

The efficacy of interventions for behaviourally inhibited preschool-aged children: A meta-analysis

Jinnie Ooi^{a,b,*}, Helen F. Dodd^c, Richard Meiser-Stedman^a, Jennifer L. Hudson^d,
Jessica Bridges^a, Laura Pass^a

^a Department of Clinical Psychology and Psychological Therapies, Norwich Medical School, University of East Anglia, Norwich NR4 7TJ, UK

^b Department of Clinical Psychology, Norfolk and Norwich University Hospitals NHS Foundation Trust, Colney Lane, Norwich NR4 7UY, UK

^c College of Medicine and Health, St Luke's Campus, Heavitree Road, University of Exeter, Exeter EX1 2LU, UK

^d Black Dog Institute, University of New South Wales, Hospital Road, Randwick, NSW 2031, Australia

ARTICLE INFO

Keywords:

Behavioural inhibition
Anxiety
Meta-analysis
Intervention
Preschool-aged

ABSTRACT

The current systematic review and meta-analyses examined the efficacy of psychological interventions targeting behavioural inhibition and anxiety in preschool-aged children, evaluated within randomised controlled trials. Web of Science, MEDLINE, PsycINFO and CINAHL were systematically searched from inception to March 2021. Ten studies ($N = 1475$ children, aged 3–7 years) were included in the current review. Separate analyses were conducted for behavioural inhibition, anxiety symptoms, and anxiety diagnosis as reported by parents, teachers, and observer-ratings. Pooled outcomes ranged from post-intervention to 12-month follow-up due to the limited number of studies. Meta-analyses revealed that intervention did not reduce behavioural inhibition as assessed by independent observers ($SMD = -0.13$, 95% $CI = -0.63$ to 0.38), but did reduce behavioural inhibition as reported by parents ($SMC = -0.64$, 95% $CI = -1.00$ to -0.27) and teachers ($SMD = -0.69$, 95% $CI = -1.02$ to -0.36). Additionally, intervention appeared to reduce the risk of anxiety disorders ($RR = 0.75$, 95% $CI = 0.62$ to 0.90), and parent-report anxiety symptoms ($SMC = -0.47$, 95% $CI = -0.83$ to -0.12) in preschool-aged children. Intervention may be efficacious in reducing anxiety in preschool-aged behaviourally inhibited children. It is less clear whether intervention leads to change in BI.

1. Introduction

Behavioural inhibition (BI) is a temperament style characterised by shy, quiet, or restrained behaviours in response to novel, unfamiliar situations (Kagan, Reznick, & Snidman, 1988). Related temperaments include anxious withdrawal (Rubin, Coplan, & Bowker, 2009), shy-inhibited temperament (Prior, Smart, Sanson, & Oberklaid, 2000) and anxious solitude (Gazelle & Ladd, 2003). Behavioural inhibition in the preschool years has been identified as a major risk factor for subsequent anxiety in a number of longitudinal studies (Chronis-Tuscano et al., 2009; Hudson, Murayama, Meteyard, Morris, & Dodd, 2019; Schwartz, Snidman, & Kagan, 1999). A recent meta-analysis concluded that behavioural inhibition in the preschool years was associated with an almost three-fold increase in the odds of developing anxiety subsequently ($OR = 2.80$, 95% $CI = 2.03$ – 3.86) (Sandstrom, Uher, & Pavlova, 2020). Several etiological models of childhood anxiety suggest a central

role for preschool behavioural inhibition (e.g., Liu & Pérez-Edgar, 2019; Rapee, Schniering, & Hudson, 2009; Rubin et al., 2009). For instance, Rapee et al. (2009) argued that behavioural inhibition may elicit and interact with environmental risk factors such as parenting behaviours and parental anxiety disorders in the development of anxiety. Similarly, Rubin et al. (2009) proposed that social withdrawal may elicit difficult peer relationships (e.g., peer victimisation, rejection, exclusion) due to poor social skills, which further increases the likelihood of developing anxiety.

Recent empirical evidence provides support for these predictions. For example, Hudson et al. (2019) found that behaviourally inhibited preschool-aged children experienced greater anxiety symptoms in early adolescence (aged 12) if their mothers were observed to exhibit high levels of overinvolved parenting at age four. Conversely, this elevated risk for anxiety in behaviourally inhibited preschool-aged children was mitigated when their mothers showed low levels of overinvolvement at

* Corresponding author at: Department of Clinical Psychology and Psychological Therapies, Norwich Medical School, University of East Anglia, Norwich NR4 7TJ, UK.

E-mail address: jinnie.ooi@uea.ac.uk (J. Ooi).

<https://doi.org/10.1016/j.janxdis.2022.102559>

Received 17 May 2021; Received in revised form 1 March 2022; Accepted 19 March 2022

Available online 24 March 2022

0887-6185/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

age four. In terms of peer relationships, Frenkel et al., (2015, pp. 12390) demonstrated that behavioural inhibition in childhood interacted with social involvement with peers in adolescence to predict risk for developing anxiety disorders in adulthood. That is, behaviourally inhibited children involved in smaller and less socially active peer networks were at a heightened risk for anxiety disorders in adulthood, compared to their behaviourally inhibited peers who were involved in larger and more socially active peer networks.

Due to the central role that preschool behavioural inhibition plays in the development of subsequent anxiety, intervention and prevention programmes targeting inhibited preschool-aged children have been developed. These aim to prevent (selective programs) or reduce the severity (indicated programs) of anxiety disorders. Interventions (selective and/or indicated programs) that have been developed so far feature two main pathways, in line with the etiological models described above. First, parent education programs (e.g., Cool Little Kids; Rapee, Kennedy, & Lau, 2010) target key parenting behaviours that interact with preschool behavioural inhibition such as overinvolvement and overcontrol/intrusion to ensure that parents promote social approach behaviours and reduce avoidance in their preschool-aged child. The other intervention pathway focuses on working directly with preschool-aged children, focusing on social skills training (e.g., Social Skills Facilitated Play program; Coplan, Schneider, Matheson, & Graham, 2010) with the aim of improving social competence and social participation in behaviourally inhibited children. More recent interventions have also begun to combine both the child-focused and parent-focused approaches (e.g., Turtle Program; Chronis-Tuscano et al., 2015).

There is initial evidence that these interventions might be efficacious in reducing anxiety and/or behavioural inhibition (e.g., Coplan et al., 2010; Kennedy, Rapee, & Edwards, 2009) but positive effects are not consistently found (e.g., Bayer et al., 2018; LaFreniere & Capuano, 1997; Rapee, Kennedy, Ingram, Edwards, & Sweeney, 2005). To date, there has been no systematic synthesis of the efficacy of these interventions for behaviourally inhibited preschool-aged children. Given that the literature on interventions for preschool inhibition is beginning to accumulate, this systematic review aimed to provide a preliminary synthesis on the efficacy of such interventions by systematically evaluating and summarising data from randomised controlled trials of selective and/or indicated psychological interventions for behaviourally inhibited preschool-aged children.

In addition, considering the preliminary nature of the field at present, this systematic review included all interventions targeting preschool behavioural inhibition, regardless of the type of intervention (i.e., parent education, social skills training). In doing so, this review provides an overview of existing interventions for these children and provides a summary of the efficacy of these interventions as well as measures of the heterogeneity between study effect sizes. By summarising the evidence in this way, the review informs the next steps for the field. Additionally, this systematic review focused on interventions targeting preschool behavioural inhibition as a risk-factor, regardless of the preschool-aged children's anxiety disorder status at baseline. This approach is distinct from a previous meta-analysis which examined prevention interventions for children and adolescents at-risk of anxiety (e.g., elevated anxiety symptoms or sensitivity, parent anxiety disorder), excluding trials where participants may already have had an anxiety disorder (Lawrence, Rooke, & Creswell, 2017). In defining efficacy, we were interested not only in whether such interventions lead to a reduction in anxiety but also whether they positively affected behavioural inhibition. Therefore, we examined whether interventions for behaviourally inhibited preschool-aged children are efficacious in reducing (a) behavioural inhibition, and (b) anxiety symptoms and diagnosis.

2. Methods

The protocol for the current meta-analysis was registered on the

International Prospective Register of Systematic Reviews (PROSPERO; protocol number: CRD42020170666) on 25 March 2020.

2.1. Search strategy

We searched four electronic databases (Web of Science, MEDLINE, PsycINFO and CINAHL) from inception to 15 March 2021. Details of the search terms and syntax for each database are available in the PROSPERO protocol (see Supplementary Material 1). No restrictions were imposed for date of publication or language. Reference lists of relevant book chapters, review articles and eligible articles were screened to identify further studies missed by the electronic search.

2.2. Eligibility criteria

Studies were included if they met the following criteria:

1. Participants were preschool-aged children (between 3 and 7 years) and their parents and/or teachers
2. Participants (children) were selected for inclusion on the basis of being behaviourally inhibited, regardless of whether they were identified as having an anxiety disorder or not. Constructs described other than behavioural inhibition (e.g., fearful temperament, shyness/inhibition) were included as long as the definition and measurement of this construct was the same or very similar to behavioural inhibition; which was defined as shyness, fear and avoidance when faced with new stimuli.
3. Reported outcomes using:
 - A validated measure or standardized laboratory observation of behavioural inhibition
 - A recognised diagnostic tool for a DSM-IV or DSM-5 anxiety disorder, or a validated measure of anxiety symptoms
4. Randomised Controlled Trial (RCT) design, comparing an intervention with a waitlist and/or active comparison condition.
5. Included an active intervention which aimed to reduce behavioural inhibition, anxiety symptoms and/or incidence of anxiety disorders in preschool-aged children.
6. Published in a peer-reviewed journal.

Studies without primary data (e.g., reviews) and those that reported qualitative data only were excluded. Additionally, universal interventions (whole populations) and studies that focused on children with intellectual disabilities, neurodevelopmental disorders or specific health conditions were excluded as the current meta-analysis focused on intervention for behaviourally inhibited children from the general population. Only peer reviewed studies were included in this systematic review to ensure that outcomes from predominantly high-quality intervention programmes were included, especially given the relatively novel field of BI interventions.

2.3. Study selection/ screening method

Fig. 1 shows a summary of the search and screening method using a Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) flowchart. Two authors (JO and JB) independently screened all ($n = 8167$) the retrieved titles and abstracts for eligibility. There was a 99.8% agreement on eligibility between raters. Inter-rater reliability on eligibility between raters was substantial, $k = 0.99$. The full texts of eligible studies were then independently reviewed and rated by JO and JB. There was an 88.9% agreement on inclusion between raters. Inter-rater reliability on inclusion between raters was substantial, $k = 0.72$. Disagreements regarding inclusion were resolved by a third member of the research team, LP. Where the same trial was reported in multiple publications (e.g., multiple follow-ups of the same sample), the publication reporting outcomes most relevant to the systematic review were chosen for inclusion to avoid repeated inclusion of data from the same

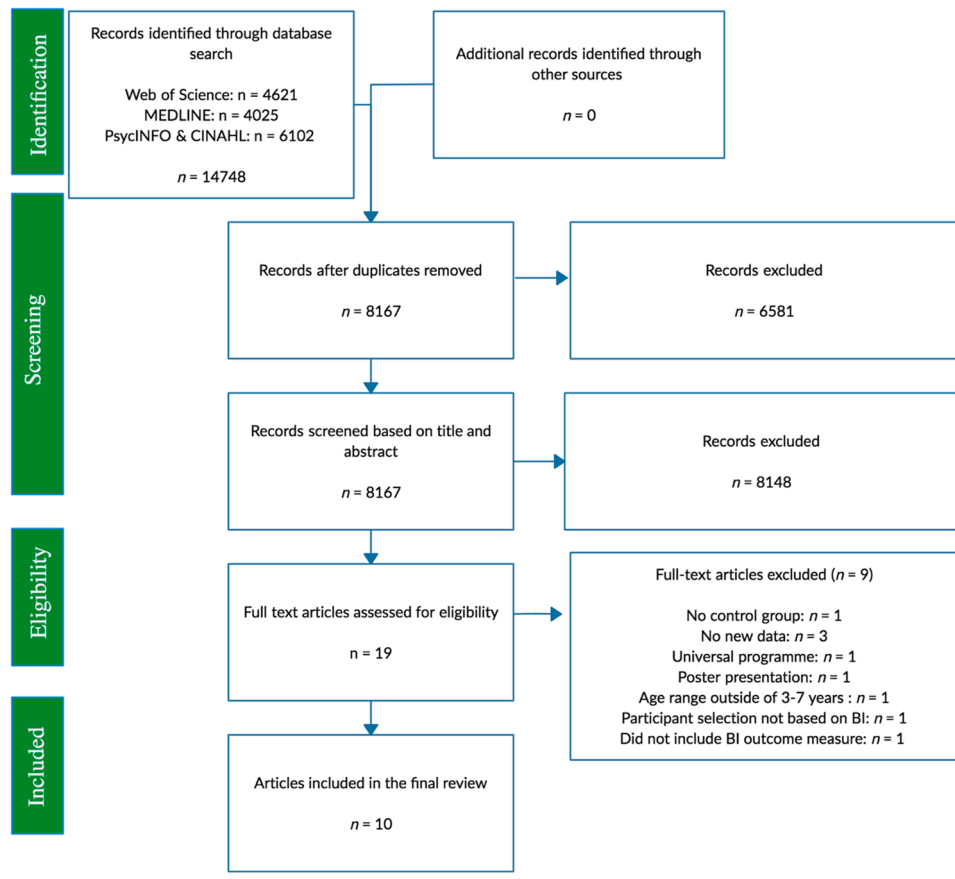


Fig. 1. PRISMA diagram of the study selection process.

Table 1
Characteristics of Included Studies.

Study	N (n Intervention)	Gender % F	M Age (Years) (range)	% Baseline AD Int (Ctrl)	Nature of Risk	Recipient	Intervention Approach	Control Condition	Intervention Target	Intervention Name
Barstead et al. (2018)	40 (18)	56	4.3 (3.5–5.0)	N/A	BI	P + C	PCIT + SST	WL	BI	Turtle Program
Bayer et al. (2018)	545 (265)	48.3	4.6 (4.0)	N/A	BI	P	CBT	UC	AD, AS	Cool Little Kids
Chronis-Tuscano et al. (2015)	40 (18)	57.5	4.4 (3.5–5.5)	77.8 (45.5)	BI	P + C	PCIT + SST	WL	BI, AD, AS	Turtle Program
Coplan et al. (2010)	28 (13)	50	4.7 (4.0–5.5)	N/A	BI	C	SST	WL	BI	SST-FP
Kennedy et al. (2010)	71 (35)	54.5	3.9 (3.0–4.8)	100 (100)	BI + Parent AD	P	CBT	WL	BI, AD, AS	Cool Little Kids
LaFreniere and Capuano (1997)	43 (21)	53.49	4.5 (2.6–5.8)	N/A	BI	P	PCIT	UC	BI	NA
Lau, Rapee, and Coplan (2017)	72 (39)	47.2	4.3 (3.0–5.4)	100 (100)	BI + High PES	P + C	CBT + SST	WL	BI, AD, AS	Cool Little Kids +SST-FP
Luke, Chan, Au, and Lai (2017)	57 (29)	38.6	3.9 (3.0–5.3)	N/A	BI	P	CBT	WL	BI, AS	Cool Little Kids
Morgan et al. (2017)	433 (215)	52.7	4.8 (3.0–6.0)	N/A	BI	P	CBT	WL	AS	Cool Little Kids Online
Rapee et al. (2005)	146 (73)	54.5	3.9 (3.0–5.2)	90.0 (91.5)	BI	P	CBT	UC	BI, AD, AS	Cool Little Kids

% Baseline AD [Int (Ctrl)]: % Baseline Anxiety Diagnosis [Intervention (Control)]; **Nature of risk**: BI = Elevated Behavioural Inhibition, High PES = High Parental Emotional Distress [at least one parent scoring ≥ 30 on the Depression Anxiety Stress Scales (Lovibond & Lovibond, 1995)], Parent AD = at least one parent meeting DSM-IV criteria for diagnosis of anxiety disorder; **Recipient**: C = child, P = parent; **Intervention Approach**: CBT = Cognitive Behavioural Therapy, SST = Social Skills Training, PCIT = Parent-Child Interaction Training; **Control Condition**: WL = Wait-List control, UC = Usual Care; **Intervention Target**: BI = Behavioural Inhibition, AD = Anxiety Diagnosis, AS = Anxiety Symptoms; **Intervention Name**: Cool Little Kids (Rapee & Kennedy, 2010), SST-FP = Social Skills Training and Facilitated Play Program (Coplan et al., 2010), Turtle Program (Danko, O'Brien, Rubin, & Chronis-Tuscano, 2018).

participants.

2.4. Data extraction

Data were extracted and coded by JO. To ensure accuracy, all the studies were cross-checked by JB, resulting in no disagreement. Information extracted were: a) study characteristics (e.g., year of publication, study location: country), b) sample characteristics (e.g., N, age, nature of risk), c) intervention characteristics and control condition (e.g., intervention recipient: child and/or parent, intervention type: parenting and/or social skills training, waitlist/care as usual), d) primary outcome data for BI, and e) secondary outcome data for anxiety diagnosis and/or symptoms (e.g., name of BI/anxiety outcome measures, respondent, percentage or Ms and SDs for each condition at post-intervention and/or follow-ups). See Tables 1 and 2 for characteristics of the included studies and summary of outcome measurement respectively. Study authors were contacted where there were insufficient data for calculating an effect size.

2.5. Assessment of study quality

Study quality was assessed using the quality assessment instrument developed by Moncrieff, Churchill, Drummond and McGuire (2001). Following consultation, this assessment tool was deemed more appropriate for the purpose of this systematic review than Kmet, Lee, and Cook (2004) Standard Quality Assessment Criteria for Evaluating Primary Research Papers for Quantitative Studies, as initially indicated in our PROSPERO protocol. This change in the quality assessment tool has been updated on PROSPERO. The Moncrieff, Churchill, Colin Drummond, and McGuire (2001) instrument was developed specifically to assess the quality of controlled trials for mental health interventions. The scale assesses specific methodological issues relevant to mental health interventions, such as clear operationalisation of the nature of the

mental health condition, including severity. The scale consists of 23 items which are rated between 0 and 2, generating a total score ranging between 0 and 46; higher scores suggests greater quality for studies. To check for reliability, JO and LP rated all the studies. Percentage agreement for the individual items in the scale was 88.10%. Inter-rater reliability for total quality score between raters was good, $k = .80$.

2.6. Data synthesis

Analyses were performed using the metafor package (Viechtbauer, 2010) in R 3.4.2 (R Core Team, 2013). Random effects models using Restricted Maximum Likelihood (REML) were used to account for the expected heterogeneity in effect sizes between trials due to the diversity in type of interventions trialled, target populations, type of measurements used, and duration of measurement (i.e., post-intervention up to 12-months follow-up).

For continuous outcome measures (i.e., behavioural inhibition and anxiety symptoms), when pre-test data was available for all trials, standardized mean change (SMC) effect sizes were calculated for each trial by subtracting the mean pre- to post-test change in the control group from the mean pre- to post-test change in the intervention group, then dividing this difference by the pooled standard deviation for the control and intervention groups at pre-test. Given that the pre- and post-test correlations for the intervention and control groups were not available, approximate values for the correlations ($r = 0.00, .45, .90$) were used as substitutes. Sensitivity analyses were then conducted to ensure that outcomes of the meta-analyses were comparable when the correlations were varied (Morris, 2007). Outcomes using a correlation of $r = 0.45$ were reported in Results, while those using correlations of $r = 0.00$ and $.90$ were reported in Supplementary Materials 2. However, when pre-test data were not available for one or more trials, standardized mean differences (SMD) were calculated for each trial by subtracting the mean of the intervention condition from the mean of the

Table 2
Outcome measures for temperament-related behaviours and duration of measurement.

Study	BI Screening Measure	Respondent for BI screening	Temperament-related Outcome Measures	Respondent of Temperament-related Outcomes	Duration of Measurement (months follow-up)
Barstead et al. (2018)	BIQBI cut-off: score of 132 or more (85th percentile and above)	P	Reticence (POS-Revised) Anxious-Fearful subscale (CBS)	CT	Post-intervention
Bayer et al. (2018)	Approach subscale of the STSCBI cut-off: score of 30 and above	P	N/A	N/A	N/A
Chronis-Tuscano et al. (2015)	BIQBI cut-off: score of 132 or more	P	BIQ	P	Post-intervention
Coplan et al. (2010)	BIQBI cut-off: highest 15%	P	Reticence-wariness (POS) Anxious-Fearful subscale of the CBS	CT	Post-intervention
Kennedy et al. (2010)	Approach subscale of the STSCBI cut-off: score of higher than 35 + Laboratory Observation	P + C	Laboratory Observation BIQ	CP	6
LaFreniere and Capuano (1997)	Anxiety-Withdrawal subscale of the SCBEBI cut-off: 1 SD above mean	T	Anxiety-Withdrawal subscale of the SCBE	T	Post-intervention
Lau et al. (2017)	Approach subscale of the STSCBI cut-off: score of 30 and above	P	BIQ	P	6
Luke et al. (2017)	BIQBI cut-off: highest 20%	T	Anxious Shyness subscale of the Chinese Shyness Scale(CSS-AS)	T	3
Morgan et al. (2017)	Approach subscale of the STSCBI cut-off: score of 30 and above	P	N/A	N/A	N/A
Rapee et al. (2005)	Approach subscale of the STSCBI cut-off: score of higher than 30 + Laboratory Observation	P + C	Laboratory ObservationSocial Inhibition subscale of the TABC-R	CP	12 24

BI screening measure: STSC = Short Temperament Scale for Children (Prior et al., 2000), BIQ = (Bishop, Spence, & McDonald, 2003), SCBE = Social Competence and Behaviour Evaluation (LaFreniere & Dumas, 1995), Laboratory Observation (Asendorpf, 1990; Kagan, 1994; Kagan, Reznick, & Gibbons, 1989; Kagan, Reznick, Clarke, Snidman, & Garcia-Coll, 1984); **Respondent for BI screening:** P = Parent, T = Teacher, C = Clinician; **BI-related Outcome Measures:** CSS = Chinese Shyness Scale (Xu, Farver, Yu, & Zhang, 2009), BIQ = (Bishop et al., 2003), POS = Play Observation Scale (Rubin, 2001), POS-Revised = Play Observation Scale - Revised (Rubin, 2008), laboratory observation (Kagan et al., 1989; Kagan, 1994); TABC-R = Temperament Assessment Battery for Children - Revised (Presley & Martin, 1994), CBS = Child Behaviour Scale (Ladd & Profilet, 1996), SCBE = Social Competence and Behaviour Evaluation (LaFreniere & Dumas, 1995); **Respondent of BI-related Outcomes:** T = Teacher, P = Parent, C = Clinician; N/A = Information not available.

control condition at post-intervention/follow-up, divided by the pooled standard deviation for the intervention and control conditions at post-intervention/follow-up. Pre-test data were available for all analyses except for laboratory observations of BI and teacher-report BI.

To calculate the pooled SMC/SMDs, the SMC/SMD and the 95% confidence interval for each trial was weighted according to sample size using random effects models. Pooled SMC/SMDs were reported using Hedges' *g*, with 0.249, 0.409, and 0.695 indicating small, moderate, and large effects respectively (Rubio-Aparicio, Marin-Martinez, Sanchez-Meca, & Lopez-Lopez, 2018). Results in favour of intervention groups were represented as a negative effect size.

For diagnostic outcome measures (i.e., anxiety disorder diagnosis), risk ratios (RR) were calculated and pooled. RR represents the likelihood that an outcome (diagnosis of one or more anxiety disorders) would occur in the intervention group, compared to the likelihood of the outcome occurring in the control group. As such, a RR of 1 suggests that the likelihood for a diagnosis of anxiety disorder are the same for both the treatment and control groups.

Estimates of heterogeneity were calculated using the Q statistic, the *I*² statistic, and the prediction interval (95% PI). A statistically significant Q statistic (*p* < .05) suggests evidence of heterogeneity. The *I*² statistic quantifies the degree of heterogeneity, with 25% indicating 'low', 50% indicating 'moderate', and 75% indicating 'high' heterogeneity (Higgins, Thompson, Deeks, & Altman, 2003). A prediction interval estimates the dispersion of the true effect sizes in similar studies that might be conducted in the future. As such, a 95% prediction interval implies that in 95% of the cases, the true effect size of a new study would fall within the predicted interval (Higgins, Thompson, & Spiegelhalter, 2009).

For primary outcomes, three meta-analyses were conducted to examine the pooled effects of interventions on BI-related behaviours, assessed using (1) laboratory observations, (2) parent-report and (3) teacher-report. Next, secondary outcomes on the pooled effects of interventions on anxiety were assessed by conducting two meta-analyses: (1) the presence of an anxiety disorder, and (2) parent-report measures of anxiety symptoms. Only two eligible studies assessed teacher-report anxiety symptoms (Chronis-Tuscano et al., 2015; Luke et al., 2017); the SMDs for each study were reported but the pooled effects were not explored, given that the type of intervention and outcome measures used were different. Similarly, as only two studies (Bayer et al., 2018; Rapee et al., 2005) reported outcomes beyond 12-month follow-up, SMDs and RRs for each study were reported without examining pooled effects. Moderation analyses were not explored due to the limited number of studies in the meta-analyses.

Effect sizes were included for the available outcome measures within the relevant meta-analysis. Where more than one outcome measure of a single outcome was included (e.g., two parent-report measures of anxiety), the primary outcome measure or the one most widely used in other studies, or with the strongest psychometric properties, was chosen. For parental measures, if paternal- and maternal-report measures were reported separately, the maternal-report measure was used to facilitate pooling of effects across studies; most studies included in this meta-analysis had mothers as the primary reporters. If more than one time-point was reported, data from the latest time-point was used as we were interested in the intervention effects over a sustained period of time. Given the limited number of studies included, it was not possible to conduct separate analyses for specific follow-up periods, which means that the outcome ranges across studies from post-intervention to 12-month follow-up.

3. Results

3.1. Study selection

Overall, 8167 studies were identified, and 10 studies met inclusion criteria (see Fig. 1). Three studies reported on BI-related behaviours only

(Barstead et al., 2018; Coplan et al., 2010; LaFreniere & Capuano, 1997), while two studies reported on anxiety only (Bayer et al., 2018; Morgan et al., 2017). The remaining five studies reported on both outcomes (Chronis-Tuscano et al., 2015; Kennedy et al., 2009; Lau et al., 2017; Luke et al., 2017; Rapee et al., 2005).

3.2. Study characteristics

Table 1 summarises the characteristics of all the studies included in the meta-analyses. The total number of participants from the included studies were 1475. Table 2 describes BI screening measures and outcome measures for BI-related behaviours, Table 3 describes outcome measures for anxiety diagnosis and symptoms.

Most (*κ* = 8) of the included studies selected preschool-aged children based on their elevated BI only, while two studies selected for preschool-aged children with elevated BI and parental mental health difficulties. Screening for elevated BI was done predominantly using two measures: the Behavioral Inhibition Questionnaire (BIQ; Bishop et al., 2003) (*κ* = 4) and the Approach subscale of the Short Temperament Scale for Children (STSC; Prior et al., 2000) (*κ* = 5). The cut-off scores used for screening elevated BI varied between studies, even when the same screening measure was used. For the BIQ (Bishop et al., 2003), three studies selected for preschool-aged children scoring on the 85th percentile and above, while one study used a lower cut-off on the 80th percentile and above. For the Approach subscale of the STSC (Prior et al., 2000), four studies used a cut-off score of 30 and above, while one study used a higher cut-off score of 35 and above.

With regards to the type of interventions, six studies evaluated parent education programs: *k* = 5 for Cool Little Kids (Rapee et al.,

Table 3
Outcome Measures for Anxiety and Duration of Measurement.

Study	Anxiety Diagnostic Tool (Respondent)	Anxiety Symptoms Measure (Respondent)	Duration of Measurement (months)
Barstead et al. (2018)	N/A	N/A	N/A
Bayer et al. (2018)	ADIS-IV-P (Clinician)	Emotional Symptoms subscale of SDQ (Parent)	12 24
Chronis-Tuscano et al. (2015)	PAPA (Clinician)	PAS (Parent)	Post-intervention
Coplan et al. (2010)	N/A	N/A	N/A
Kennedy et al. (2009)	ADIS-IV-P (Clinician)	PAS-R (Parent)	6
LaFreniere and Capuano (1997)	N/A	N/A	N/A
Lau et al. (2017)	ADIS-IV-P (Clinician)	PAS-R (Parent)	6
Luke et al. (2017)	N/A	Internalising construct of CBS (Teacher)	3
Morgan et al. (2017)	OAPA (Parent)	PAS-R (Parent)	36
Rapee et al. (2005)	ADIS-IV-P (Clinician)	PAS (Parent)	12 24 36 132

Anxiety Diagnostic Tool: ADIS-IV-P = Anxiety Disorders Interview Schedule for DSM-IV Parent version (Any Anxiety Disorders) (Silverman & Albano, 1996), OAPA = Online Assessment of Preschool Anxiety (Any Anxiety Diagnosis) (Morgan et al., 2019), PAPA = Preschool Age Psychiatric Assessment (Any Anxiety Diagnosis) (Egger, Ascher, & Angold, 1999); **Anxiety Symptoms Measure:** SDQ = Strength and Difficulties Questionnaire (Goodman, 2001), PAS-R = Preschool Anxiety Scale Revised (Edwards, Rapee, Kennedy, & Spence, 2010), CBS = Child Behaviour Scale (Ladd, 2010), PAS = Preschool Anxiety Scale (Spence, Rapee, McDonald, & Ingram, 2001); N/A = Information not available.

2010), $k = 1$ for Parent-Child Interaction Training (LaFreniere & Capuano, 1997). One study evaluated a social skills training program: Social Skills Training and Facilitated Play (SST-FP; Coplan et al., 2010). Finally, three studies evaluated programs which combined both parent education and social skills training: $k = 2$ for Turtle Program (Danko et al., 2018), $k = 1$ for combination of the Cool Little Kids and the SST-FT programs. Parents were the primary recipients for parent education programs, while preschool-aged children were the primary recipients for social skills training programs.

Additionally, the duration of measurement also varied across studies (see Tables 2 and 3). Four studies reported post-intervention data only. For follow-ups, only one study provided data for 3-month follow-up, while three studies reported 6-month follow-up data as their latest time-point. Out of the two studies that reported data for 12-month follow-up, one study reported mid- and longer-term follow-up periods (i.e., 2-year, 3-year, and 11-year follow-ups) (Rapee, 2013; pp. 8, 1204; Rapee et al., 2005; pp. 8, 1204; Rapee, Kennedy, Ingram, Edwards, & Sweeney, 2010; pp. 8, 1204), while the remaining study recently reported their 2-year follow-up data (Bayer et al., 2020). Due to the limited duration of measurement reported in the other studies in this review, only the 12-month follow-up data from both the Rapee et al. (2005) and Bayer et al. (2018) studies were included in the current meta-analyses. Subsequent follow-ups of these studies were discussed qualitatively.

In terms of outcome measures, the measures used to assess temperament-related outcomes at post-intervention/follow-ups were varied between studies. Out of the four studies that conducted laboratory observations, two studies used the Reticence/Reticence-Wariness scores from the Play Observation Scale (POS & POS-R; Rubin, 2008) while the remaining two studies used the procedure developed by Kagan and colleagues (Kagan et al., 1989; Kagan, 1994). For parent-reported temperament-related outcomes, three out of the four studies used the BIQ (Bishop et al., 2003), while one study used the Social Inhibition subscale of the Temperament Assessment Battery for Children – Revised (Presley & Martin, 1994). Similarly, for teacher-reported temperament-related outcomes, two out of the four studies used the Anxious-Fearful subscale of the Child Behaviour Scale (CBS; Ladd & Profilet, 1996), while each of the two remaining studies used the Anxiety-Withdrawal subscale of the Social Competence and Behaviour Evaluation (SCBE; LaFreniere & Dumas, 1995) and the Anxious Shyness subscale of the Chinese Shyness Scale (Xu, Farver, Chang, Zhang, & Yu, 2007; CSS; 2009).

There was greater consistency across studies in the outcome measures used to assess anxiety at post-intervention/follow-ups. For anxiety diagnosis, the majority of studies ($k = 4$) used the Anxiety Disorders Interview Schedule for DSM-IV Parent version (ADIS-IV-P; Silverman & Albano, 1996), while each of the remaining two studies used the Online Assessment of Preschool Anxiety (OAPA; Morgan et al., 2019) and the Preschool Age Psychiatric Assessment (PAPA; Egger, Ascher & Angold, 1999 respectively). When anxiety symptoms were reported by parents, five out of the six studies used the Preschool Anxiety Scales (PAS & PAS-R; Edwards et al., 2010; Spence et al., 2001).

3.3. The effect of intervention on preschool-aged children's behavioural inhibition (post-intervention to 12-month follow-up)

For laboratory observations of BI, there was a non-significant effect of intervention (SMD = -0.13 , 95% CI = -0.63 to $.38$, 95% PI = -1.09 to $.83$, $p = .62$, $k = 4$) (See Fig. 2A). Statistical heterogeneity in effect sizes across studies was moderate ($Q = 9.43$, $p = .02$, $I^2 = 68.5\%$). In contrast, parent-report measures showed a significant moderate effect of intervention (SMC = -0.64 , 95% CI = -1.00 to -0.27 , 95% PI = -1.17 to -0.11 , $p < .01$, $k = 4$). Heterogeneity between studies in this analysis was low ($Q = 4.40$, $p = .22$, $I^2 = 27.77\%$) (See Fig. 2B). Finally, teacher-report measures showed a significant moderate effect of intervention (SMD = -0.69 , 95% CI = -1.02 to -0.36 , 95% PI = -1.02 to -0.36 ,

$p < .001$, $k = 4$). Statistical heterogeneity between studies in this analysis was low ($Q = 1.48$, $p = .69$, $I^2 = 0.00\%$) (See Fig. 2C).

3.4. The effect of intervention on preschool-aged children's anxiety (post-intervention to 12-month follow-up)

Intervention significantly reduced behaviourally inhibited preschool-aged children's risk of anxiety disorders (RR = 0.75 , 95% CI = 0.62 to $.90$, 95% PI = 0.54 to 1.04 , $p < .01$, $k = 5$). Statistical heterogeneity between studies was moderate ($Q = 7.81$, $p = .10$, $I^2 = 46.44\%$) (See Fig. 3A).

In contrast, parent-report anxiety symptom measures showed a significant moderate effect of intervention (SMC = -0.47 , 95% CI = -0.83 to -0.12 , 95% PI = -1.28 to $.34$, $p < .01$, $k = 6$). Statistical heterogeneity between studies was moderate ($Q = 12.25$, $p = .03$, $I^2 = 68.34\%$) (See Fig. 3B). For teacher-report anxiety symptoms, only two studies were identified. As such, results from these studies were not included in a meta-analysis and effect sizes of the individual studies were described instead. Chronis-Tuscano et al. (2015) reported a moderate effect size in favour of intervention (Hedges' $g = -0.64$, 95% CI = -1.32 to 0.05). In contrast, Luke et al. (2017) reported a small effect size in the opposite direction to intervention (Hedges' $g = 0.07$, 95% CI = -0.45 to $.59$).

3.5. The effect of intervention on mid- and longer-term outcomes (beyond 12-month follow-up)

Two studies included in this systematic review reported outcomes beyond 12-month follow-up (Bayer et al., 2018; Rapee et al., 2005). These outcomes were described qualitatively given that they were not the primary focus of the current systematic review, and due to the limited number of studies available.

Bayer and colleagues (2020) recently reported outcomes from their 2-year follow-up, when the children were approximately aged 6. There was a small effect of intervention for both anxiety diagnosis (RR = 0.88) and parent-report anxiety symptoms (Hedges' $g = -0.18$). For comparison, Rapee and R. M (2010) reported larger effects for anxiety diagnosis (RR = 0.55) but similar effects for parent-report anxiety symptoms (Hedges' $g = -0.17$, small effect) at 2-year follow-up.

Rapee and colleagues monitored the cohort of behaviourally inhibited preschool-aged children for 11 years, up to middle adolescence when they were approximately 15 years old (Rapee, 2013; pp. 8, 1204; Rapee et al., 2005; pp. 8, 1204; Rapee et al., 2010; pp. 8, 1204). For anxiety diagnosis, the risk of being diagnosed with an anxiety disorder for preschool-aged children in the intervention group, compared to those in the control group decreased from 1-year follow-up (RR = 0.81) to 3-year follow-up (RR = 0.58 , medium effect). In terms of parent-report anxiety symptoms, the effect of intervention increased from a small effect at 1-year follow-up (Hedges' $g < 0.01$) to a medium effect at 3-year follow-up (Hedges' $g = -0.45$). At 11-year follow-up, girls in the intervention group were less likely to be diagnosed with internalising disorders (anxiety and depressive) and have lower parent-report anxiety symptoms than those in the control group (Rapee, 2013, pp. 12048). On the other hand, behavioural inhibition remained largely comparable over time.

3.6. Quality ratings

Scores from the Moncrieff et al. (2001) quality rating indicated common methodological problems in the design and reporting of studies. Likely due to limited resources, half of the studies ($k = 5$) were rated as having inadequate sample size (n per group < 50). Similarly, only half of the studies ($k = 5$) were rated as having a trial duration (including follow-up) that was long enough to assess longer-term outcomes (≥ 6 months). In terms of reporting, only 3 out of the 10 studies reported details of the power calculation, while 4 out of the 10 studies explicitly reported 'intention to treat' analyses. Additionally, only 3

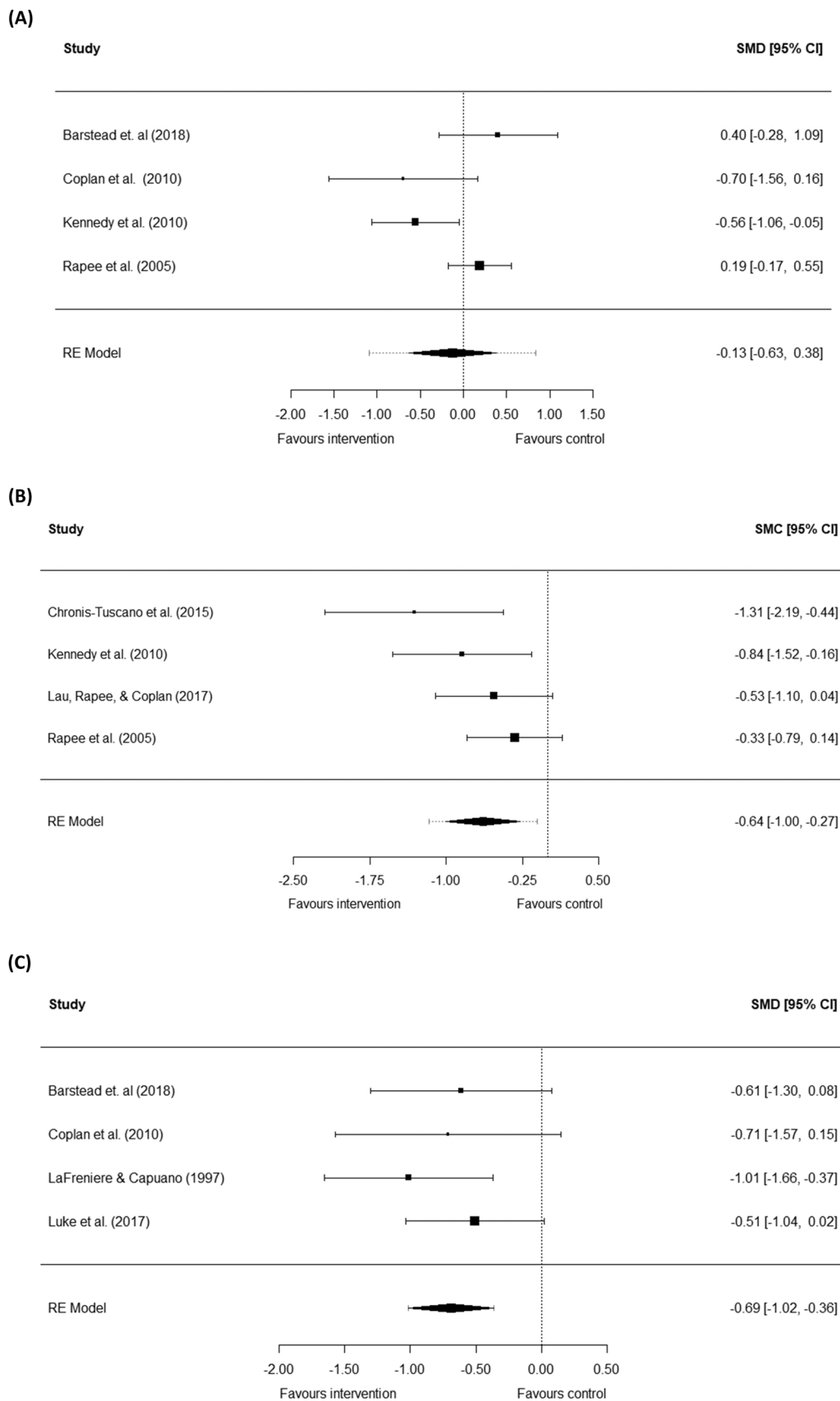


Fig. 2. Forest plot of the effect of intervention on young children’s behavioural inhibition. **Note:** A: Laboratory observations; B: Parent-report measures; C: Teacher-report measures.

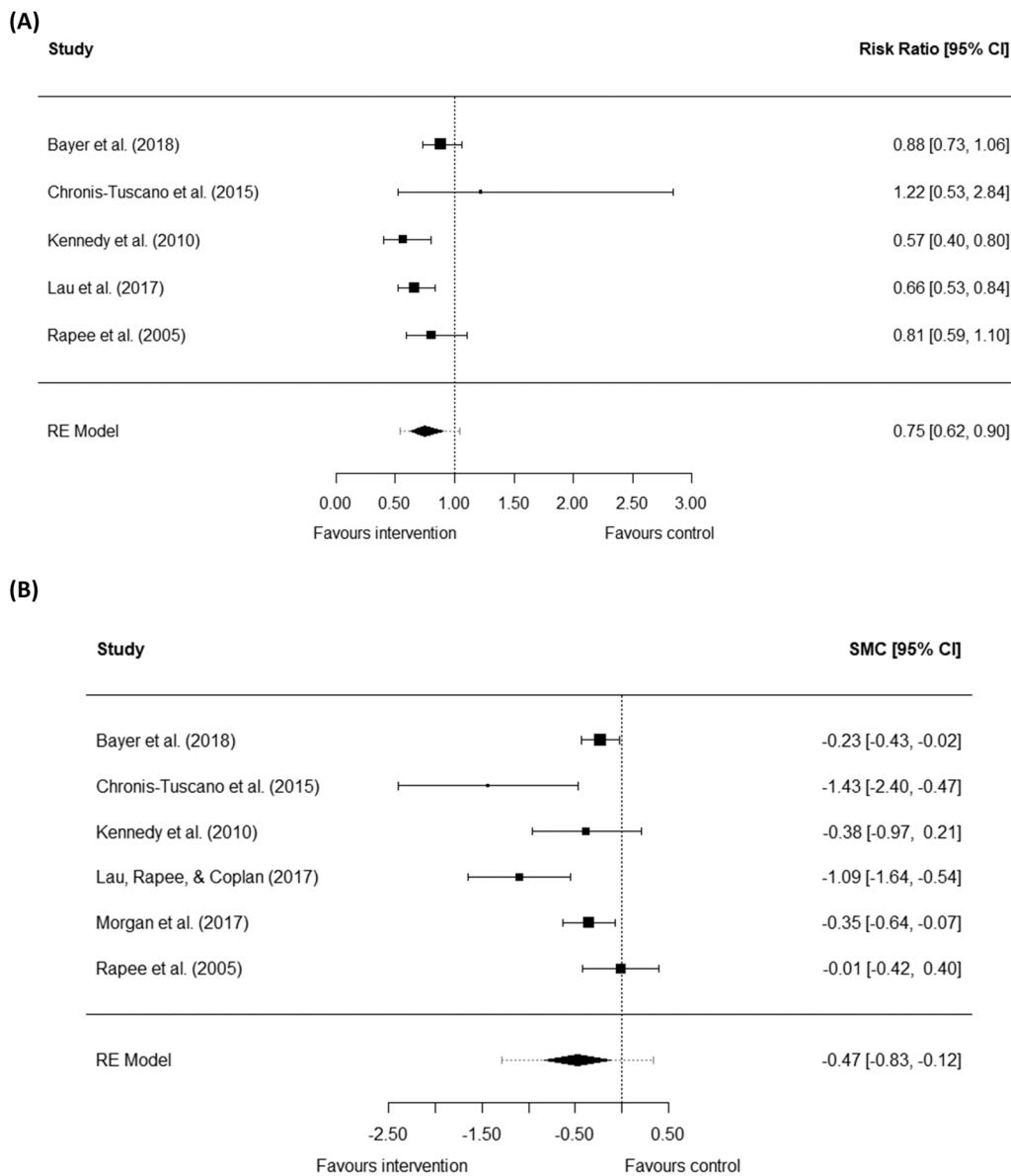


Fig. 3. Forest plot of the effect of interventions on young children’s anxiety. **Note:** A: Diagnosis of Anxiety Disorder; B: Parent-report measures of anxiety symptoms.

studies reported the number of withdrawals by group, including the reason for withdrawal, while the remaining 7 studies reported on the number of withdrawals only, without reporting on the reason for withdrawal. Correlations between the quality rating and study effect size was not explored due to the limited number of studies included in this review.

3.7. Publication bias

Given the limited number of studies included in the analyses, it was not possible to reach firm conclusions about publication bias.

4. Discussion

The current meta-analysis aimed to provide a preliminary synthesis on the effectiveness of intervention for behaviourally inhibited preschool-aged children. Firstly, we examined whether intervention was efficacious in reducing behavioural inhibition, as reported by the following informants: (a) laboratory observers (b) parents, and (c) teachers. Findings on the efficacy of such interventions were mixed.

Observer-ratings of structured laboratory observations ($SMD = -0.13$) indicated a non-significant effect of intervention. In contrast, parent-report ($SMC = -0.64$) and teacher-report ($SMD = -0.69$) measures of behavioural inhibition both showed significant moderate effects of intervention, in favour of the intervention conditions. Overall, although intervention did not reduce behavioural inhibition when assessed by laboratory observers, it was efficacious in reducing behavioural inhibition when reported by parents and teachers.

Next, we explored whether intervention was effective in reducing anxiety disorders and anxiety symptoms as reported by parents and teachers. Compared to controls, intervention significantly decreased the risk of anxiety disorders by 25% in the intervention group ($RR = 0.75$). Additionally, parents reported a significant, moderate reduction in anxiety symptoms in the intervention conditions, compared to the control conditions ($SMD = -0.47$). Given that only two studies included teacher-report measures, the effect size of each study was described. While Chronis-Tuscano et al. (2015) reported a moderate effect of teacher-report anxiety symptoms in favour of the intervention condition (Hedges’ $g = -0.64$), Luke et al. (2017) found a very small effect favouring the control condition (Hedges’ $g = 0.07$). In summary,

intervention appeared to reduce the risk of anxiety disorders, as assessed by laboratory observers, and parent-report anxiety symptoms in preschool-aged children. However, evidence on teacher-report anxiety symptoms is currently limited.

4.1. Conceptual and clinical implications

As noted previously, studies in this meta-analysis only reported outcomes between post-intervention and 1-year follow-up. It is therefore important to stress that findings should be interpreted as *short-term* outcomes of intervention for inhibited young children. Accordingly, the conceptual and clinical implications should be interpreted with this limited interval in mind.

Findings revealed that some aspects of preschool behavioural inhibition may be more amenable than previously thought (Buss & Plomin, 1984; Kagan, 1994), which is consistent with longitudinal evidence that temperament fluctuates across development (Pérez-Edgar & Fox, 2005; Sanson, 1996). However, the evidence is not yet compelling. Although parents and teachers reported a reduction in preschool-aged children's behavioural inhibition following intervention, this change was not observed in structured laboratory observations.

Current evidence supports the risk/vulnerability model (Rapee & Coplan, 2010), in which temperament is considered distinct from psychopathology and affects a child's likelihood of developing an internalising disorder (Dodd, Hudson, & Rapee, 2017). Rapee and Bayer (2018) argued that interventions may be altering the more transient expression of anxiety, while temperamental inhibition remains unchanged. Based on our findings, it is possible that the reductions observed in parent- and teacher-report measures of inhibition reflected changes in preschool-aged children's expression of anxiety. Meanwhile, the lack of evidence for changes in behavioural inhibition based on laboratory observations may indicate that true inhibition remained unchanged by intervention. Alternatively, it is possible that the effects of intervention were not substantial enough in the current meta-analysis to meet the high threshold for detecting significant change using structured laboratory observations (e.g., Kagan, 1994; Kagan et al., 1989), which is typically considered the 'gold standard' for assessing inhibition due to its methodological rigour. Parent- and teacher-report measures, on the other hand, may be able to detect more subtle changes in certain features of inhibition that were altered by intervention. Additionally, it is also possible that changes in inhibition may be more apparent in familiar contexts where children feel relatively comfortable. Therefore, such changes may be more observable to parents and teachers. In contrast, children with a history of inhibition may revert to more typical ways of responding in unfamiliar contexts, such as in laboratory observations. Finally, findings on parent-reported changes in inhibition should be interpreted with caution given that it was not possible to keep parents blinded from the condition that their children were assigned to due to the nature of the interventions (e.g., parenting education vs waitlist control). However, the concordant evidence from independent sources (i.e., parents and teachers) on the effect of intervention is encouraging, especially given that teachers in all the studies were unaware of the children's condition allocations.

This meta-analysis also demonstrated that intervention was effective in decreasing the risk of having an anxiety disorder diagnosis, as well as the severity of parent-report anxiety symptoms in behaviourally inhibited preschool-aged children. These findings are encouraging given that the effects of intervention were observable at preschool-aged children's diagnosis status, as well as at the symptom severity level, at least within the duration measured in this meta-analysis (post-intervention to 12-month follow-up).

Beyond the short-term perspective (up to 12-month follow-up) explored in this meta-analysis, studies reporting mid- and longer-term outcomes (Bayer et al., 2018; Rapee et al., 2005) indicate that at least for anxiety, intervention may continue to yield benefits later on. However, behavioural inhibition remained largely unchanged over time.

Rapee and colleagues (2010) postulated that the preventive effects of their intervention on anxiety could be mediated by other factors such as reductions in life interference resulting from changes in parenting behaviours, as observed in Kennedy et al. (2009), instead of reductions in BI as initially expected. Alternatively, the intervention (i.e., Cool Little Kids) may serve as a treatment for anxiety given that it was based on a cognitive behavioural therapy (CBT) programme for child anxiety, and therefore have an effect on anxiety symptoms and diagnoses, but not behavioural inhibition.

For a meaningful consideration of clinical implications, it is important to recognise that findings from this review are limited to short-term outcomes and may well underestimate the overall effects of intervention. This is reflected in the quality rating of studies in this review where half of the studies ($\kappa = 5$) measured outcomes (including follow-up) for less than 6 months. At best, the evidence is tentative and preliminary, and interpretation requires the consideration that this effort is an encouraging first step to a longer-term endeavour in examining the efficacy of intervention for preschool-aged children at risk of developing anxiety. Therefore, longer-term follow-up of interventions is needed to inform the longitudinal clinical implications of intervention, although such an approach raises ethical considerations about withholding treatment from children in the control condition for a protracted period of time. There is also debate about whether intervention should focus on changing inhibited temperament given that approximately half of inhibited preschool-aged children do not go on to develop anxiety later on (Fox, Barker, White, Suway, & Pine, 2013, pp. 12101). Additionally, evidence for population level intervention is currently limited. A recent population-delivered parenting intervention found modest participation from parents, with only 29.4% of eligible parents attending most sessions offered and only 20.5% of parents reporting using the skills with their children frequently in the first year following intervention (Bayer et al., 2018). These findings suggest that such interventions, at least the parenting programmes, could be more suitable as treatment options for families actively seeking help to prevent anxiety in their preschool-aged children rather than as population level prevention programmes.

An additional limitation of this meta-analysis was that specific factors that impact on the efficacy of intervention could not be explored due to the limited number of studies currently available in the literature. As such, exploration of methodological heterogeneity (e.g., nature of risk: severity of behavioural inhibition, parental mental health, type of intervention: parenting and/or social skills training, recipient of intervention: parent and/or child, duration of outcome measurement: post-intervention, mid- and longer-term follow-ups) through moderation and subgroup analyses could not be carried out. Moreover, scores from the quality rating of the studies also highlighted common methodological problems that might impact on the findings of this review. For instance, half of all studies ($\kappa = 5$) were rated as having an inadequate sample size, which may result in limited statistical power. Indeed only 3 out of the 10 studies reported details of the power calculation. Finally, due to the limited number of studies currently available, the potential effects of publication bias could not be estimated. As such, adjustment for potential overestimation of treatment effect estimates resulting from small-study effects was not possible.

Nevertheless, a key strength of this systematic review is the summary of heterogeneity between study effect sizes. Given the preliminary nature of the field currently, this summary of evidence highlights areas of improvements for future efforts, leading to key recommendations as outlined below.

4.2. Future directions

The findings of this review lead to key recommendations for further intervention research. First, there was substantial variation across studies on how preschool behavioural inhibition was defined and measured. The field would benefit from bringing together the various strands of research that examine constructs associated with inhibited

temperament, including behavioural inhibition, anxious-withdrawal, shy-inhibited, and anxious solitude. Improving consensus on the definition of inhibited temperament would promote greater consistency in the measurement of inhibition, ideally arriving with a set of mutually agreed multimethod assessment tool (i.e., structured lab observations, parent- and teacher report measures) that can be used across the board (Rapee R.M, 2010), in line with recent efforts by the National Institute of Mental Health (NIMH) and the Wellcome Trust calling for greater consensus on outcome measurement in mental health research (The Lancet Psychiatry, 2020).

Second, outcomes from various intervals (post-intervention, and 3-month, 6-month and 12-month follow-ups) were clustered together in the current review partly due to the limited number of available studies, but also from the varied intervals in which outcomes were reported (e.g., post-intervention only vs. first time-point reported at 3-month or 6-month follow-up without post-intervention outcomes). Given that psychological interventions aim to have an enduring impact on preschool-aged children's well-being and functioning, measuring outcomes at more consistent intervals and ideally over the long term would improve our understanding of potential benefits at different stages of the intervention (i.e., short-, medium- and long-term).

Finally, as further evidence continues to accumulate, future efforts could consider exploring factors that may moderate and mediate the effects of intervention. Exploring intervention characteristics (e.g., type, duration, number of sessions, format of delivery and recipient of intervention), as well as child (e.g., gender, severity of behavioural inhibition, social skills), and environmental factors (e.g., parenting behaviours, parental mental health) would enhance our understanding of factors that moderate the efficacy of intervention. For instance, exploring whether the type of intervention (e.g., parent-education programmes vs. social skills training vs. combination approach) moderates intervention outcomes could enhance our understanding of whether a particular approach is more efficacious than others, which would allow for a more concerted effort in reducing behavioural inhibition and anxiety in preschool-aged children. Additionally, exploring how specific treatment components/processes (e.g., exposure, parent training) are associated with change in preschool-aged children's behavioural inhibition and anxiety could enhance the efficacy of intervention.

4.3. Conclusion

Preliminary evidence from this meta-analysis indicated that intervention targeted at behaviourally inhibited preschool-aged children may be effective in reducing behavioural inhibition and anxiety, but this change was not consistently observed across all outcomes. Further work is needed to gain a more comprehensive understanding on how to best support preschool-aged children identified as at-risk for anxiety.

Declaration of interest

none.

Acknowledgements

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

References

Asendorpf, J. B. (1990). Development of inhibition during childhood: Evidence for situational specificity and a two-factor model. *Developmental Psychology*, 26(5), 721–730. <https://doi.org/10.1037/0012-1649.26.5.721>

Barstead, M. G., Danko, C. M., Chronis-Tuscano, A., O'Brien, K. A., Coplan, R. J., & Rubin, K. H. (2018). Generalization of an Early Intervention for Inhibited Preschoolers to the Classroom Setting. *Journal of Child and Family Studies*, 27(9), 2943–2953. <https://doi.org/10.1007/s10826-018-1142-0>

Bayer, J. K., Beatson, R., Bretherton, L., Hiscock, H., Wake, M., Gilbertson, T., ... Rapee, R. M. (2018). Translational delivery of Cool Little Kids to prevent child

internalising problems: Randomised controlled trial. *The Australian and New Zealand Journal of Psychiatry*, 52(2), 181–191. <https://doi.org/10.1177/0004867417726582>

Bayer, J. K., Prendergast, L. A., Brown, A., Harris, L., Bretherton, L., Hiscock, H., ... Rapee, R. M. (2020). Cool Little Kids translational trial to prevent internalising: Two-year outcomes and prediction of parent engagement. *Child and Adolescent Mental Health*, 26, 211–219. <https://doi.org/10.1111/camh.12420>

Bishop, G., Spence, S. H., & McDonald, C. (2003). Can parents and teachers provide a reliable and valid report of behavioral inhibition? *Child Development*, 74(6), 1899–1917. (<http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med5&NEWS=N&AN=14669903>).

Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale: Lawrence Erlbaum Associates, Inc.

The Lancet Psychiatry. (2020). A good enough measure. *The Lancet Psychiatry*, 7(10), 825. ([https://doi.org/10.1016/S2215-0366\(20\)30395-3](https://doi.org/10.1016/S2215-0366(20)30395-3)).

Chronis-Tuscano, A., Degnan, K. A., Pine, D. S., Perez-Edgar, K., Henderson, H. A., Diaz, Y., ... Fox, N. A. (2009). Stable early maternal report of behavioral inhibition predicts lifetime social anxiety disorder in adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry*, 48(9), 928–935. <https://doi.org/10.1097/CHI.0b013e3181ae09df>

Chronis-Tuscano, A., Rubin, K. H., O'Brien, K. A., Coplan, R. J., Thomas, S. R., Dougherty, L. R., ... Wimsatt, M. (2015). Preliminary evaluation of a multimodal early intervention program for behaviorally inhibited preschoolers. *Journal of Consulting and Clinical Psychology*, 83(3), 534–540. <https://doi.org/10.1037/a0039043>

Coplan, R. J., Schneider, B. H., Matheson, A., & Graham, A. (2010). 'Play skills' for shy children: Development of a social skills facilitated play early intervention program for extremely inhibited preschoolers. *Infant and Child Development*, 19(3), 223–237. <https://doi.org/10.1002/icd.668>

Danko, C. M., O'Brien, K. A., Rubin, K. H., & Chronis-Tuscano, A. (2018). The turtle program: PCIT for young children displaying behavioral inhibition. *Handbook of Parent-Child Interaction Therapy* (pp. 85–98). Springer.

Dodd, H. F., Hudson, J. L., & Rapee, R. M. (2017). Temperament in youth internalizing disorders. *Treatments for Psychological Problems and Syndromes* (pp. 504–524). Wiley Blackwell. (<https://doi.org/10.1002/9781118877142.ch31>).

Edwards, S. L., Rapee, R. M., Kennedy, S. J., & Spence, S. H. (2010). The assessment of anxiety symptoms in preschool-aged children: the revised preschool anxiety scale. *Journal of Clinical Child & Adolescent Psychology*, 39(3), 400–409.

R Core Team, Rf. (2013). *R: A Language and Environment for Statistical Computing*. Austria: R foundation for statistical computing Vienna.

Egger, Ascher, B. H., & Angold, A. (1999). Preschool age psychiatric assessment (PAPA). Durham (North Carolina): Duke University Medical Center.

Fox, N. A., Barker, T. V., White, L. K., Suway, J. G., & Pine, D. S. (2013). Commentary: To intervene or not? Appreciating or treating individual differences in childhood temperament – remarks on Rapee (2013). *Journal of Child Psychology and Psychiatry*, 54(7), 789–790. <https://doi.org/10.1111/jc>

Frenkel, T. L., Fox, N. A., Pine, D. S., Walker, O. L., Degnan, K. A., & Chronis-Tuscano, A. (2015). Early childhood behavioral inhibition, adult psychopathology and the buffering effects of adolescent social networks: A twenty-year prospective study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 56(10), 1065–1073. <https://doi.org/10.1111/jc>

Gazelle, H., & Ladd, G. W. (2003). Anxious solitude and peer exclusion: A diathesis-stress model of internalizing trajectories in childhood. *Child Development*, 74(1), 257–278.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(11), 1337–1345.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *Bmj*, 327(7414), 557–560.

Higgins, J. P. T., Thompson, S. G., & Spiegelhalter, D. J. (2009). A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society: Series A Statistics in Society*, 172(1), 137–159.

Hudson, J. L., Murayama, K., Meteyard, L., Morris, T., & Dodd, H. F. (2019). Early childhood predictors of anxiety in early adolescence. *Journal of Abnormal Child Psychology*, 47(7), 1121–1133. <https://doi.org/10.1007/s10802-018-0495-6>

Kagan, J. (1994). *Galen's Prophecy: Temperament in Human Nature*. New York: Westview Press.

Kagan, J., Reznick, J. S., Clarke, C., Snidman, N., & Garcia-Coll, C. (1984). Behavioral inhibition to the unfamiliar. *Child Development*, 55(6), 2212–2225. <https://doi.org/10.2307/1129793>

Kagan, J., Reznick, J. S., & Gibbons, J. (1989). Inhibited and uninhibited types of children. *Child Development*, 60, 838–845.

Kagan, J., Reznick, J. S., & Snidman, N. (1988). Biological bases of childhood shyness. *Science*, 240(4849), 167–171. (<http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med3&NEWS=N&AN=3353713>).

Kennedy, S. J., Rapee, R. M., & Edwards, S. L. (2009). A selective intervention program for inhibited preschool-aged children of parents with an anxiety disorder: Effects on current anxiety disorders and temperament. *Journal of the American Academy of Child and Adolescent Psychiatry*, 48(6), 602–609. <https://doi.org/10.1097/CHI.0b013e31819f6fa9>

Kmet, L. M., Lee, R. C., Cook, L. S. (2004). Standard quality assessment criteria for evaluating primary research papers from a variety of fields. Alberta Heritage Foundation for Medical Research. Retrieved from (<http://www.ihc.ca/Document s/HTA-FR13.Pdf>).

Ladd, G. W. (2010). *The Child Behavior Scale: Applications and Research Findings*. Arizona: Parkview Publications.

- Ladd, G. W., & Profilet, S. M. (1996). The child behavior scale: A teacher-report measure of young children's aggressive, withdrawn, and prosocial behaviors. *Developmental Psychology*, 32(6), 1008–1024.
- LaFreniere, P. J., Dumas, J. E. (1995). Social competence and behavior evaluation scale: Preschool edition (SCBE). Western Psychological Services.
- LaFreniere, P. J., & Capuano, F. (1997). Preventive intervention as means of clarifying direction of effects in socialization: Anxious-withdrawn preschoolers case. *Development and Psychopathology*, 9(3), 551–564. (<http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med4&NEWS=N&AN=9327239>).
- Lau, E. X., Rapee, R. M., & Coplan, R. J. (2017). Combining child social skills training with a parent early intervention program for inhibited preschool children. *Journal of Anxiety Disorders*, 51, 32–38. <https://doi.org/10.1016/j.janxdis.2017.08.007>
- Lawrence, P. J., Rooke, S. M., & Creswell, C. (2017). Prevention of anxiety among at-risk children and adolescents—a systematic review and meta-analysis. *Child and Adolescent Mental Health*, 22(3), 118–130.
- Liu, P., & Pérez-Edgar, K. E. (2019). Developmental pathways from early behavioral inhibition to later anxiety: an integrative review of developmental psychopathology research and translational implications. *Adolescent Research Review*, 4(1), 45–58.
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the depression anxiety stress scales (DASS) with the beck depression and anxiety inventories. *Behaviour Research and Therapy*, 33(3), 335–343.
- Luke, F., Chan, C. C., Au, A., & Lai, S. M. K. (2017). Adaptive parenting for alleviating young children's shyness: A randomized controlled trial of an early intervention program (n/a-N.PAG) *Infant & Child Development*, 26(6). <https://doi.org/10.1002/icd.2030>.
- Moncrieff, J., Churchill, R., Colin Drummond, D., & McGuire, H. (2001). Development of a quality assessment instrument for trials of treatments for depression and neurosis. *International Journal of Methods in Psychiatric Research*, 10(3), 126–133. <https://doi.org/10.1002/mpr.108>
- Morgan, A. J., Rapee, R. M., Salim, A., Goharpey, N., Tamir, E., McLellan, L. F., & Bayer, J. K. (2017). Internet-delivered parenting program for prevention and early intervention of anxiety problems in young children: randomized controlled trial. *e1 Journal of the American Academy of Child and Adolescent Psychiatry*, 56(5), 417–425. <https://doi.org/10.1016/j.jaac.2017.02.010>.
- Morgan, A. J., Tamir, E., Rapee, R. M., Lyneham, H. J., McLellan, L. F., & Bayer, J. K. (2019). Online assessment of preschool anxiety: Description and initial validation of a new diagnostic tool. *Child & Adolescent Mental Health*, 24(3), 259–265. <https://doi.org/10.1111/camh.12324>
- Morris, S. B. (2007). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, 11(2), 364–386. <https://doi.org/10.1177/1094428106291059>
- Pérez-Edgar, K., & Fox, N. A. (2005). Temperament and anxiety disorders. *Child Adolescent Psychiatric Clinics*, 14(4), 681–706.
- Presley, R., & Martin, R. P. (1994). Toward a structure of preschool temperament: Factor structure of the temperament assessment battery for children. *Journal of Personality*, 62(3), 415–448.
- Prior, M., Smart, D., Sanson, A. N. N., & Oberklaid, F. (2000). Does shy-inhibited temperament in childhood lead to anxiety problems in adolescence? *Journal of the American Academy of Child & Adolescent Psychiatry*, 39(4), 461–468.
- Rapee, R. M., Kennedy, S. J., Lau, E. X. (2010). Cool Little Kids: Anxiety Prevention Program. Centre for Emotional Health, Macquarie University.
- Rapee, R. M. (2013). The preventative effects of a brief, early intervention for preschool-aged children at risk for internalising: follow-up into middle adolescence. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 54(7), 780–788. <https://doi.org/10.1111/jc>
- Rapee, R. M., & Bayer, J. K. (2018). Behavioural inhibition and the prevention of internalising distress in early childhood. *Behavioral Inhibition* (pp. 337–355). Springer.
- Rapee, R. M., & Coplan, R. J. (2010). Conceptual relations between anxiety disorder and fearful temperament. *New Directions for Child and Adolescent Development*, 2010(127), 17–31.
- Rapee, R. M., Kennedy, S., Ingram, M., Edwards, S., & Sweeney, L. (2005). Prevention and early intervention of anxiety disorders in inhibited preschool children. *Journal of Consulting and Clinical Psychology*, 73(3), 488–497. (<http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med6&NEWS=N&AN=15982146>).
- Rapee, R. M., Kennedy, S. J., Ingram, M., Edwards, S. L., & Sweeney, L. (2010). Altering the trajectory of anxiety in at-risk young children. *The American Journal of Psychiatry*, 167(12), 1518–1525. <https://doi.org/10.1176/appi.ajp.2010.09111619>
- Rapee, R. M., Schniering, C. A., & Hudson, J. L. (2009). Anxiety disorders during childhood and adolescence: Origins and treatment. *Annual Review of Clinical Psychology*, 5, 311–341.
- Rubin, K. H. (2008). The Play Observation Scale (Revised 2008). University of Maryland.
- Rubin, K. H. (2001). *The Play Observation Scale. The Center for Children, Relationships and Culture*. College Park: University of Maryland.
- Rubin, K. H., Coplan, R. J., & Bowker, J. C. (2009). Social withdrawal in childhood. *Annual Review of Psychology*, 60, 141–171.
- Rubio-Aparicio, M., Marin-Martinez, F., Sanchez-Meca, J., & Lopez-Lopez, J. A. (2018). A methodological review of meta-analyses of the effectiveness of clinical psychology treatments. *Behavior Research Methods*, 50(5), 2057–2073.
- Sandstrom, A., Uher, R., & Pavlova, B. (2020). Prospective association between childhood behavioral inhibition and anxiety: a meta-analysis. *Journal of Abnormal Child Psychology*, 48(1), 57–66. (<https://doi.org/10.1007/s10802-019-00588-5>).
- Sanson, A. (1996). Shyness ratings: stability and correlates in early childhood. *International Journal of Behavioral Development*, 19(4), 705–724.
- Schwartz, C. E., Snidman, N., & Kagan, J. (1999). Adolescent social anxiety as an outcome of inhibited temperament in childhood. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38(8), 1008–1015. (<http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med4&NEWS=N&AN=10434493>).
- Silverman, W. K., & Albano, A. M. (1996). *Anxiety Disorders Interview Schedule for DSM-IV: Child and Parent Versions*. Oxford University Press.
- Spence, S. H., Rapee, R., McDonald, C., & Ingram, M. (2001). The structure of anxiety symptoms among preschoolers. *Behaviour Research and Therapy*, 39(11), 1293–1316.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–48.
- Xu, Y., Farver, J. A. M., Chang, L., Zhang, Z., & Yu, L. (2007). Moving away or fitting in? Understanding shyness in Chinese children. *Merrill-Palmer Quarterly*, 53(4), 527–556. <https://doi.org/10.1353/mpq.2008.0005>
- Xu, Y., Farver, J. A. M., Yu, L., & Zhang, Z. (2009). Three types of shyness in Chinese children and the relation to effortful control. *Journal of Personality and Social Psychology*, 97(6), 1061–1073. <https://doi.org/10.1037/a0016576>