: 12–17 EG: 33 M/9F, IQ 58–69; CG: 38 M/7F, IQ 55–68 Target: Trauma Location: USA	Typically, 8 to 20 sessions but not reported Psychoeducation around PTSD and cognitive restructuring	Significant change in CG (except somatic subscale) Significant change in Aggressive Behaviour and Rule Breaking subscales for EG only	None
Jahoda, et al. [29] [a]	Two groups EG1: behavioural	EG1: *N* = 84, avg. age = 40.1, 38 M/39F, avg. IQ 58.34	EG1: 12 weeks, 1–2 h EG2: 8 weeks, 1–1.5 h EG1: Focus on purposeful regular	Significant improvements in depression for both groups	12 months Effects maintained

(*continued on next page*)

Table 1 (*continued*)

First author (year)	Study design	Participants	Intervention	Outcome	Follow-up
	activation EG2: guided self-help	EG2: *N* = 77, avg. age 40.3, 38 M/46F, avg. IQ 55.44 Mild/moderate intellectual disabilities Target: Depression Location: UK	activities; EG2: psychoeducation Group size: Individual At participant's home	No significant differences between groups	
Karatzias, et al. [32]	Two groups EG: eye movement desensitisation and reprocessing (EMDR) CG: TAU	n = 29 (EG: 15; CG: 14) Avg. age: 42 Mild to moderate intellectual disabilities Target: Trauma Location: UK	8 1-h sessions History taking, preparation (affect management and psychoeducation), assess components of distressing memory (self-referencing negative belief, desired positive belief, current emotional and physiological components of image and belief), bilateral stimulation whilst attention directed toward assessment phase, closure, re-evaluation Group size: Individual	**Post-Traumatic Stress Disorder Checklist** EG significant improvement 16 in EG diagnosis free (PTSD), 4 in CG	3 months Some became diagnosis free at follow-up but number not reported
Klein, et al. [33]	Two groups EG: cognitive bias modification CG: neutral control training	N = 69 (EG = 33; CG = 36) Avg. age: 14.4 Sex not reported IQ range 60–85 (based on school entry requirements) Mild/borderline intellectual disabilities Target: Social phobia Location: Netherlands	Ambiguous scenarios of 3 sentences – word in last sentence removed. Imagine self as central character – missing final word provided with one letter missing, fill in the letter. Comprehension question. Positive social situations for EG; neutral situations for CG Group size: Individual	**Social Anxiety Scale for Children - Revised** Significant time, and group x time interactions for all measures for EG – interpretation recognition task (IREC-T), social anxiety scale for children – revised (SASC-R)	10 weeks Continued improvement in positive group for performance on interpretation recognition task and social anxiety Improvement in CG on task
Lawrence [35] [a]	Two groups EG1: reality therapy EG2: support group	EG1: *N* = 16, avg. age = 39.56, 8 M/8F, avg. IQ 57.69 EG2: *N* = 14, avg. age = 44.50, 7 M/7F, avg. IQ 56.50 Mild/moderate intellectual disabilities Target: Self-determination Location: US	6 weeks 1 h sessions EG1: understanding needs and responsibility of behaviour, make plan to improve aspects of life EG2: share and receive ideas from group. Group size: 8 Setting: classroom of vocational workshop	Increase in self-regulation and self-realisation in EG1. Not sig diff in autonomy and psychological empowerment	None
[78]	Two groups EG: anger management CG: wait list	EG: *N* = 33, avg. age = 28.5, 25% F, avg. IQ 65.4 CG: N = 14, avg. age = 23.9, 42.85% F, IQ 66.2 Target: Anger Location: UK	40 weeks 40–60 min sessions Relaxation, stress inoculation, discussions about emotion/anger, problem solving, roleplay	**Dundee Provocation Inventory (DPI)** Significant improvement on DPI, reduced self-report anger, roleplay data for EG only No significant group differences on DPI	3, 9, 15 month follow-up, sparse data for 21, 27 and 30 month follow-up. Further reduction at 3 months, no further reduction after but remains under baseline
Lindsay, et al. [38]	Two groups EG: CBT CG: TAU	N = 24 (12 per group) 6 M/6F per group EG: avg. age = 29.9, avg. IQ 62.4 CG: avg. age = 33.1, avg. IQ 63.9 Mild intellectual disabilities Target: Emotional disorders Location: UK	Average 10.75 weeks role-play, discussion of real-life scenarios, goal setting, personalised handbook. Hospital and community residential settings	**Glasgow Severity Index (GSI)** EG better outcome on GSI, marginally significant better outcome for anxiety No effects on depression	3 and 6 months Effects maintained
Mann [39] [a]	Two groups EG: group counselling CG: study group sessions	N = 36 (18 per group) Age range: 9–13 All male IQ range: 56–80 Target: Self-concept Location: US	12 1-h sessions Counselling sessions in structured supportive environment Group size: 6	EG positive change on self-concept, manifest anxiety and behaviour rating Age and IQ not significant factors	None
Matson [40] [b]	Two groups EG: participant modelling CG: wait list	N = 24 (12 per group) Age range 25–45 12 M/12F Mild/moderate intellectual disabilities Target: Phobia Location: US	3 months (3× a week) 1 h Practiced grocery shopping scenarios before attempting real life scenario Avg. group size: 5 Setting: local mental health clinic and stores in community	Approach behaviour to feared object and grocery shopping skills measured. EG more effective at decreasing fear	4 months Reported data unclear
Matson & Senatore [41]	Three groups EG1: psychotherapy	N = 32 (EG1: 11; EG2: 11; CG: 10)	No. of sessions unclear 1-h sessions	**Nurses' Observation Scale for Inpatient Evaluation – 30**	

(*continued on next page*)

Table 1 (*continued*)

First author (year)	Study design	Participants	Intervention	Outcome	Follow-up
	EG2: social skills training CG: no treatment	Mean age: 34 21 M/11F Mild/moderate intellectual disabilities Target: Social skills Location: US	EG1: therapy groups. Discussion of workshop program. EG2: group discussion but focused on teaching target social skills behaviours. Setting: workshop and local clinic	Significant improvements in EG2, only improvements in role play for EG1	3 months Some decrease in effect of EG2 but not significant
McCabe, McGillivray, and Newton [42]	Two groups EG: CBT CG: wait list	EG: N = 19, avg. age = 34.05, 10 M/9F CG: N = 15, avg. age = 39.80, 6 M/9F Mild/moderate intellectual disabilities Target: Depression Location: Australia	5 weeks 2-h sessions groups sessions. Social skills, recreation, positive self-talk, problem solving, goal-setting, etc. 3–5 per group Setting: workplace	**Beck Depression Inventory** EG significant lower depression scores, higher social comparison, lower negative automatic thoughts. Self-esteem not reported.	3 months. No significant changes.
McGaw, Ball, and Clark [43]	Two groups EG: group intervention with home-based intervention CG: no group intervention but home-based intervention	EG: N = 12, avg. age = 29.1, 3 M/9F, avg. IQ 73.1 CG: *N* = 10, avg. age = 30.4, 4 M/6F, avg. IQ 71.7 Mild/borderline intellectual disabilities Target: Social awareness and understanding Location: UK	14 weeks 2-h sessions Semi-structured, cognitive behavioural methods. Setting: at home	**Judson Self Rating Scale – self-concept subscale** Self-concept became more positive with EG at follow-up, significant difference. CG increased significantly on judgment. EG did not significantly change in feelings about children or judgment of children's capabilities. No sig diff in either group on interactions with professionals and others	27 weeks Larger effect of self-concept in EG
McGillivray, McCabe, and Kershaw [47]	Two groups EG: Staff-administered community-based intervention CG: wait list	EG: N = 20, avg. age = 38.5, 13 M/7F CG: *N* = 29, avg. age = 31.15, 19 M/8F Mild intellectual disabilities Target: Depression Location: Australia	12 weeks 2-h sessions Group-based CBT 5–6 per group	**Beck Depression Inventory** EG group significantly fewer depressive symptoms and negative automatic thoughts	3 months No significant change
McGillivray and Kershaw [44] [a]	Three groups EG1: CBT + GP referral EG2: CBT only CG: GP referral only	*N* = 82 (EG1: 32; EG2: 24; CG: 24) Mean age: 37 47 M/35F IQ range 50–70 Mild intellectual disabilities Target: Depression Location: Australia	12 weeks 1.5-h sessions Think happy, feel happy. Focus on feelings, thoughts and social interactions 6–7 per group	EG1 greatest reduction in depressive symptoms; EG2 greatest reduction for negative automatic thoughts; increase automatic thoughts for CG Sig. effect of time and group x time No reduction on depressive symptoms for CG; no sig. Effect of group without time	8 months EGs effective at reducing depressive symptoms and automatic thoughts; slight reduction in depressive symptoms for CG but increase of automatic thoughts
McGillivray and Kershaw [45] [a]	Three groups EG1: cognitive behavioural strategies EG2: cognitive only EG3: behavioural only	*N* = 70 (EG1: 23; EG2: 23; EG3: 24) Mean age: 36 42 M/28F Mild intellectual disabilities Target: Depression Location: Australia	12 weeks 1.5 h Think happy, feel happy. Focus on feelings, thoughts and social interactions	Greatest reduction in EG1 for depressive symptoms and automatic thoughts. Only time sig. For depressive symptoms, group x time for automatic thoughts Minimal change on automatic thoughts for EG3	6 months Continued reduction of depressive symptoms for EG1 (slight increase for EG2 and EG3). Increase of automatic thoughts for EG1, further reduction for EG2
Rose, Loftus, Flint, and Carey [56]	Two groups EG: group anger intervention CG: wait list	*N* = 61 (EG: 50; CG: 36) 71 M/15F Receptive vocabulary Target: Anger Location: UK	16 weeks 2-h sessions Role play, videos, group discussion.	**Anger Inventory** EG reduced anger	3–6 months Effects maintained
Rose, Dodd and Rose [55]	Two groups EG: CBT CG: wait list	EG: N = 20, avg. age = 37.05 CG: *N* = 21, avg. age = 37.14 29 M/12F Mild/borderline intellectual disabilities Target: Anger Location: UK	14–18 weeks 30–60 min sessions Emotional recognition and problem solving Group size: Individual	**Anger Inventory** Significant reduction in anger in EG. Higher baseline score associated with larger decrease.	3–6 months. Effects maintained
Rose, O'Brien, and Rose [58]	Three groups EG1: group CBT EG2: individual CBT CG: wait list	*N* = 64 (EG1: 25; EG2: 18; CG: 21) 43 M/21F British picture vocabulary scale Target: Aggression Location: UK	14–18 weeks 30–60 min sessions EG1: role play, thought stopping, identifying negative thoughts, etc. 16 2 h weekly sessions EG2: same as EG1 but less role play.	**Anger Inventory** Significant reduction in anger in EG. Higher baseline score associated with larger decrease.	None

(*continued on next page*)

Table 1 (*continued*)

First author (year)	Study design	Participants	Intervention	Outcome	Follow-up
Silvestri [61] [b]	3 groups EG: implosive therapy CG1: pseudotreatment CG2: no treatment	N = 24 (8 per group) Mean age 21 Average 5 M/3F per group Avg. IQ 70.3 Borderline/mild intellectual disabilities Target: Anxiety Location: US	2 interviews, 8 sessions EG1 45 min sessions EG1: imagining of traumatic scenes, role play. EG2: recount dreams and fantasies	Improvement in EG across all measures (behavioural, clinical and subjective) compared to CGs	6 weeks Effects not maintained but did not fall below baseline
Taylor, Novaco, Guinan, and Street [64]	Two groups EG: anger treatment CG: wait list	*N* = 48 Mean age 30.99 All male Avg. IQ 69.40 Target: Anger Location: UK	18 sessions Cognitive behavioural treatment – cognitive restructuring, arousal reduction and behavioural skills training	**Imaginal Provocation Index – anger composite subscale** EG sig diff anger and behavioural reactions, anger composite Non-sig diff in anger regulation	None
Taylor, Novaco, Gillmer, Robertson, and Thorne [63]	Two groups EG: anger treatment CG: wait list	EG: N = 16, avg. age = 29.4, avg. IQ 67.1 CG: *N* = 20, avg. age = 29.9, avg. IQ 70.7 Target: Anger Location: UK	18 sessions Individual sessions tailored to anger problems of individual	**Provocation Inventory** Significant reduction on anger x time for EG Other differences not significant but in line with predicted trend	4 months Effects maintained
Willner, Jones, Tams, and Green [73]	Two groups EG: anger management CG: wait list	N = 14 (7 per group) Mean age: EG = 31.4; CG = 30.4 9 M/5F Mild intellectual disabilities Target: Anger Location: UK	9 weeks 2-h sessions Group sessions. Relaxation techniques, behavioural and cognitive strategies, role play. Group size: 7	**Provocation Index** Significant reduction in anger ratings in EG x time. CG tended to increase in anger ratings	3 months Further decrease in EG anger ratings
Willner, Brace, and Phillips [72] [b]	Two groups EG: anger management CG: wait list	N = 17 (EG = 8; CG = 9) Mean age: EG = 44.8; CG = 31.5 12 M/5F Receptive vocabulary Target: Anger Location: UK	12 weeks 2-h sessions Group sessions, relaxation techniques and discussion 8–9 per group Location: day service	Significant reduction in anger ratings in EG compared to CG, and anger coping skills.	6 months Further decrease in EG anger rating, other effects maintained
Willner, et al. [74]	Two groups EG: CBT CG: TAU	EG: *N* = 91, avg. age = 37, 71.4% M, avg. IQ 59 CG: *N* = 90, avg. age = 38.5, 70% M, avg. IQ 55 Target: Anger Location: UK	12 sessions Psychoeducational group sessions 4–9 per group	**Provocation Index** Effects on secondary measures No significant effect of EG on self-reported anger	10 months Still no effect, smaller

Note. Studies and outcome measures in bold were included in the meta-analysis.
[a] = control/comparison group not eligible for meta-analysis, [b] = data for meta-analysis not reported.
CG = control group, EG = experimental group, F = female, M = male, PPVT = Peabody Picture Vocabulary Test, TAU = treatment as usual.
Outcome measure included if study eligible for meta-analysis.

using a Galbraith plot, Supplementary Materials: Appendix 3, led to the removal of Hartley, et al. [23] which decreased effect size to g = 0.84, 95% CI [0.34, 1.35], *N* = 262 and reduced I^2 to 68%, x^2 (5) = 15.72, *p* = 0.008 and the prediction interval to [−0.74, 2.43], followed by the removal of Rose et al. [55] and Rose et al. [58] which decreased the effect size further to g = 0.72, 95% CI [0.16, 1.29], *N* = 196 and I^2 to 65%, x^2 (4) = 11.55, *p* = 0.02, and an associated wide prediction interval [−1.14, 2.58]. The final study identified as an outlier, McGaw et al. [43] was then removed leading to an increase in effect size, g = 0.90, 95% CI [0.47, 1.32], *N* = 174, and reduced I^2 to 31%, x^2 (3) = 4.37, *p* = 0.22, with a continuing wide prediction interval was calculated [−0.51, 0.2.31]. Together, these results suggest substantial uncertainty regarding the effect size generated from non-randomised studies due to clear problems with heterogeneity and decreasing sample sizes as outliers were removed from the analysis.

3.4.3. Group- vs. individual-based interventions

Group-based interventions alone were associated with a large effect size of g = 0.83, 95% CI [0.38, 1.28], *N* = 566, with a high I^2 of 81%, x^2 (10) = 52.10, *p* <0.001 and a wide prediction interval [−73, 2.38], Fig. 6. Most of the overall studies used group-based interventions with a weighting of 57.3%. Visualisation using a Galbraith plot, Supplementary Materials: Appendix 4, led initially to the removal of Rose et al. [58], reduced the effect size to, g = 0.76, 95% CI [0.29, 1.22], *N* = 520, with a slight reduction in I^2 to 80%, x^2 (9) = 45.45, *p* <. 001, and a prediction interval of [−0.81, 2.32]. This was followed by the removal of McCabe et al. [42], reducing the effect size further to, g = 0.65, 95% CI [0.18, 1.12], *N* = 471, with a slight reduction in I^2 to 78%, x^2 (8) = 36.42, *p* <. 001, and a prediction interval of [−0.86, 2.16], followed by the removal of Rose et al. [56] reducing the effect size to, g = 0.66, 95% CI [0.11, 1.22], *N* = 385, while I^2 increased to 80%, x^2 (7) = 34.72, *p* <. 001, and a widening of the prediction interval to [−1.14, 2.47]. Finally, Hartley et al. [23] was removed leading to a reduction in effect size, g = 0.37, 95% CI [0.05, 0.68], *N* = 361, with a large reduction in I^2 to 38%, x^2 (6) = 9.69, *p* = 0.14, and a prediction interval that narrowed, but continued to cross zero [−0.40, 1.14].

Individualised interventions were associated with a moderate effect size, g = 0.58, 95% CI [0.15, 1.01], *N* = 292, which was not significantly smaller than group-based interventions, x^2 (1) = 0.62, *p* = .43, Fig. 7. For individual interventions, I^2 was 67%, x^2 (7) = 21.38, *p* = 0.003, and

Table 2
Summary of quality appraisal findings.

Author (year)	Internal validity	External validity	Comments
Benson et al. [2]	+	+	Selection: stratified by verbal ability and anger, unclear if randomised Allocation concealment: not reported Blinding: raters blind for some outcomes Power: underpowered – sample spread across many groups Other: only reported significant results
Cooney et al. [10]	++	++	Selection: simple randomisation; from 1 centre Allocation concealment: not concealed Blinding: no blinding Power: sufficiently powered for medium effect size Other: harms mentioned
Dowling et al. [16]	+	++	Selection: cluster randomisation by residential homes, simple randomisation for day centre service users Allocation concealment: by independent researcher Blinding: raters not blind Power: underpowered Other: gatekeeping source of bias in integrated intervention, high withdrawal, length of exposure underreported, intention to treat analysis used
Hagiliassis et al. [21]	+	++	Selection: stratified randomisation by sex and region, from 1 centre Allocation concealment: by external researcher Blinding: no blinding Power: underpowered
Hartley et al. [23]	+	+	Selection: not randomised, unclear Allocation concealment: not reported Blinding: no blinding Power: underpowered Other: measured understanding of CBT components
Hassiotis et al. [24]	++	+	Selection: simple randomisation Allocation concealment: by research assistant Blinding: blind assessment, $n = 2$ unintentional unmasking Power: sufficiently powered Other: fidelity check completed; economic assessment; self-report measure not specific to intellectual disabilities
[77]	+	+	Selection: based on state requirement; from 1 agency Allocation concealment: no concealment Blinding: teachers (raters) blind Power: sufficiently powered Other: outcomes only focused on behavioural problems; unbalanced groups
Jahoda et al. [29]	++	++	Selection: stratified block randomisation by centre and use of anti-depressant Allocation concealment: concealed Blinding: blinded Power: sufficiently powered Other: protocol available; fidelity check completed, harms assessed, economic assessment, intention to treat analysis
Karatzias et al. [32]	++	++	Selection: simple randomisation Allocation concealment: computer generated Blinding: assessors blind Power: underpowered Other: therapist supervision, harms discussed in interviews, intention to treat analysis
Klein et al. [33]	+	++	Selection: simple randomisation Allocation concealment: not concealed but based on pre-screening allocation Blinding: not reported Power: sufficiently powered Other: harms discussed
Lawrence [35]	+	+	Selection: simple random, from 1 centre Allocation concealment: not reported Blinding: not reported Power: likely underpowered Other: variability in attendance
[78]	−	−	Selection: first come, first serve Allocation concealment: not concealed Blinding: only assessors of role-play blind but process not detailed Power: likely underpowered Other: group differences in age and sex, self-reports not specific to intellectual disabilities, difficulties in completing daily diaries and role-plays, no follow-up for CG
Lindsay et al. [38]	+	+	Selection: matched pairs Allocation concealment: not concealed Blinding: no blinding Power: likely underpowered Other: no follow-up for CG
Mann [39]	−	+	Selection: simple randomisation Allocation concealment: not concealed Blinding: no blinding Power: likely underpowered Other: same investigator across interventions, outcome measures not appropriate for intellectual disabilities, analytical methods poorly described
Matson [40]	+	+	Selection: matched pairs Allocation concealment: not reported Blinding: not reported Power: likely underpowered Other: high prevalence of schizophrenia and neurosis, precision of effects incalculable
Matson and Senatore [41]	+	−	Selection: matching triads Allocation concealment: not reported Blinding: raters blind Power: likely underpowered Other: selection based on staff report
[42]	+	+	Selection: simple randomisation Allocation concealment: not reported Blinding: not reported Power: likely underpowered
McGaw et al. [43]	+	+	Selection: first come, first serve; from 1 service Allocation concealment: not concealed

(continued on next page)

Table 2 (*continued*)

Author (year)	Internal validity	External validity	Comments
			Blinding: not reported Power: likely underpowered Other: harms observed with malaise inventory
McGillivray et al. [47]	++	+	Selection: random by agency Allocation concealment: staff naïve Blinding: independent assessor Power: likely underpowered
McGillivray and Kershaw [44]	+	++	Selection: by agency and recommendation Allocation concealment: not reported Blinding: independent rater Power: likely sufficiently powered
McGillivray and Kershaw [45]	++	+	Selection: cluster randomisation Allocation concealment: not reported Blinding: not blind but all self-report Power: sufficiently powered Other: therapist supervision
Rose et al. [56]	+	+	Selection: first come, first serve Allocation concealment: not concealed Blinding: not reported Power: sufficiently powered Other: inclusion of CG participants in EG, follow-up only for EG, differences in characteristics between groups not reported/ analysed
Rose et al.[55]	++	++	Selection: first come, first serve; from 1 centre Allocation concealment: not concealed Blinding: not reported Power: sufficiently powered Other: weekly supervision of therapists, variable follow-up times
Rose et al. [58]	+	+	Selection: by treatment availability; from 1 centre Allocation concealment: not concealed Blinding: not reported Power: sufficiently powered Other: age and measure of intellectual disability not reported
Silvestri [61]	–	+	Selection: matched pairs, from 1 centre Allocation concealment: not reported Blinding: some blinding Power: likely underpowered Other: incomplete outcome scales, no controlling for multiple analyses
Taylor et al. [64]	+	+	Selection: not randomised, not specified Allocation concealment: not concealed Blinding: not blinded Power: likely underpowered Other: therapist supervision
Taylor et al. [63]	++	++	Selection: randomised by anonymised list Allocation concealment: concealed Blinding: blinded Power: likely underpowered Other: all male participants, outcomes not specific to intellectual disabilities
Willner et al. [73]	+	+	Selection: randomisation with some swapping Allocation concealment: not concealed Blinding: not blinded Power: likely underpowered Other: recruitment by staff referral, outcomes not validated
Willner et al. [72]	+	++	Selection: by staff and participant preference Allocation concealment: not concealed Blinding: not blinded Power: likely underpowered Other: outcomes not validated but specific to intellectual disabilities, minor issues in completion of outcome measures
Willner et al. [74]	++	++	Selection: randomised by centre Allocation concealment: concealed Blinding: blinded Power: sufficiently powered

Note. CG = control group, EG = experimental group.
++ = study designed to minimise bias. + = not all sources of bias addressed/ some aspects unclear.
- = significant sources of bias.

the prediction intervals was [−0.76, 1.92]. Inspection of the Galbraith plot, Supplementary Materials: Appendix 4, identified Rose et al. [55] as an outlier initially, and following removal the effect size reduced to small, g = 0.47, 95% CI [0.04, 0.90] $N = 251$, with a reduction in I^2 to 62%, x^2 (6) = 15.63, p =. 02, and a prediction interval of [−0.80, 1.74]. Lindsay et al. [38] was then removed which resulted in further reduction in effect size, g = 0.32, 95% CI [.-0.07, 0.71], $N = 227$, with a further reduction to $I^2 = 50\%$, x^2 (5) = 9.94, $p = 0.08$, and a prediction interval of [−0.61, 1.91]. Klein et al. [33] was them removed, increasing the effect size, g = 0.45, 95% CI [0.06, 0.84], $N = 163$, with a reduction in I^2 to 31%, x^2 (4) = 5.81, $p = 0.21$, and a prediction interval of [−0.56, 1.46], and finally, Taylor et al. [64] was removed. The resulting effect size estimate was non-significant, g = 0.32, 95% CI [−0.01, 0.65], N = 146, while I^2 reduced to 0%, x^2 (3) = 1.13, $p = 0.77$, and a prediction interval of [−0.40, 1.04]. The analysis suggested substantial uncertainly around the estimate of effect size for psychotherapy delivered to individuals due to problems with heterogeneity and conclusions about effectiveness of these interventions at the current time are unclear.

3.4.4. Clinical presentation: Anger, depression/anxiety, interpersonal functioning

As Karatzias, et al. [32] was the only study to focus on trauma, it was not included in the clinical presentation subgroup analysis (Fig. 5). A large proportion of participants were based in studies that focused on anger, with a weighting of 48.7% for anger-based interventions. A moderate effect size was found for this subgroup, g = 0.76, 95% CI [0.43, 1.10], $N = 476$ with an I^2 of 57%, x^2 (8) = 16.18, $p = 0.02$, and a prediction interval of [−0.16, 1.68], Fig. 7.

Visualisation using a Galbraith plot, Supplementary Materials: Appendix 5, led initially to the removal of Rose et al. [58] and Rose et al. [55], and reduced the effect size to, g = 0.61, 95% CI [0.33, 0.89], N = 410, with a substantial reduction in I^2 to 32%, x^2 (6) = 8.83, p =0.18, and a prediction interval of [−0.03, 1.25]. Rose et al. [56] was then removed, and the estimate of effect size reduced slightly to, g = 0.60, 95% CI [0.26, 0.93], $N = 324$, with a slight increase in I^2 to 35%, x^2 (5) = 7.67, $p = 0.18$, and a prediction interval of [−0.23, 1.42]. No further studies were removed.

Most studies in the 'Depression and Anxiety' subgroup had outcome measures of depression, except for Lindsay, et al. [38] who used a global severity index measure of psychopathology and Klein, et al. [33] who

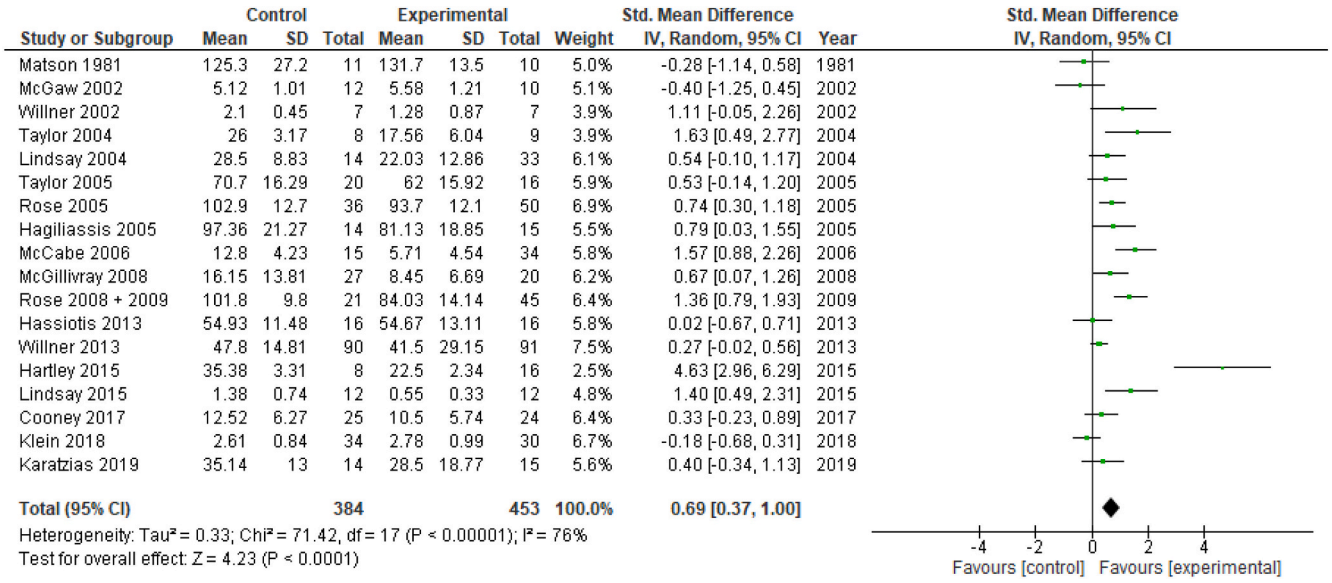

Fig. 2. Forest plot of all studies included in meta-analysis.

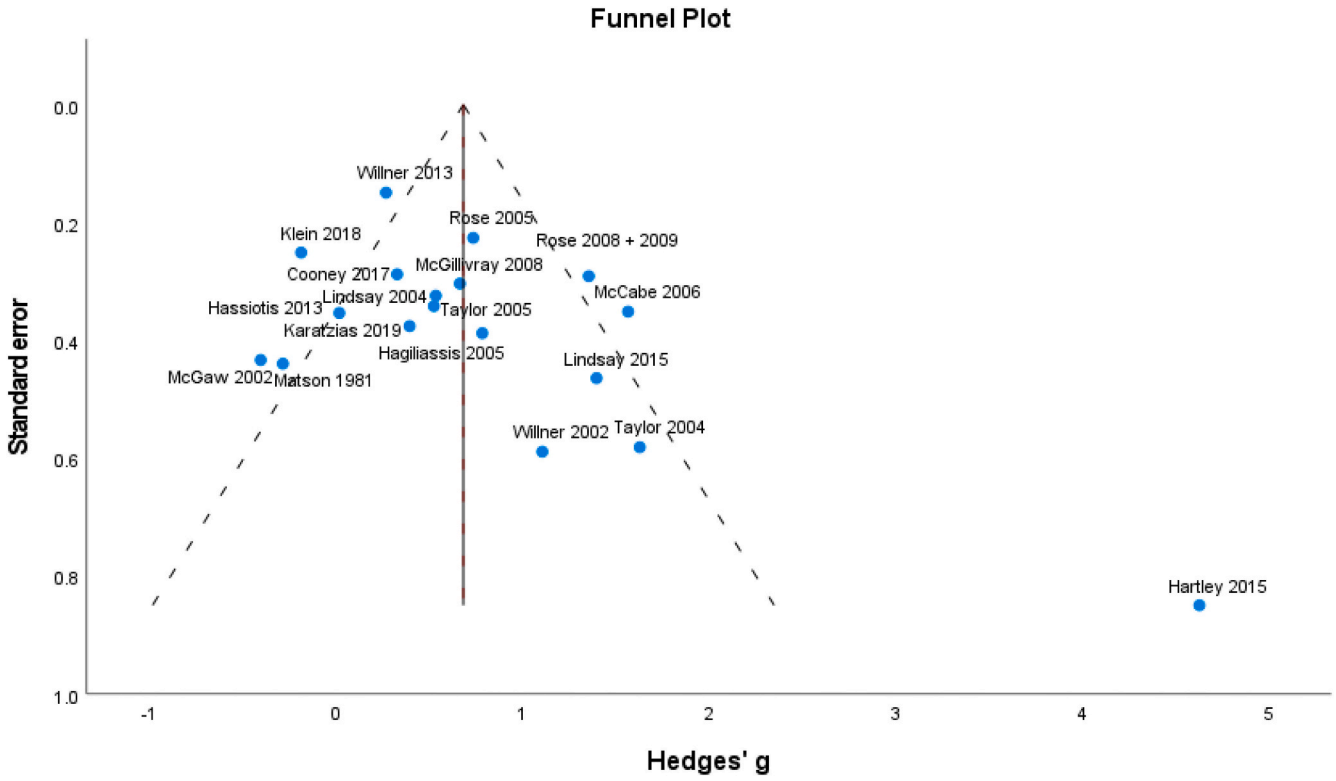

Fig. 3. Funnel plot for all studies included within the meta-analysis.

used a social anxiety measure. The subgroup analysis resulted in a large effect size, g = 0.98, 95% CI [0.25, 1.71], $N = 289$, with $I^2 = 87\%$, x^2 (6) = 46.31, $p \leq 0.001$, and a large prediction interval of [−1.52, 3.48], and an overall weighting of 40.5%, Fig. 9.

Following visualisation of heterogeneity, Supplementary Materials: Appendix 5, McCabe et al. [42] was removed from this analysis and the effect size decreased to, g = 0.87, 95% CI [0.09, 1.64], $N = 240$, with a small reduction in I^2 to 86%, x^2 (5) = 36.80, $p <$. 001, and a wide prediction interval of [−1.81, 3.54]. Hartley et al. [23] was then removed which led to a reduction in the effect size estimate, which was non-significant, g = 0.38, 95% CI [−0.10, 0.85], $N = 216$, with a reduction in I^2 to 65%, x^2 (4) = 11.41, $p = 0.02$, and a prediction interval of [−1.20, 1.95] which crossed zero. No further studies were removed at this stage. There was evidence of substantial uncertainly about the effect size estimate for studies that targeted depression and anxiety.

Interpersonal functioning studies resulted in a negative effect size of g = −0.34, CI [−0.95, 0.26], $N = 43$, Fig. 10. Further analysis was not possible as the subgroup only contained two studies. There was a significant difference between subgroups within the clinical presentation analysis, $x^2(2) = 11.22$, $p = .004$.

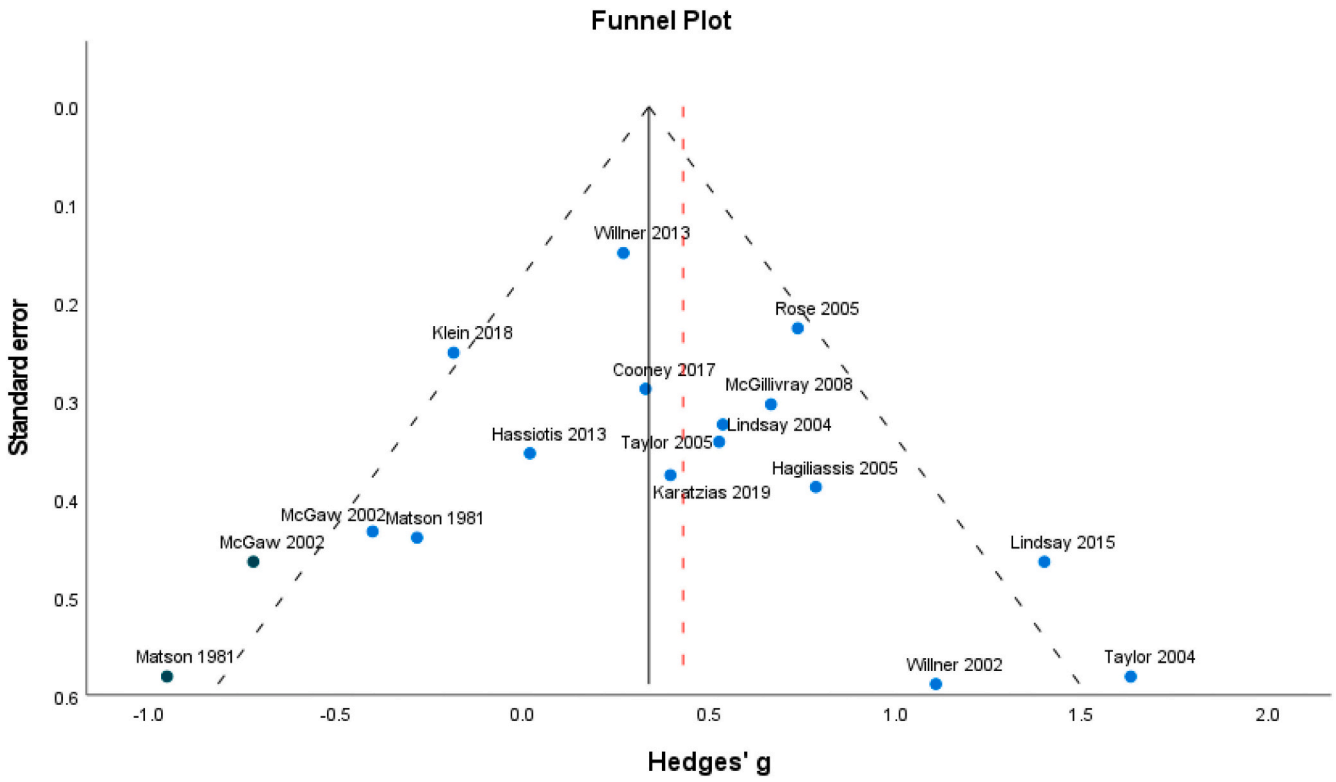

Fig. 4. Funnel plot following the removal of studies judged to be contributing unduly to heterogeneity following visualisation using a Galbraith plot (see Supplementary Materials: Appendix 2).

Study or Subgroup	Control Mean	Control SD	Control Total	Experimental Mean	Experimental SD	Experimental Total	Weight	Std. Mean Difference IV, Random, 95% CI	Year
1.1.1 Randomised									
Matson 1981	125.3	27.2	11	131.7	13.5	10	5.0%	-0.28 [-1.14, 0.58]	1981
Willner 2002	2.1	0.45	7	1.28	0.87	7	3.9%	1.11 [-0.05, 2.26]	2002
Taylor 2005	70.7	16.29	20	62	15.92	16	5.9%	0.53 [-0.14, 1.20]	2005
Hagiliassis 2005	97.36	21.27	14	81.13	18.85	15	5.5%	0.79 [0.03, 1.55]	2005
McCabe 2006	12.8	4.23	15	5.71	4.54	34	5.8%	1.57 [0.88, 2.26]	2006
McGillivray 2008	16.15	13.81	27	8.45	6.69	20	6.2%	0.67 [0.07, 1.26]	2008
Hassiotis 2013	54.93	11.48	16	54.67	13.11	16	5.8%	0.02 [-0.67, 0.71]	2013
Willner 2013	47.8	14.81	90	41.5	29.15	91	7.5%	0.27 [-0.02, 0.56]	2013
Cooney 2017	12.52	6.27	25	10.5	5.74	24	6.4%	0.33 [-0.23, 0.89]	2017
Klein 2018	2.61	0.84	34	2.78	0.99	30	6.7%	-0.18 [-0.68, 0.31]	2018
Karatzias 2019	35.14	13	14	28.5	18.77	15	5.6%	0.40 [-0.34, 1.13]	2019
Subtotal (95% CI)			**273**			**278**	**64.3%**	**0.43 [0.14, 0.72]**	
Heterogeneity: Tau² = 0.13; Chi² = 24.03, df = 10 (P = 0.008); I² = 58%									
Test for overall effect: Z = 2.90 (P = 0.004)									
1.1.2 Non-randomised									
McGaw 2002	5.12	1.01	12	5.58	1.21	10	5.1%	-0.40 [-1.25, 0.45]	2002
Lindsay 2004	28.5	8.83	14	22.03	12.86	33	6.1%	0.54 [-0.10, 1.17]	2004
Taylor 2004	26	3.17	8	17.56	6.04	9	3.9%	1.63 [0.49, 2.77]	2004
Rose 2005	102.9	12.7	36	93.7	12.1	50	6.9%	0.74 [0.30, 1.18]	2005
Rose 2008 + 2009	101.8	9.8	21	84.03	14.14	45	6.4%	1.36 [0.79, 1.93]	2009
Lindsay 2015	1.38	0.74	12	0.55	0.33	12	4.8%	1.40 [0.49, 2.31]	2015
Hartley 2015	35.38	3.31	8	22.5	2.34	16	2.5%	4.63 [2.96, 6.29]	2015
Subtotal (95% CI)			**111**			**175**	**35.7%**	**1.20 [0.50, 1.89]**	
Heterogeneity: Tau² = 0.68; Chi² = 35.20, df = 6 (P < 0.00001); I² = 83%									
Test for overall effect: Z = 3.37 (P = 0.0008)									
Total (95% CI)			**384**			**453**	**100.0%**	**0.69 [0.37, 1.00]**	
Heterogeneity: Tau² = 0.33; Chi² = 71.42, df = 17 (P < 0.00001); I² = 76%									
Test for overall effect: Z = 4.23 (P < 0.0001)									
Test for subgroup differences: Chi² = 3.96, df = 1 (P = 0.05), I² = 74.8%									

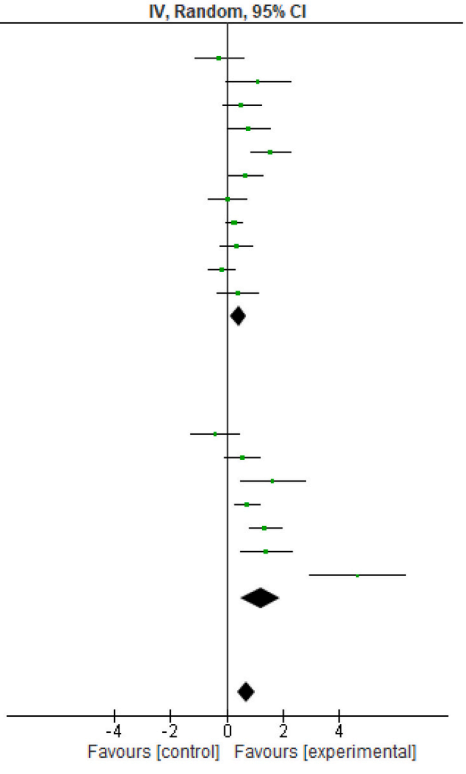

Fig. 5. Forest plot of randomised vs. non-randomised studies subgroup analysis.

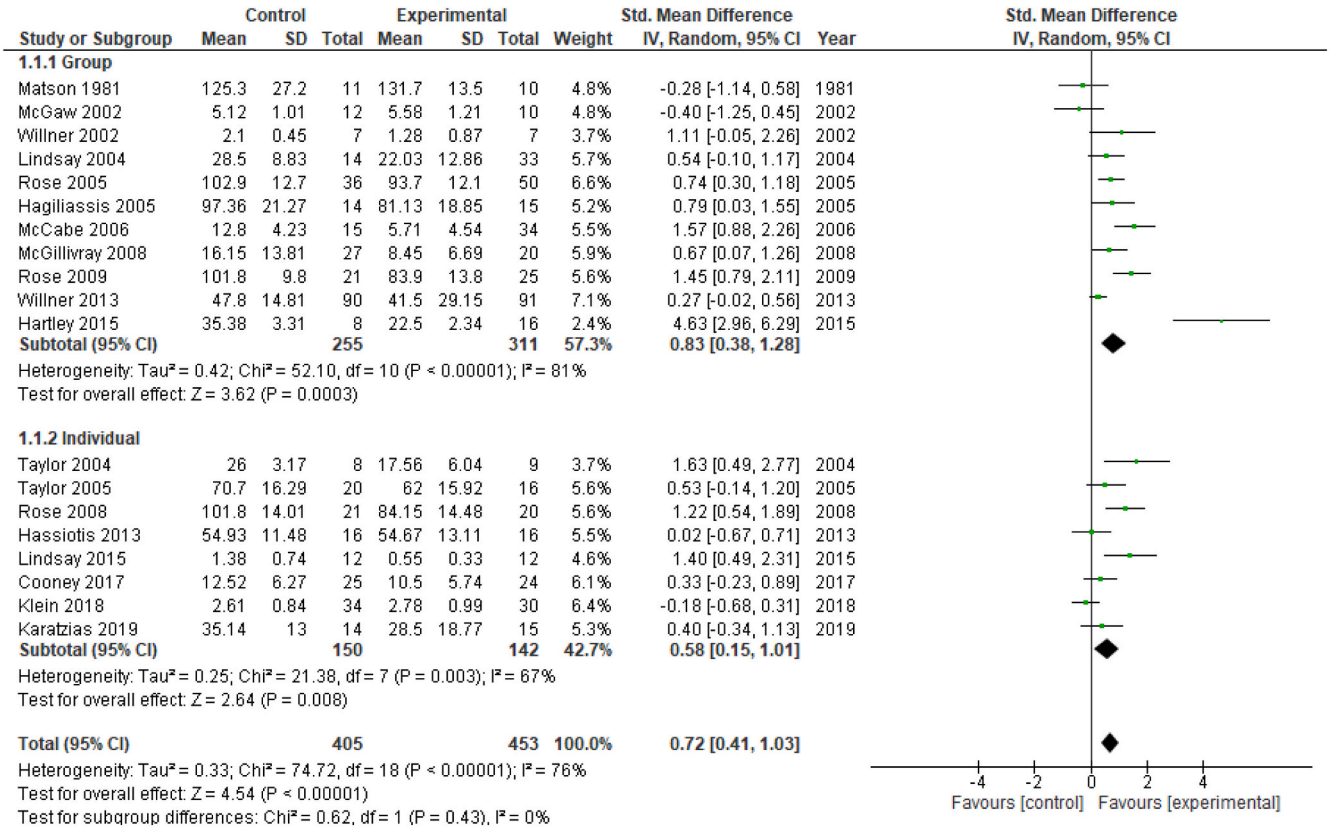

Fig. 6. Forest plot of group vs. individual based interventions subgroup analysis.

4. Discussion

Initially, a moderate effect size was found for the use of psychotherapies for mental health problems when used with people with intellectual disabilities, g = 0.69, *n* = 837, but there was evidence to indicate a large proportion of the observed variance was not attributable to sampling error, and the estimate of true effect sizes varied widely. There was also evidence to suggest that studies with a small number of participants had an undue impact upon the estimate of effect size. This points toward substantial uncertainty about the estimated effect size. Removal of studies associated with heterogeneity led to a reduction in effect size, g = 0.43, *n* = 698, which was small, and increased in the proportion of observed variance due to sampling error, but the estimate of true effect size continued to vary widely. Our further Trim and Fill analysis following the removal of outliers identified missing studies, indicating publication bias, and there was evidence to indicate an undue effect of small studies. Imputation of missing studies led to a further reduction in effect size to, g = 0.34, which was small.

Our subgroup analysis of randomised vs non-randomised studies indicated that the effect size calculated using randomised studies was small, g = 0.43, *n* = 551. Removal of outliers led to a reduction in the effect size to, g = 0.36, *n* = 438, with all of the observed variance being accounted for by sampling error, while variation in our estimte of true effect sizes did not cross zero, and ranged from small to moderate. It is important to note that this estimate was based upon only 438 participants. Comparing this to non-randomised studies indicated that a lack of randomisation was a clear source of heterogeneity within our meta-analysis. A large effect size was found for non-randomised studies, g = 1.20, based upon only 286 participants, with a large proportion of the variance of observed variance being attributable to variance amongst true effect sizes, with wide variation in the estimated true effect sizes. Removal of outliers reduced the effect size to, g = 0.90, based upon only 174 participants. The amount of observed variance attributable to sampling error increased, while variation in the estimated true effect sizes remained wide.

Further sub-group analysis by whether therapy was delivered within a group or individually revealed that group-based interventions were associated with a large effect size, g = 0.83, *n* = 566, and there were problems with observed variance being attributable to variation in true effect sizes and a wide prediction interval. Removal of outliers decreased the effect size estimate to small, g = 0.37, *n* = 361, which remained significant, and led to an increase in the degree of observed variance due to sampling error and narrowed the prediction interval. Studies that delivered therapy to individuals were initially associated with a moderate effect size, g = 0.58, *n* = 292, with associated problems with the estimate being affected by small studies, a wide prediction interval, and the observed variance being attributable to variation in true effect sizes. Again, removal of outliers decreased the effect size to small, g = 0.32, *n* = 146, which was non-significant. It is not possible to conclude that individually delivered psychotherapy to people with intellectual disabilities is effective based upon the studies included within our meta-analysis.

Finally, we undertook a further subgroup analysis by clinical presentation. Studies targeting anger were initially associated with a moderate effect size, g = 0.76, *n* = 476, and removal of outliers reduced this estimate, which remained moderate, g = 0.60, *n* = 324, with an associated prediction interval indicating variation amongst the estimated true effect sizes. There were problems with the reliability of our initial estimated effect size, g = 0.87, *n* = 289, for psychotherapy targeting anxiety and depression and removal of outliers reduced the estimate to small and non-significant, g = 0.38, *n* = 216, with variation in the estimate of true effect sizes. As a consequence, it is not possible to conclude that psychotherapy for depression and anxiety is effective for people with intellectual disabilities based upon the studies included within our meta-analysis. Finally, psychotherapy targeting interpersonal functioning were associated with a negative effect size, g = −0.34, *n* =

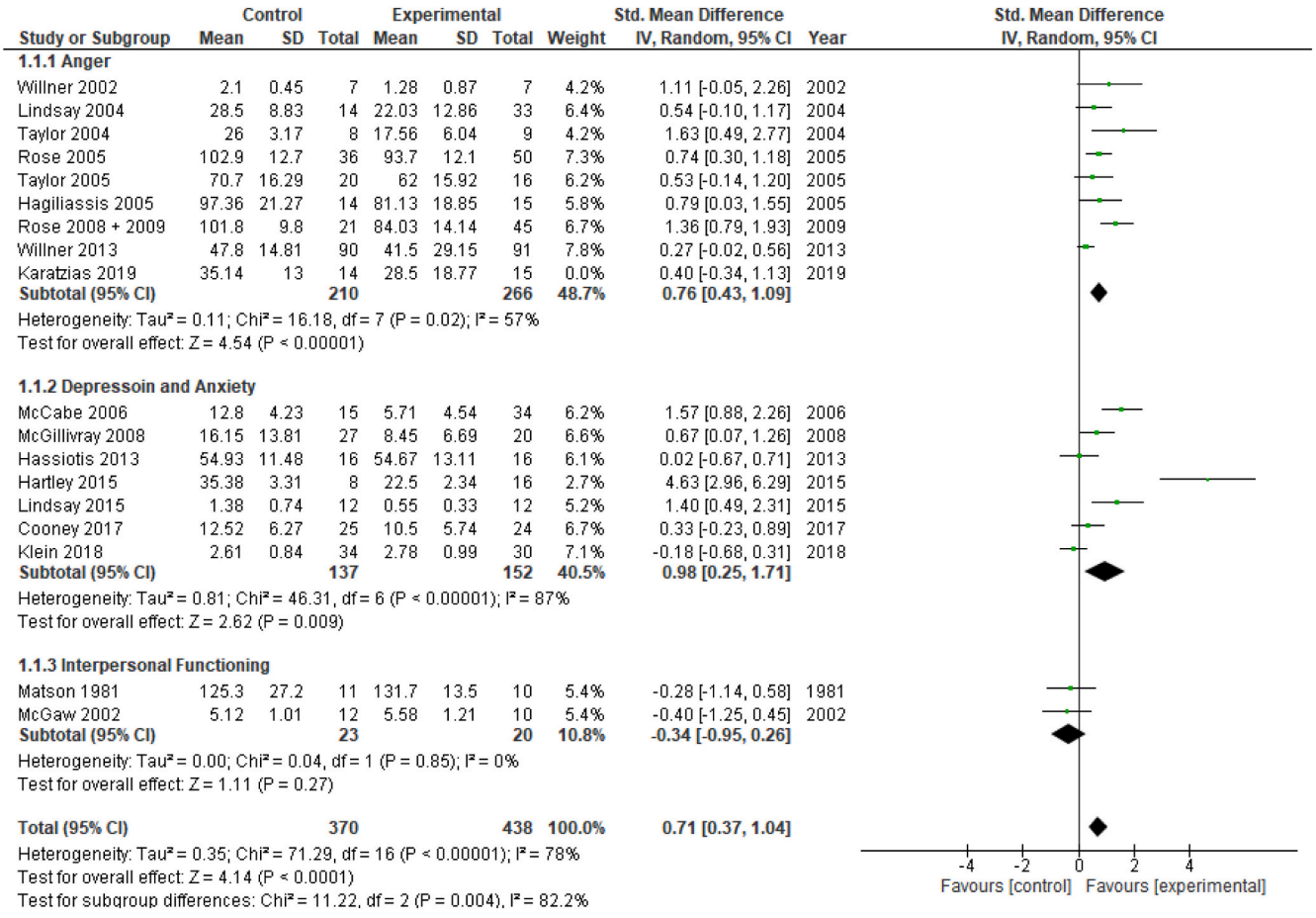

Fig. 7. Forest plot of clinical presentation subgroup analysis.

43. The results of this sub-group analysis indicated that there was evidence to indicate that psychotherapy for anger is associated with a moderate effect size, while estimates for depression and anxiety, and interpersonal functioning are unreliable based upon the current evidence.

It is worth noting that the I^2 statistic was initially always above 40%, indicating that variance in observed effects is attributable to variation in true effects, rather than sampling error [14]. Outlier removal reduced I^2 in all cases, and within our subgroup analysis for randomised studies, this reduced to 0%, while it reduced to below 40% following the removal of outliers within our subgroup analysis for non -randomised studies, and studies using group-based interventions, and studies targeting anger. The prediction interval for most analyses included zero, with the exception of our subgroup analysis involving randomised studies following the removal of outliers, meaning the likely effectiveness of future clinical trials is unclear. However, prediction intervals may not be suitable when a small number of studies are included in a meta-analysis [14]. As many subgroup analyses in this review contained a small number of participants, all estimates should be interpreted with caution.

Making comparisons to Vereenooghe and Langdon [67], they estimated the effect size calculated using randomised studies as small, g = 0.39, n = 388, following the exclusion of the study by McCabe et al. [42], which was also identified as an outlier in the current meta-analysis. In our current meta-analysis, we also removed Klein et al. [33], leading to a current estimated effect size that was similar and small, g = 0.36, n = 438. However, Vereenooghe and Langdon [67] went on to remove two further studies that were not identified as outliers within our current meta-analysis leading to an estimated moderate effect size, g = 0.65, n = 367, following the removal of Matson and Senatore [41], and g = 0.64, n = 358, following the removal of Hassiotis et al. [24]. Vereenooghe and Langdon [67] reported a large effect size for non-randomised studies, g = 0.85, n = 275, while we also calculated a large effect size, g = 0.90, n = 174, for non-randomised studies following the removal of outliers. It should be noted that Vereenooghe and Langdon [67] did not visualise and remove outliers from their analysis of non-randomised studies.

We calculated a small effect size for group-based interventions after the removal of outliers, g = 0.37, n = 361, compared to individual-based interventions which were associated with a non-significant and small effect size, g = 32, n = 146. This was not the case in Vereenooghe and Langdon's [67] review, in which individual-based interventions had a larger effect size than group-based interventions. It should be noted that Vereenooghe and Langdon [67] reported an estimated effect size for individually delivered psychotherapy based upon four studies, n = 124. They did not remove any outliers but noted variability in the estimated effect size across studies. Four studies in the individual-based subgroup ([10]; [33,38]) were not included in Vereenooghe and Langdon's [67] original review. Lindsay, et al. [38] reported marginally significant effects, which was identified as an outlier within our current analysis, whilst [10] reported no significant differences following treatment, and Klein, et al. [33] reported significant results at the 10-week follow-up but not at post-treatment, which again was removed from our analysis. Our findings indicated that there is evidence that group-based interventions appear effective in treating mental health problems in adults with intellectual disabilities, predominantly anger problems, while this is not the case for interventions delivered to individuals following the removal of outliers.

Initially, a moderate effect size was found for the anger-focused intervention subgroup, g = 0.76, n = 476, reducing to, g = 0.60, n =

324, following the removal of outliers which is inconsistent with Vereenooghe and Langdon [67], who reported a large effect size, g = 0.83, n = 494. However, this difference is likely due to the decision in the current review to combine the data of two of Rose's studies [55,58] while Vereenooghe and Langdon [67] did not do this, nor did they consider the impact of potential outliers. Graser et al. [20] also found a moderate effect size for anger interventions in their meta-analysis, although they focused on CBT interventions which excluded studies by Rose et al. [55,56,58].

The large effect size for depression and anxiety interventions reduced considerably after outlier removal and became non-significant in the current meta-analysis. Vereenooghe & Langdon [67] reported a moderate effect size associated with the treatment of depression, g = 0.74, n = 126, generated from only three studies and cautioned that the variability between studies was high. Graser et al. [20] and Koslowski et al. [34] examined anxiety and depression separately in their analyses, both reporting a small effect size for the treatment of anxiety and a medium effect size for depression. Anxiety and depression treatment outcomes were examined together in the current meta-analysis to remain consistent with Vereenooghe and Langdon's [67] original meta-analysis, and because some of the studies attempted to target both when delivering treatment. However, most outcome measures selected in our anxiety and depression subgroup measured depression, and so the moderate effect size likely reflects the findings of the depression intervention subgroups in other meta-analyses.

4.1. Limitations of evidence

The quality appraisal process identified widespread issues including a lack of allocation concealment and blinding, small sample sizes, and a lack of diversity in the recruitment process, and are clear sources of bias limiting our ability to draw firm conclusions about the effectiveness of psychotherapy for people with intellectual disabilities. Further, grouping together studies that targeted different clinical presentations, and made use of different methods to deliver the treatment, likely contributed to heterogeneity. There was clear evidence from out subgroup analysis of randomised vs non-randomised studies that poorer study design was associated with inflated effect sizes. Further, the majority of our included studies made use of small samples and there was evidence of a relationship between study size and the estimated effect size as visualised via funnel plots. As a consequence, many of the included studies were underpowered and/or biased making conclusions about causality problematic. This reduced the certainty of our estimated effect sizes and drawing definitive conclusions about the effectiveness of psychotherapy for people with intellectual disabilities is not currently possible. Further, outside of the quality appraisal process, it was noted that very few studies included participants with severe intellectual disabilities. Most studies recruited participants with mild intellectual disabilities, limiting our understanding for those with a greater degree of intellectual disability. A few studies also included people with borderline intellectual disabilities, but authors did not always make it clear how many participants had an IQ > 69. It would be helpful if future studies reported an estimate of level of general intellectual functioning to enable future meta-analyses to investigate the relationship between treatment outcome and general intellectual functioning. Very few studies included children or adolescents, which remains disappointing, as Vereenooghe and Langdon [67] noted. Finally, reporting of ethnicity and race of participants was uncommon, making it hard to determine the representativeness of studies in this respect.

It must be noted that, difficulties can arise when recruiting participants with intellectual disabilities into clinical trials of psychotherapy. Obtaining consent from individuals with intellectual disabilities can be challenging [1] as additional communication strategies are often needed to facilitate the process of gaining consent [27]. Additional processes may also be required in some jurisdictions when an individual lacks capacity to make an informed decision to partake in research (e.g., [80] in England and Wales). There are also challenges with gatekeeping by carers and clinicians which may unduly affect recruitment [19] as they may try to protect others and themselves or misunderstand research, policies, and practices [71]. Many of these issues have been recognised by others who have argued for the increased inclusion of people with disabilities within clinical trials, noting difficulties with the law in some jurisdictions governing the inclusion of people who lack capacity to consent to take part in clinical research [5,6,26,30].

4.2. Limitations of the review

The primary limitations of this review were around the inclusion criteria. Only studies in English were included. There were also no restrictions around reporting of intellectual functioning. Our aim was to make the review inclusive, considering the paucity of research within this area. However, some studies included people with borderline intellectual disabilities in their samples, and some studies only measured receptive vocabulary rather than Full-Scale IQ. It may therefore be difficult to generalise results due to inconsistencies in inclusion criteria and measures of intellectual functioning. A further limitation is the pooling of studies that targeted anxiety and/or depression together. The rationale for this was that some of these studies attempted to treat both when delivering psychological therapy (e.g., [38]). Nevertheless, this is likely to have contributed to heterogeneity.

4.3. Implications

Following the previous review [67], 11 further eligible studies were identified, with Full-Scale IQ used as a measure of intellectual functioning in 5 of these studies. Not all these studies were eligible for the meta-analysis, meaning we were unable to explore the relationship between intellectual functioning and treatment outcome. Full-Scale IQ was not always reported consistently, and some data were missing. For example, [10] obtained IQ from existing records, but this was not available for all participants. More controlled studies using Full-Scale IQ as a measure of intellectual functioning for all included participants are required to conduct a meaningful meta-regression to address the important question of how degree of intellectual disability may be associated with treatment outcome.

Whilst adaptations to psychotherapies in clinical trials were reported, they were not always reported in detail. NICE [50] recommended adapting psychotherapies according to each individual, including preferences, level of understanding, and strengths and needs whilst taking into account other aspects of disability (e.g., physical or sensory impairment and communication needs). These adaptations could be achieved through increased collaboration with carers. It is unclear to what extent adaptations in clinical trials of psychotherapy may have followed these recommendations and so their effectiveness is unclear. There is a clear need to develop a greater understanding about the nature and degree of adaptations that are made to psychotherapy to meet the needs of people with intellectual disabilities and whether they are effective.

Areas for improvement still include the need for further clinical trials examining a range of mental health problems, other than anger, with both adults and children and adolescents with intellectual disabilities. Future clinical trials should include fidelity checks and considerations of potential harms. [24] study assessed treatment fidelity by having an independent expert rate randomly selected audiotaped session. Consideration of harms was most evident through use of secondary outcome measures, such as intervention satisfaction measures [24] or clinical outcomes evaluations [32]. Efforts should also be made to ensure blinding and allocation concealment, and to report demographics such as race and ethnicity.

5. Conclusions

The results of the current systematic review and meta-analysis were generally consistent with Vereenooghe and Langdon's [67] previous review. There was evidence of bias affecting our estimate of effect size. Larger studies are needed incorporating blinding and allocation concealment along with fidelity checks and consideration of potential harms. Further study of the use of psychotherapy with people with intellectual disabilities is important to increase the quality of research in this area and to enable further analysis into variables related to treatment outcomes, including adaptations made to therapy. It should be noted that there is evidence that increasingly larger and well-design trials are being conducted. For example, one well designed large trial was found that could not be included in our meta-analysis because they compared two psychological treatments for depression or low mood and did not include a treatment-as-usual comparison [29]. As larger and well-designed clinical trials are published over the coming years, it will be important to conduct a further meta-analysis.

Credit statement

Katherine Tapp: methodology, validation, project administation, formal analysis, methodology, investigation, writing - original draft, visualisation. **Leen Vereenooghe:** validation, investigation, writing - review and editing. **Olivia Hewitt:** validation, investigation, writing - review and editing. **Emma Scripps:** validation, investigation, writing, review and editing. **Kylie Gray:** supervision, project administration, methodology, writing - review and editing **Peter Langdon:** conceptualisation, methodology, validation, formal analysis, investigation, project administration, visualisation, supervision, writing - review and editing

Role of funding sources

The review was conducted as part of a PhD project. The project is funded by the University of Warwick Collborative Fellowship (WCF) scheme in collaboration with Herefordshire and Worcestershire Health and Care NHS Trust. The funder had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript, or the decision to submit the paper for publication.

Contributors

KT conducted literature searches and screened studies for eligibility independently alongside ES. KT and OH independently extracted data from the eligible studies. Quality appraisal was carried out independently for all eligible papers by KT and LV. The first draft of the manuscript was written by KT and all authors contributed to and approved the final manuscript. The process was supervised by PEL and KMG.

Declaration of Competing Interest

The authors report no conflicts of interest.

Acknowledgements

The authors would like to thank Samantha Jones, Academic Support Librarian in Psychology, for reviewing and assisting with the development of the search terms.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.comppsych.2023.152372.

References

[1] Arscott K, Dagnan D, Kroese BS. Consent to psychological research by people with an intellectual disability. J Appl Res Intellect Disabil 1998;11(1):77–83.
[2] Benson BA, Rice CJ, Miranti SV. Effects of anger management training with mentally retarded adults in group treatment. J Consult Clin Psychol 1986;54(5):728.
[3] Bloom MK. The effectiveness of modern psychoanalytic interventions with mentally retarded subjects. United States: The Union Institute; 1995.
[4] Bond L, Carroll R, Mulryan N, O'Dwyer M, O'Connell J, Monaghan R, et al. The association of life events and mental ill health in older adults with intellectual disability: results of the wave 3 intellectual disability supplement to the Irish longitudinal study on ageing. J Intellect Disabil Res 2019;63(5):454–65.
[5] Bunning K, Heywood R, Killett A, Shiggins C, Langdon PE. Assent: Final report. Including adults with capacity and communication difficulties in ethically-sound research. The Nuffield Foundation; 2021.
[6] Bunning K, Jimoh OF, Heywood R, Killett A, Ryan H, Shiggins C, et al. How are adults with capacity-affecting conditions and associated communication difficulties included in ethically sound research? A documentary-based survey of ethical review and recruitment processes under the research provisions of the mental capacity act (2005) for England and Wales. BMJ Open 2022;12(3):e059036.
[7] Charlot L, Beasley JB. Intellectual disabilities and mental health: United States–based research. J Ment Health Res Intellect Disabil 2013;6(2):74–105.
[8] Cheung MWL. A guide to conducting a meta-analysis with non-independent effect sizes. Neuropsychol Rev 2019;29(4):387–96.
[9] Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Erlbaum; 1988.
[10] Cooney P, Jackman C, Coyle D, O'Reilly G. Computerised cognitive–behavioural therapy for adults with intellectual disability: randomised controlled trial. Br J Psychiatry 2017;211(2):95–102.
[11] Cooper SA, Smiley E, Morrison J, Williamson A, Allan L. Mental ill-health in adults with intellectual disabilities: prevalence and associated factors. Br J Psychiatry 2007;190(1):27–35.
[12] Dagnan D, Chadwick P, Proudlove J. Toward an assessment of suitability of people with mental retardation for cognitive therapy. Cogn Ther Res 2000;24(6):627–36.
[13] Deb S, Thomas M, Bright C. Mental disorder in adults with intellectual disability. 1: prevalence of functional psychiatric illness among a community-based population aged between 16 and 64 years. J Intellect Disabil Res 2001;45(6):495–505.
[14] Deeks JJ, JPT Higgins, Altman DG. Chapter 10: Analysing data and undertaking meta-analyses. In: JPT Higgins, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022). Cochrane; 2022. Available from, http://www.training.cochrane.org/handbook.
[16] Dowling S, Hubert J, White S, Hollins S. Bereaved adults with intellectual disabilities: a combined randomized controlled trial and qualitative study of two community-based interventions. J Intellect Disabil Res 2006;50(4):277–87.
[17] Duval S, Tweedie R. Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000;56(2):455–63.
[18] Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315(7109):629–34.
[19] Forbat L, McCann L. Adults with intellectual disabilities affected by cancer: critical challenges for the involvement agenda. Eur J Cancer Care 2010;19(1):91–7.
[20] Graser J, Göken J, Lyons N, Ostermann T, Michalak J. Cognitive-behavioral therapy for adults with intellectual disabilities: a meta-analysis. Clin Psychol Sci Pract 2022. https://doi.org/10.1037/cps0000077. Advance online publication.
[21] Hagiliassis N, Gulbenkoglu H, Di Marco M, Young S, Hudson A. The anger management project: a group intervention for anger in people with physical and multiple disabilities. J Intellect Dev Disabil 2005;30(2):86–96.
[22] Hamers PCM, Festen DAM, Hermans H. Non-pharmacological interventions for adults with intellectual disabilities and depression: a systematic review. J Intellect Disabil Res 2018;62(8):684–700.
[23] Hartley SL, Esbensen AJ, Shalev R, Vincent LB, Mihaila I, Bussanich P. Cognitive behavioral therapy for depressed adults with mild intellectual disability: a pilot study. J Ment Health Res Intellect Disabil 2015;8(2):72–97.
[24] Hassiotis A, Serfaty M, Azam K, Strydom A, Blizard R, Romeo R, et al. Manualised individual cognitive behavioural therapy for mood disorders in people with mild to moderate intellectual disability: a feasibility randomised controlled trial. J Affect Disord 2013;151(1):186–95.
[25] Hastings RP, Hatton C, Lindsay WR, Taylor JL. Psychological therapies for adults with intellectual disabilities. Psychological therapies for adults with intellectual disabilities267; 2013.
[26] Heywood R, Ryan H, Killett A, Langdon P, Plenderleith Y, Shiggins C, et al. Lost voices in research: exposing the gaps in the mental capacity act 2005. Med Law Int 2019;19(2–3):81–112.
[27] Ho P, Downs J, Bulsara C, Patman S, Hill AM. Addressing challenges in gaining informed consent for a research study investigating falls in people with intellectual disability. Br J Learn Disabil 2018;46(2):92–100.
[28] Hulbert-Williams L, Hastings RP. Life events as a risk factor for psychological problems in individuals with intellectual disabilities: a critical review. J Intellect Disabil Res 2008;52(11):883–95.
[29] Jahoda A, Hastings R, Hatton C, Cooper SA, Dagnan D, Zhang R, et al. Comparison of behavioural activation with guided self-help for treatment of depression in adults with intellectual disabilities: a randomised controlled trial. Lancet Psychiatry 2017;4(12):909–19.
[30] Jimoh OF, Ryan H, Killett A, Shiggins C, Langdon PE, Heywood R, et al. A systematic review and narrative synthesis of the research provisions under the

mental capacity act (2005) in England and Wales: recruitment of adults with capacity and communication difficulties. PLoS One 2021;16(9):e0256697.
[31] Joyce T, Globe A, Moody C. Assessment of the component skills for cognitive therapy in adults with intellectual disability. J Appl Res Intellect Disabil 2006;19(1):17–23.
[32] Karatzias T, Brown M, Taggart L, Truesdale M, Sirisena C, Walley R, et al. A mixed-methods, randomized controlled feasibility trial of eye movement desensitization and reprocessing (EMDR) plus standard care (SC) versus SC alone for DSM-5 posttraumatic stress disorder (PTSD) in adults with intellectual disabilities. J Appl Res Intellect Disabil 2019;32(4):806–18.
[33] Klein AM, Salemink E, de Hullu E, Houtkamp E, Papa M, van der Molen M. Cognitive bias modification reduces social anxiety symptoms in socially anxious adolescents with mild intellectual disabilities: a randomized controlled trial. J Autism Dev Disord 2018;48(9):3116–26.
[34] Koslowski N, Klein K, Arnold K, Koesters M, Schuetzwohl M, Salize HJ, et al. Effectiveness of interventions for adults with mild to moderate intellectual disabilities and mental health problems: systematic review and meta-analysis. Br J Psychiatry 2016;209(6):469–74.
[35] Lawrence DH. The effects of reality therapy group counseling on the self-determination of persons with developmental disabilities. United States: Wayne State University; 2003.
[36] Lindsay WR, Baty FJ. Group relaxation training with adults who are mentally handicapped. Behav Cogn Psychother 1989;17(1):43–51.
[37] Lindsay WR, Skene DD. The Beck depression inventory II and the Beck anxiety inventory in people with intellectual disabilities: factor analyses and group data. J Appl Res Intellect Disabil 2007;20(5):401–8.
[38] Lindsay WR, Tinsley S, Beail N, Hastings RP, Jahoda A, Taylor JL, et al. A preliminary controlled trial of a trans-diagnostic programme for cognitive behaviour therapy with adults with intellectual disability. J Intellect Disabil Res 2015;59(4):360–9.
[39] Mann PH. Modifying the behavior of negro educable mentally retarded boys through group counseling procedures. J Negro Educ 1969;38(2):135–42.
[40] Matson JL. A controlled outcome study of phobias in mentally retarded adults. Behav Res Ther 1981;19(2):101–7.
[41] Matson JL, Senatore V. A comparison of traditional psychotherapy and social skills training for improving interpersonal functioning of mentally retarded adults. Behav Ther 1981;12(3):369–82.
[42] McCabe MP, McGillivray JA, Newton DC. Effectiveness of treatment programmes for depression among adults with mild/moderate intellectual disability. J Intellect Disabil Res 2006;50(4):239–47.
[43] McGaw S, Ball K, Clark A. The effect of group intervention on the relationships of parents with intellectual disabilities. J Appl Res Intellect Disabil 2002;15(4):354–66.
[44] McGillivray JA, Kershaw MM. The impact of staff initiated referral and intervention protocols on symptoms of depression in people with mild intellectual disability. Res Dev Disabil 2013;34(2):730–8.
[45] McGillivray JA, Kershaw M. Do we need both cognitive and behavioural components in interventions for depressed mood in people with mild intellectual disability? J Intellect Disabil Res 2015;59(2):105–15.
[46] McGillivray JA, McCabe MP. Pharmacological management of challenging behavior of individuals with intellectual disability. Res Dev Disabil 2004;25(6):523–37.
[47] McGillivray JA, McCabe MP, Kershaw MM. Depression in people with intellectual disability: an evaluation of a staff-administered treatment program. Res Dev Disabil 2008;29(6):524–36.
[48] Munder T, Flückiger C, Leichsenring F, Abbass AA, Hilsenroth MJ, Luyten P, et al. Is psychotherapy effective? A re-analysis of treatments for depression. Epidemiol Psychiatr Sci 2019;28(3):268–74.
[49] National Institute for Health and Care Excellence (NICE). Methods for the Development of NICE Public Health Guidance [NICE Processes and Methods Guidance No. 4], https://www.nice.org.uk/process/pmg4; 2012.
[50] National Institute for Health and Care Excellence (NICE). Mental health problems in people with learning disabilities: prevention, assessment and management [NICE Guidance No. 54], https://www.nice.org.uk/guidance/ng54; 2016.
[51] Norcross JC. An eclectic definition of psychotherapy. In: What is Psychotherapy. 4; 1990. p. 218–20.
[52] Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev 2016;5(1):1–10.
[53] Reiss S, Levitan GW, Szyszko J. Emotional disturbance and mental retardation: diagnostic overshadowing. Am J Ment Defic 1982;86(6):567–74.
[54] Review Manager (RevMan). [Computer program]. Version 5.4. The Cochrane Collaboration; 2020.
[55] Rose JL, Dodd L, Rose N. Individual cognitive behavioral intervention for anger. J Ment Health Res Intellect Disabil 2008;1(2):97–108.
[56] Rose J, Loftus M, Flint B, Carey L. Factors associated with the efficacy of a group intervention for anger in people with intellectual disabilities. Br J Clin Psychol 2005;44(3):305–17.
[57] Rose N, O'Brien A, Rose J. Investigating staff knowledge and attitudes towards working with adults with learning disabilities and mental health difficulties. Adv Ment Health Learn Disabil 2007;1(3):52–9.
[58] Rose J, O'Brien A, Rose D. Group and individual cognitive behavioural interventions for anger. Adv Ment Health Learn Disabil 2009;3(4):45–50.
[59] Rosenthal R. The file drawer problem and tolerance for null results. Psychol Bull 1979;86(3):638–41.
[60] Shechtman Z, Pastor R. Cognitive-behavioral and humanistic group treatment for children with learning disabilities: a comparison of outcomes and process. J Couns Psychol 2005;52(3):322.
[61] Silvestri R. Implosive therapy treatment of emotionally disturbed retardates. J Consult Clin Psychol 1977;45(1):14.
[62] Sturmey P. Against psychotherapy for people who have mental retardation. Ment Retard 2005;43(1):55–7.
[63] Taylor JL, Novaco RW, Gillmer BT, Robertson A, Thorne I. Individual cognitive-behavioural anger treatment for people with mild-borderline intellectual disabilities and histories of aggression: a controlled trial. Br J Clin Psychol 2005;44(3):367–82.
[64] Taylor JL, Novaco RW, Guinan C, Street N. Development of an imaginal provocation test to evaluate treatment for anger problems in people with intellectual disabilities. Clin Psychol Psychother 2004;11(4):233–46.
[65] Tsakanikos E, Bouras N, Costello H, Holt G. Multiple exposure to life events and clinical psychopathology in adults with intellectual disability. Soc Psychiatry Psychiatr Epidemiol 2007;42(1):24–8.
[66] Vereenooghe L, Gega L, Reynolds S, Langdon PE. Using computers to teach people with intellectual disabilities to perform some of the tasks used within cognitive behavioural therapy: a randomised experiment. Behav Res Ther 2016;76:13–23.
[67] Vereenooghe L, Langdon PE. Psychological therapies for people with intellectual disabilities: a systematic review and meta-analysis. Res Dev Disabil 2013;34(11):4085–102.
[68] Vereenooghe L, Reynolds S, Gega L, Langdon PE. Can a computerised training paradigm assist people with intellectual disabilities to learn cognitive mediation skills? A randomised experiment. Behav Res Ther 2015;71:10–9.
[70] Whittle EL, Fisher KR, Reppermund S, Lenroot R, Trollor J. Barriers and enablers to accessing mental health services for people with intellectual disability: a scoping review. J Ment Health Res Intellect Disabil 2018;11(1):69–102.
[71] Williams P. 'It all sounds very interesting, but we're just too busy!': Exploring why 'gatekeepers' decline access to potential research participants with learning disabilities. Eur J Spec Needs Educ 2020;35(1):1–14.
[72] Willner P, Brace N, Phillips J. Assessment of anger coping skills in individuals with intellectual disabilities. J Intellect Disabil Res 2005;49(5):329–39.
[73] Willner P, Jones J, Tams R, Green G. A randomized controlled trial of the efficacy of a cognitive-behavioural anger management group for clients with learning disabilities. J Appl Res Intellect Disabil 2002;15(3):224–35.
[74] Willner P, Rose J, Jahoda A, Kroese BS, Felce D, MacMahon P, et al. A cluster randomised controlled trial of a manualised cognitive behavioural anger management intervention delivered by supervised lay therapists to people with intellectual disabilities. Health Technol Assess (Winch Eng) 2013;17(21):1.
[75] Witwer AN, Rosencrans ME, Held MK, Cobranchi C, Crane J, Chapman R, et al. Psychotherapy treatment outcome research in adults with ID: Where do we go from here? Clin Psychol Sci Pract 2022. https://doi.org/10.1037/cps0000053. Advance online publication.
[76] Wood JA. Methodology for dealing with duplicate study effects in a meta-analysis. Organ Res Methods 2008;11(1):79–95.
[77] Holstead J, Dalton J. Utilization of Trauma-Focused Cognitive Behavioral Therapy (TF-CBT) for children with cognitive disabilities. J Public Child Welfare 2013;7(5):536–48.
[78] Lindsay WR, Allan R, Parry C, Macleod F, Cottrell J, Overend H, et al. Anger and aggression in people with intellectual disabilities: treatment and follow-up of consecutive referrals and a waiting list comparison. Clin Psychol Psychother: An Int J Theory Practice 2004;11(4):255–64.
[79] Mason J, Scior K. Diagnostic overshadowing'amongst clinicians working with people with intellectual disabilities in the UK. J Appl Res Intellect. Disabilit 2004;17(2):85–90.
[80] Department of Health. Mental Capacity Act. London: HMSO; 2005.