Case Study 1: An Evidence-Based Practice Review Report

Theme: School (setting) based interventions for children with special educational needs (SEN)

The effectiveness of 'Cool Kids', a manualised CBT intervention in the reduction of diagnosed anxiety in school aged pupils with a recognised anxiety disorder

Summary

Anxiety is the most prevalent psychiatric problem in childhood (Mazzone et al., 2007) and this problem has been exacerbated by the COVID-19 global pandemic with up to one in five young people experiencing clinically increased symptoms of anxiety (Racine et al., 2021). Research suggests that Cognitive Behavioural Therapy (CBT) is an efficacious treatment in minimising or, at times, eliminating the effects of anxiety on a child or young person (James et al., 2015). This review examines the effectiveness of a manualised CBT intervention developed in Australia known as 'Cool Kids' (Johnsen et al., 2019) and considers if the intervention would be appropriate for use in a UK school-based context. The review found that in Australia and Denmark the intervention produced promising results in reducing anxiety levels in children and young people. Recent evidence (Scaini et al., 2022) suggests that the programme is effective in a school-based setting, however, further research in a UK school-based context is necessary before such claims can be made in the UK.

Introduction

Anxiety in children and young people

Anxiety, as defined by Foa et al. (2017), can be seen as challenges that prevent or limit normal adaptive behaviour that is developmentally expected. Anxiety is one of the most common types of psychiatric problem in childhood (Mazzone et al., 2007) and is considered to affect an estimated 117 million youths worldwide (Polanczyk et al., 2015). It is important to address the issues of anxiety in children early on as, left untreated, anxiety can lead to further challenges in adulthood such as poorer physical and mental health and poorer life opportunities (Hawton et al., 2012; Swift et al., 2014). It can impact on social interaction and the ability to make friends as well as academic success in school (Creswell & Willetts, 2019). Since the outbreak of the COVID-19 pandemic in early 2020, young people's wellbeing has been impacted greatly, with many reporting higher levels of depression and anxiety (Duan et al., 2020; Ravens-Sieberer et al., 2021; Walters et al., 2021). Recommendations suggest that to minimise the impact of mental health problems increased by the pandemic, early intervention and preventative support should be offered (Loades et al., 2020). This has implications for the work of Educational Psychologists (EPs) as they work with staff and children, on the frontline, who are experiencing the day to day impact of these issues.

Efficacy of Cognitive Behavioural Therapy (CBT) Interventions

A meta-analysis of 55 studies by Reynolds et al. (2012) suggests that CBT is an efficacious intervention for children and young people struggling with anxiety, as does the most recent Cochrane Review (James et al., 2015).

CBT finds its theoretical basis in behaviour therapy and cognitive therapy (Beck, 1989). It focuses on the links between thoughts, emotions and behaviours (Fenn & Byrne, 2013). The theory states that it is the perceptions of the individual around a situation that evokes a positive or negative feeling rather than the situation itself. Therefore, learning to address the emotions around a variety of situations can equip individuals to better cope for the future. The 'Cool Kids Anxiety Program' (Johnsen et al., 2019) was developed to include psychoeducation around anxiety (developing an understanding of information related to anxiety and why it can occur), cognitive restructuring considering the likelihood of feared events, parent skills (enabling parents to better support their child), in-vivo exposure helping the child to gradually face their feared events, social skills and improved coping strategies. The overall aim of the programme is to teach children and young people to recognise negative emotion, restructure negative automatic thoughts (NATs) and gradually confront feared situations (Arendt et al., 2016b). The programme is typically delivered over a 12-week period with varied session length still showing a reduction in anxiety symptoms. The current review suggests that the programme can be delivered by professional therapists, student therapists or in collaboration with a trained facilitator. Mychailyszyn (2017) reviewed 16 studies to assess the effectiveness of the Cool Kids programme in reducing symptomatic anxiety and found strong empirical support that it does reduce anxiety. The current review focuses solely on the reports of children and young people to assess the effectiveness of the Cool Kids programme whereas the review by Mychailyszyn (2017) analysed both child and parent reports.

In the UK, EPs are important in supporting mental health provision within school based settings (Zafeiriou & Gulliford, 2020). As anxiety is so prevalent and seems to increase as pupils age (ImpactED, 2021), it is important that EPs are equipped to offer evidence based support early on to prevent continued issues.

This programme incorporates the psychological underpinnings of CBT with practical activities to effectively reduce anxiety symptoms in children and young people, in a cost-effective manner. It will be useful for EPs to be aware of its potential uses and to know that what they recommend is grounded in a firm evidence base. Therefore, a systematic literature review around popular CBT interventions, such as the 'Cool Kids Anxiety Program', will be of benefit to current EP practice. In light of the current UK financial situation, EPs must provide cost and time effective interventions as schools are often struggling for budget amidst a rise in mental health challenges.

Review Question

'The effectiveness of 'Cool Kids', a manualised CBT intervention in the reduction of diagnosed anxiety in school aged pupils with a recognised anxiety disorder.'

This review considers the Cool Kids Anxiety Programme and its effectiveness in reducing reported anxiety symptoms from the perspective of the child or young person. The aim of this review is to critically review the research base for the programme and to consider the implications for practising Educational Psychologists and future research.

Critical Review of the Evidence Base

Literature Search

A systematic search of literature was concluded on 1st February 2022 to answer the review question. The following databases were used to conduct a search using the key terms shown in Table 1: PsycINFO, Education Resources Information Centre (ERIC), Web of Science and Scopus. The key search term used was 'Cool Kids'. This term was used as the review question looks specifically at the 'Cool Kids' manualised CBT programme. Computerised database searching was utilised first due to its precision (Conn et al., 2003). Alternative search items were explored, including specific participant groups, however, they did not yield any useful results. Due to this, an ancestral search was conducted to ensure all relevant literature was included, however, no relevant research was found. The 38 studies identified through searching were screened firstly by their title, the date and then by their abstract. A full-text review was used to ensure the final studies found were appropriate. Table 2 details the inclusion and exclusion criteria used for this review. From the six studies which underwent a full text review and were assessed against the criteria, one did not meet the criteria as shown in Appendix A. The five studies that met the parameters of the criteria were selected for analysis as part of this review. References for the studies selected can be found in Table 3. A PRISMA diagram illustrates the systematic search and screening process undertaken in Figure 1.

Table 1

Database and Search Terms

Database	Search Terms	*n*
PsycINFO	Cool Kids	45
ERIC	“cool kids”	48
Web of Science	“Cool Kids”	9
Scopus	“cool kids”	49

Figure 1: PRISMA diagram depicting the database search and screening process

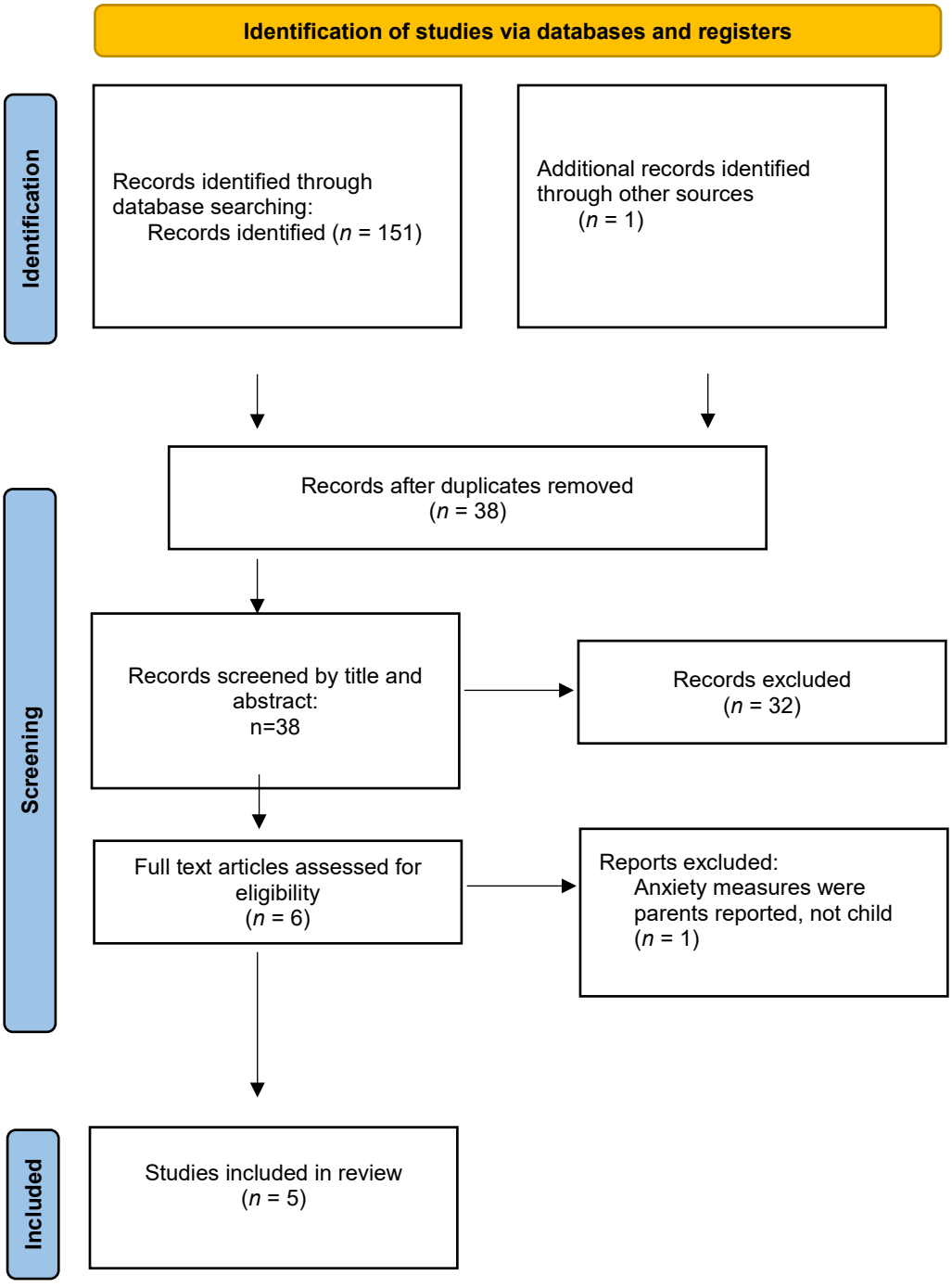

Table 2

Inclusion and Exclusion Criteria for current review

	Study Feature	Inclusion Criteria	Exclusion Criteria	Rationale
1	Diagnosis	Children/ young people met the criteria for a diagnosed anxiety disorder using an internationally recognised assessment scale	Children/ young people did not meet the criteria for a diagnosed anxiety disorder using an internationally recognised assessment scale	The review question looks at the effectiveness of reducing diagnosed anxiety in children and young people
2	Types of studies	The studies are either a randomised control trial, cohort study, Quasi-experimental study, or Open trial design study	Studies that are not a randomised control trial, cohort study, Quasi-experimental study, or Open trial design study	Randomised control trials, cohort studies, Quasi-experimental studies and Open trial design studies have been suggested to be the 'best' study type for measuring effectiveness (Petticrew & Roberts, 2003)
3	Language	Study published in the English language	Studies published in a language other than English and not translated into English	Only studies written in English could be used as there were no translation services available

4	Participants	School aged pupils between 7 – 18 years	Participants younger than 7 years and older than 18 years	The review looks at an intervention aimed at school aged pupils
5	Intervention	The study uses the 'Cool Kids' Anxiety Program (Johnsen et al., 2019)	Studies that do not use the 'Cool Kids' Anxiety Program	This review looks specifically at the 'Cool Kids' Anxiety Program
6	Report Measuring	The study reports the child or young person's anxiety rating	Anxiety measures were reported by parents	This review is considering the use of the 'Cool Kids' programme and focuses on child reported anxiety
7	Types of Publications	Peer-reviewed journal articles	Non-peer reviewed journal articles	To ensure that the articles used are of a high quality and credible (Kelly et al., 2014)
8	Country of study	Study conducted in an Organisation for Economic Co-operation and Development (OECD) country	Studies conducted in non-OECD countries	OECD (2021) countries share contextual similarities with the United Kingdom which may therefore allow for more generalisability to a UK context
9	Date (post 2015)	Journal articles published between 2016-2021	Studies published pre-2015	A systematic literature review was published in 2017 (Mychailyszyn, 2017) in which

				they included studies published up to 11/06/2015. Therefore, this review looked at studies published from 2016 onwards
10	Outcome	The study must evaluate pre and post outcomes of anxiety symptoms	The study does not evaluate pre and post outcomes of anxiety symptoms	This review is looking at the reduction of anxiety in children

Table 3

The final 5 studies included in the systematic literature review

1	Johnsen, D. B., Arendt, K., & Thastum, M. (2019). The efficacy of manualized Cognitive Behavior Therapy conducted by student-therapists treating Danish youths with anxiety using a benchmark comparison. *Scandinavian journal of child and adolescent psychiatry and psychology*, *7*, 68–80. https://doi.org/10.21307/sjcapp-2019-010
2	Sciberras, E., Mulraney, M., Anderson, V., Rapee, R. M., Nicholson, J. M., Efron, D., Lee, K., Markopoulos, Z., & Hiscock, H. (2018). Managing Anxiety in Children With ADHD Using Cognitive-Behavioral Therapy: A Pilot Randomized Controlled Trial. *Journal of Attention Disorders*, *22*(5), 515–520. https://doi.org/10.1177/1087054715584054
3	Arendt, K., Thastum, M., & Hougaard, E. (2016a). Homework Adherence and Cognitive Behaviour Treatment Outcome for Children and Adolescents with Anxiety Disorders. *Behavioural and Cognitive Psychotherapy*, *44*(2), 225–235. https://doi.org/10.1017/S1352465815000429
4	Arendt, K., Thastum, M., & Hougaard, E. (2016b). Efficacy of a Danish version of the Cool Kids program: A randomized wait-list controlled trial. *Acta Psychiatrica Scandinavica*, *133*(2), 109–121. https://doi.org/10.1111/acps.12448
5	McLellan, L. F., Andrijic, V., Davies, S., Lyneham, H. J., & Rapee, R. M. (2017). Delivery of a Therapist-Facilitated Telecare Anxiety Program to Children in Rural Communities: A Pilot Study. *Behaviour change*, *34*(3), 156–167. https://doi.org/10.1017/bec.2017.11

Weight of Evidence (WoE)

The Weight of Evidence Framework (Gough, 2007) was used to appraise the quality of the studies selected for this review. This framework is used to evaluate the relevance of each piece of evidence in relation to how well it contributes to answering the review question. The framework considers the methodological quality of the research, the relevance of the methodology and how well the research contributes to the review question. The framework divides these considerations into three areas called Weight of Evidence A, B and C.

Weight of Evidence A (WoE A) is a non-review specific judgement considering the quality of the studies in comparison to others of its type (Gough, 2007). This review used a modified version of the Critical Review Form for Quantitative Studies developed by Law et al (1998a). Omissions and rationale for these can be found in Appendix C.

Weight of Evidence B (WoE B) is a review-specific judgement in which consideration of the relevance of the methodology is used for the research identified (Gough, 2007). The criteria selected, based on the Typology of Evidence criteria by Petticrew and Roberts (2003), can be found in Appendix D.

Weight of Evidence C (WoE C) is a review-specific judgement that considers how relevant the research is in relation to the review question (Gough, 2007). The criteria used for WoE C can be found in Appendix E.

The considerations of WoE A, B and C are averaged to culminate in one final weight of evidence score called Weight of Evidence D (WoE D). This is an overall assessment of how well the evidence presented in the selected

studies helps in answering the review question (Gough, 2007). Table 4 details the scores for the five studies used in this review. Table 1 (Appendix C) details the criteria used to define the overall quality of the research. A rating of ≤1.0 was considered 'low', 1.1-1.5 was considered 'medium' and ≥ 1.6 was considered 'high'.

Table 4

WoE D - Overall Weight of Evidence

Study	WoE A	WoE B	WoE C	WoE D
Johnsen et al. (2019)	High (2)	Low (1)	Low (1)	Medium (1.33)
Sciberras et al. (2018)	Medium (1.14)	High (2)	Low (0.75)	Medium (1.30)
Arendt et al. (2016b)	Medium (1.29)	Low (1)	Low (1)	Medium (1.10)
Arendt et al. (2016a)	High (1.86)	Low (1)	Low (1)	Medium (1.29)
McLellan et al. (2017)	High (1.58)	Low (1)	High (2)	Medium (1.53)

Note: WoE D ratings are described as 'Low' for scores ≤1.0, 'Medium' for scores 1.1 - 1.5, and 'High' for scores ≥ 1.6.

Study Participants

In total, the five studies reviewed included 288 participants between the ages of 7-16 years old and were from either Australia (McLellan et al., 2017; Sciberras et al., 2018) or Denmark (Arendt et al., 2016a; Arendt et al., 2016b; (Johnsen et al., 2019). Studies conducted in Australia received a higher WoE C rating due to the similarities with the UK in their education system and

transition stages. Although studies from Denmark received a lower WoE C rating for criterion D 'Location', they were still considered relevant as this country forms part of the Organisation for Economic Co-Operation and Development. Participants in these studies were recruited via seeking help from a university department (Arendt et al., 2016a; Johnsen et al., 2019), an ADHD assessment clinic (Sciberras et al., 2018) and through adverts in rural schools (McLellan et al., 2017). One study did not detail how participants were recruited (Arendt et al., 2016b). Each study reported the percentage of males and females except Sciberras et al. (2018) who only recruited male participants. Each study varied in detail with regards to other participant demographics which resulted in different sample rating for WoE A. For example, Arendt et al. (2016a) gave very limited detail regarding participant characteristics which contributed to the low WoE A score for the 'Sample' criterion.

Study Design

Due to their similarities, this review question applied the quasi-experimental criteria to the cohort (open trial) design. There are potential issues with drawing causal inferences with this design as doing so would follow the post hoc ergo propter hoc fallacy (Barker et al., 2015). However, if sufficient contextual information is provided within the study, including post-treatment ratings, the results are still worthy of consideration (Cook, 1979). Each study followed the 'Cool Kids' manualised programme and reports the amount and length of sessions, including booster sessions. Drop out and attrition rates, which are important in understanding the comparability of pre and post test

results, were reported in each study (Gersten et al., 2005). One study included used a randomised controlled trial design (RCT) (Sciberras et al., 2018) which resulted in a higher WoE B rating for study design as RCTs are seen to be the 'Gold Standard' when considering the effectiveness of interventions (Petticrew & Roberts, 2003). One study used a randomised wait-list controlled trial design but did not compare the wait-list group with the intervention group, deciding to pool the data into overall results instead (Arendt et al., 2016a). Any missing data points were explained in the studies.

Interventions

The Cool Kids Anxiety Program, developed at Macquarie University, Sydney, in 1993, is grounded in the principles of Cognitive Behavioural Therapy (CBT) to teach children and their parents' practical skills for coping with anxiety. Typically, it is delivered as 10, two-hour sessions over 12 weeks with one in vivo session recommended to take place in a public setting (e.g. a shopping mall). One study followed this format with no adaptations to the programme (Johnsen et al., 2019) and two studies followed the format with the only adaptation being translating the resources (Arendt et al., 2013a; Arendt et al., 2013b). Sciberras et al. (2018) adapted the programme by using: shorter sessions, activity schedules and positive reinforcement for on task behaviour, giving 'brain breaks', shortening and repeating key concepts and using visual aids in sessions and at home. The most relevant adaptations for this review question were made by McLellan et al. (2017) who adapted the programme for use in a school via video software so that the child had access to a professional therapist. A higher WoE C was given for

criterion A 'Intervention and Setting'. This is because this review aimed to consider the applicability of the Cool Kids intervention in a school-based setting, therefore, if adaptations to the programme still produced a significant reduction in anxiety ratings, it would suggest the programme is adaptable whilst still being generalisable. Originally, the programme was developed for a professional therapist, however, some of these studies used only student therapists (Johnsen et al., 2019), a mix of either trained psychologists and graduate students (Ardent et al., 2016a; Ardent et al., 2016b) or a combination of a professional therapist and an identified school adult (McLellan et al., 2017). This is reflected in the WoE C criterion C 'Instructor' rating.

Measures

For this review, there were three main measures focussed upon. The Anxiety Disorders Interview Schedule for Children IV (ADIS-C-IV) (Silverman et al., 2001) was used in each study to ensure participants had a recognised anxiety diagnosis. This measure has been shown to have excellent test-retest reliability for children as well as concurrent validity when compared to self-report anxiety measures (Lyneham & Rapee, 2005).

Two self-report measures were used to measure the severity of anxiety symptoms and symptom impairment. As this review considers the reduction of anxiety in children (C) and youths (Y), these are the measures that are focussed on. Although reported in the reviewed studies, parent measures were not considered. The Spence Children's Anxiety Scale (SCAS-C/Y) (Spence, 1997) is a 38-item questionnaire providing an anxiety score. This

measure was normed on a mostly English speaking, Caucasian population based in Australia with a sample of 4,916 children between the ages of 8-15 years. This measure was found to have convergent validity, high internal consistency and a modest level of test-retest reliability (Spence et al., 2003). The Child Anxiety Life Interference Scale (CALIS-C/Y) (Lyneham et al., 2013) is a nine-item scale measuring the functional impact of anxiety symptoms including social, academic and home life (McLellan et al., 2017). The CALIS has demonstrated internal consistency and moderate to high test-retest reliability, significant interrater reliability and good convergent and divergent validity (Johnsen et al., 2019; Lyneham et al., 2013). Arendt et al. (2016a) only reported outcomes using the SCAS-C. All studies apart from two (McLellan et al., 2017; Sciberras et al., 2018) reported the reliability and validity of the measures used. This is reflected in the WoE dimension 5 'Outcomes' rating. Due to the reliability and validity of the measures, the studies were rated two for WoE A.

Outcomes and Effect sizes

Effect sizes for the SCAS and CALIS from the studies reviewed can be found in Table 5. Statistical significance was reported in three of the five studies (Arendt et al., 2016b; Johnsen et al., 2019; McLellan et al., 2017. See Appendix B). Three studies (Arendt et al., 2016a; Arendt et al., 2016b; Johnsen et al., 2019) used pre-test, post-test and follow up measures which aids in measuring the effectiveness of the intervention. One study compared measures against a control group who received treatment as usual (Sciberras et al., 2018) and one study only reported pre-treatment and post-

treatment data (McLellan et al., 2017). This is reflected in the WoE A dimension 5 'Outcomes' rating. Johnsen et al. (2019) reported Hedges' *g* for effect size which is comparable to Cohen's *d* as they both use pooled standard deviation as the standardiser (Lakens, 2013) and use the same parameters for descriptors. The author of this review noted that the effect sizes for some studies were not reported. Therefore, these were manually calculated using an effect size calculator (Lenhard, 2016). Effect sizes for two studies were calculated by the reviewer using the pooled standard deviations (Arendt et al., 2016b; McLellan et al., 2017). The effect sizes for all five studies reviewed can be found in Table 5.

Each study reported a large effect size for the SCAS and medium effect size for the CALIS except Scibberas et al. (2018) who reported a small effect size for each. This variance may be because they only reported the effect size for the five-month follow up and the effect may be expected to diminish overtime. The effect sizes reported by McLellan et al. (2017) were very high (SCAS *d* = 5.96 and CALIS *d*= 2.93) and were subsequently recalculated (SCAS *d* = 1.45 and CALIS *d*= 0.65). These results still showed a large and medium effect size, respectively. However, these results should be interpreted with caution due to the study's small sample size which limits generalisability to the population. Importantly, the results of each of the reviewed studies showed a decrease in reported anxiety. One study did not report any follow-up measures (McLellan et al., 2017). This is reflected in the WoE A Dimension 5 'Outcomes' rating where a higher score was given for studies with follow-up measures as these enable researchers to have a greater understanding of the long-term effects of the intervention (Gersten et al.,

2005) and may help to reduce the issue of confounding variables impacting the intervention (Barker et al., 2015).

Table 5

Effect Sizes and Descriptors for Statistically Significant Findings

	Study	*n*	Measures	Age of respondent	Effect Size	Descriptor (*g* or d***)	WoE D
1	Johnsen et al. (2019)	110	CALIS-Y & SCAS-Y	7- to 16-year-olds	SCAS-Y *g* = 1.08* [b c] CALIS-Y *g* = .99* [b c]	Large Large	Medium (1.33)
2	Sciberras et al. (2018)	12	CALIS-C & SCAS-C	8- to 12-year-olds	SCAS-C *d* = 0.2** [a] CALIS-C *d* = 0.4*** [a]	Small Small	Medium (1.30)
3	Arendt et al. (2016b)	98	SCAS-C	7- to 16-year-olds	SCAS-C *d* = 0.86**** [b]	Large	Medium (1.10)
4	Arendt et al. (2016a)	109	CALIS-Y & SCAS-Y	7- to 16-year-olds	SCAS-Y *d* = 1.08 [a] CALIS-Y *d* = 0.63 [a]	Large Medium	Medium (1.29)
5	McLellan et al. (2017)	16	CALIS-C & SCAS-C	9 – 12-year-olds	SCAS-C *d* = 1.45 [b] CALIS-C *d* = 0.65**** [b]	Large Medium	Medium (1.53)

*Note: *g* is for Hedges' g which is a measure of effect size comparable to Cohen's *d.* The parameters for both are 'Small' is 0.2 'Medium' is 0.5 and 'Large' is 0.8 (Cohen, 1988). ** *d* is an effect size referred to as Cohen's d (see parameters in *g* for descriptors). *** Effect sizes for Sciberras et al. (2018) were reported for five - month follow up only. **** This effect size was manually calculated by the author of this review. [a] Effect size refers to interaction effects (between group, pre-post). [b] Effect size refers to within group effects (pre-post change measured). [c] This effect size refers to student therapist led intervention as this review is considering the impact of minimal training for the interventionist to consider if the intervention can be carried about by teaching staff in schools.

Conclusion

This review explored five studies which used the CBT based 'Cool Kids Anxiety Program'. Results of the studies were drawn from child and youth reported measures considering how their anxiety had been impacted after receiving the intervention. All of the studies scored medium for WoE D which indicates that they are methodologically sound and relevant for this review. Results suggest the intervention led to a significant reduction in reported anxiety levels. Arendt et al. (2016a) found not only reduced anxiety levels but for some participants such reduction that they no longer met the criteria for their primary anxiety diagnosis. They also reported that some participants had either reduced or stopped their use of psychopharmacological medication. The results of these studies add to the previous evidence that the 'Cool Kids' intervention is efficacious in treating anxiety problems in children and young people. Results from Arendt et al. (2016b) suggest that at post treatment and follow up, anxiety levels were significantly lower than at the pre-intervention measures. This research also found that completion (or lack thereof) of homework tasks linked to the programme did not impact the reduction of anxiety in the participants. One study (McLellan et al., 2017) found that, despite minimal parental involvement, reduced anxiety outcomes were still achieved at a statistically significant level. Results from Johnsen et al. (2019) specifically considered the impact of the interventionist and found that student therapists with minimal training in the intervention were still able to produce statistically significant results in the reduction of anxiety. The strong WoE A scores for methodology give strength to the conclusions drawn from this review. However, the lower WoE C scores for context suggest that

more research is needed within a UK school setting with the interventionist being a member of school staff, rather than a therapist. All of these conclusions are important for this particular review as the reviewer wanted to consider the application of the 'Cool Kids' intervention in a UK school-based setting. The studies reviewed conclude that: adaptations to the programme, minimal parental involvement, lack of homework adherence and minimal interventionist training do not limit the interventions impact in reducing anxiety for children and young people. Therefore, it may be possible that the intervention could be adapted to work in a UK school-based setting. However, these findings do need to be considered in the light of various limitations of the studies reviewed which will now be discussed.

Limitations

The studies reviewed were conducted in either Australia (McLellan et al., 2017; Sciberras et al., 2018) or Denmark (Arendt et al., 2016a; Arendt et al., 2016b; Johnsen el al., 2019). Although these are OECD countries, there are limitations to the generalisability of the findings to the UK context. Therefore, further research, in the UK, is necessary to allow for generalisability of the results to UK school-based settings. All the studies took place in a more clinical setting except for McLellan et al (2017) which took place in a school-based setting via a therapist-facilitated telecare adaptation to the Cool Kids intervention. This research suggests that adaptations to the intervention can still lead to successful reductions in anxiety when run in a school setting. However, further research is needed in UK school-based settings to ensure that the intervention both works in the UK and can be delivered successfully

via minimally trained staff in a school setting (see Johnsen et al., 2019 for more of an understanding on the impact of the interventionist).

Recommendations for further research

One of the defining features of EPs in the UK is their ability to support the incorporation evidence-based practice into the day to day practicalities of the classroom. This, alongside the observation of the impact of evidence informed practice, is much appreciated by teachers (Coldwell et al., 2017). Knowing that an intervention is efficacious and grounded in research can save time and money in looking for 'what works'; two resources which are often in short supply in schools. The results of the research reviewed in this paper suggest that the 'Cool Kids' programme is effective in reducing anxiety in school-aged pupils. However, these results need to be replicated in a UK school-based setting before claims of its effectiveness in the UK can be made. The adaptations reviewed in the five papers suggest there is scope for changing the programme so that it may be delivered successfully in a UK school with minimal interventionist training. Once the question of efficacy has been answered in a UK setting, further research should consider using a randomised-control trial design and sufficient sample size to allow for a greater level of generalisability of the results.

References

Arendt, K., Thastum, M., & Hougaard, E. (2016a). Efficacy of a Danish version of the Cool Kids program: A randomized wait-list controlled trial. *Acta Psychiatrica Scandinavica*, *133*(2), 109–121. https://doi.org/10.1111/acps.12448

Arendt, K., Thastum, M., & Hougaard, E. (2016b). Homework Adherence and Cognitive Behaviour Treatment Outcome for Children and Adolescents with Anxiety Disorders. *Behavioural and Cognitive Psychotherapy*, *44*(2), 225–235. https://doi.org/10.1017/S1352465815000429

Barker, C., Pistrang, N., & Elliott, R. (2015). Research Methods in Clinical Psychology. *Research Methods in Clinical Psychology*. https://doi.org/10.1002/9781119154082

Beck, A. T. (1989). *Cognitive therapy and the emotional disorder*. Penguin.

Coldwell, M., Greany, T., Higgins, S., Brown, C., Maxwell, B., Stielle, B., Stoll, L., Willis, B., & Burns, H. (2017). *Evidence-informed teaching: an evaluation of progress in England* (pp. 1–75). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/625007/Evidence-informed_teaching_-_an_evaluation_of_progress_in_England.pdfConn, V. S., Isaramalai, S.-A., Rath, S., Jantarakupt, P., Wadhawan, R., & Dash, Y. (2003). Clinical Scholarship Beyond MEDLINE for Literature Searches Rationale for Comprehensive Literature Searching. *Journal of Nursing Scholarship Second Quarter*, *35*(2), 177–182. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.617.3885&re

p=rep1&type=pdf#:~:text=Ancestry searching refers to the,to other potentially useful studies.

Cook, T. D. (1979). *Quasi-experimentation : design & analysis issues for field settings / Thomas D. Cook, Donald T. Campbell.* (Donald T. Campbell (ed.). Houghton Mifflin.

Creswell, C., & Willetts, L. (2019). *Helping Your Child With Fears and Worries* (2nd ed.). Robinson.

Duan, L., Shao, X., Wang, Y., Huang, Y., Miao, J., Yang, X., & Zhu, G. (2020). An investigation of mental health status of children and adolescents in china during the outbreak of COVID-19. *Journal of Affective Disorders*, *275*(June), 112–118. https://doi.org/10.1016/j.jad.2020.06.029

Fenn, K., & Byrne, M. (2013). The key principles of cognitive behavioural therapy. *InnovAiT: Education and Inspiration for General Practice*, *6*(9), 579–585. https://doi.org/10.1177/1755738012471029

Foa, E. B., Franklin, M., McLean, C., McNally, R. J., & Pine, D. (2017). Defining anxiety disorders. In D. L. Evans, E. B. Foa, R. E. Gur, H. Hendin, C. P. O'Brien, D. Romer, M. E. P. Seligman, & B. T. Walsh (Eds.), *Treating and preventing adolescent mental health disorders: What we know and what we don't know: A research agenda for improving the mental health of our youth* (pp. 189–211). Oxford University Press. https://doi.org/10.1093/med-psych/9780199928163.003.0009

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., &

Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, *71*(2), 149–164. https://doi.org/10.1177/001440290507100202

Gough, D. (2007). Weight of evidence: A framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, *22*(2), 213–228. https://doi.org/10.1080/02671520701296189

Hawton, K., Saunders, K. E. A., & O'Connor, R. C. (2012). Self-harm and suicide in adolescents. *The Lancet*, *379*(9834), 2373–2382. https://doi.org/10.1016/S0140-6736(12)60322-5

ImpactED. (2021). *Impact in Practice: Lessons from schools in supporting pupil learning and wellbeing*. https://drive.google.com/file/d/12vXwHb0eJwEwDnutTkxBm_2EnLAC-T8N/view

James, A. C., Reardon, T., Soler, A., James, G., & Creswell, C. (2015). Cognitive behavioural therapy for anxiety disorders in children and adolescents. *Cochrane Database of Systematic Reviews*, *2*. https://doi.org/https://doi.org/10.1002/14651858.CD004690.pub4

Johnsen, D B, Arendt, K., & Thastum, M. (2019). The efficacy of manualized Cognitive Behavior Therapy conducted by student-therapists treating Danish youths with anxiety using a benchmark comparison. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology*, *7*, 68–80. https://doi.org/10.21307/sjcapp-2019-010

Kelly, J., Sadeghieh, T., Adeli, K., & Biochemistry, C. (2014). *Peer review in scientific publications: benefits, critiques, & a survival guide*.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*(NOV), 1–12. https://doi.org/10.3389/fpsyg.2013.00863

Law, M., Stewart, D., Pollock, N., Letts, L., Bosch, J., & Westmorland, M. (1998a). *Critical Review Form - Quantitative Studies (Protocol)* (pp. 1–3). McMaster University.

Law, M., Stewart, N., Pollock, D., Letts, L., Bosch, J., & Westmorland, M. (1998b). *Guidelines for Critical Review Form - Quantitative Studies*. Guidelines for Critical Review Form - Quantitative Studies. http://srs-mcmaster.ca/wp-content/uploads/2015/04/Guidelines-for-Critical-Review-Form-Quantiative-Studies-English.pdf

Lenhard, W. & L. A. (2016). *Computation of effect sizes*.

Loades, M. E., Chatburn, E., Higson-Sweeney, N., Reynolds, S., Shafran, R., Brigden, A., Linney, C., McManus, M. N., Borwick, C., & Crawley, E. (2020). Rapid systematic review: The impact of social isolation adolescents in the context of COVID-19. *Journal of the American of Child & Adolescent Psychiatry*, *59*(11), 1218–1239.

Lyneham, H. J., & Rapee, R. M. (2005). Agreement between telephone and in-person delivery of a structured interview for anxiety disorders in children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*(3), 274–282. https://doi.org/10.1097/00004583-200503000-00012

Lyneham, H. J., Sburlati, E. S., Abbott, M. J., Rapee, R. M., Hudson, J. L., Tolin, D. F., & Carlson, S. E. (2013). Psychometric properties of the

Child Anxiety Life Interference Scale (CALIS). *Journal of Anxiety Disorders*, *27*(7), 711–719. https://doi.org/10.1016/j.janxdis.2013.09.008

Mazzone, L., Ducci, F., Scoto, M. C., Passaniti, E., D'Arrigo, V. G., & Vitiello, B. (2007). The role of anxiety symptoms in school performance in a community sample of children and adolescents. *BMC Public Health*, *7*, 1–6. https://doi.org/10.1186/1471-2458-7-347

McLellan, L. F., Andrijic, V., Davies, S., Lyneham, H. J., & Rapee, R. M. (2017). Delivery of a Therapist-Facilitated Telecare Anxiety Program to Children in Rural Communities: A Pilot Study. *Behaviour Change*, *34*(3), 156–167. https://doi.org/10.1017/bec.2017.11

Mychailyszyn, M. P. (2017). "Cool" youth: A systematic review and comprehensive meta-analytic synthesis of data from the Cool Kids family of intervention programs. *Canadian Psychology*, *58*(2), 105–115. https://doi.org/10.1037/cap0000101

OECD. (2021). *OECD Secretary-General's Report to Ministers*. https://doi.org/10.1787/8cd95b77-en.

Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: horses for courses. *Journal of Epidemiology and Community Health*, *57*(7), 527 LP – 529. https://doi.org/10.1136/jech.57.7.527

Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *56*(3), 345–365. https://doi.org/10.1111/jcpp.12381

Racine, N., McArthur, B. A., Cooke, J. E., Eirich, R., Zhu, J., & Madigan, S. (2021). Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19. *JAMA Pediatrics*, *175*(11), 1142–1150. https://doi.org/10.1001/jamapediatrics.2021.2482

Ravens-Sieberer, U., Kaman, A., Erhart, M., Devine, J., Schlack, R., & Otto, C. (2021). Impact of the COVID-19 pandemic on quality of life and mental health in children and adolescents in Germany. *European Child and Adolescent Psychiatry*. https://doi.org/10.1007/s00787-021-01726-5

Reynolds, S., Wilson, C., Austin, J., & Hooper, L. (2012). Effects of psychotherapy for anxiety in children and adolescents: A meta-analytic review. *Clinical Psychology Review*, *32*(4), 251–262. https://doi.org/10.1016/j.cpr.2012.01.005

Scaini, S., Rossi, F., Rapee, R. M., Bonomi, F., Ruggiero, G. M., & Incerti, A. (2022). The Cool Kids as a School-Based Universal Prevention and Early Intervention Program for Anxiety: Results of a Pilot Study. *International Journal of Environmental Research and Public Health*, *19*(2), 941. https://doi.org/10.3390/ijerph19020941Sciberras, E., Mulraney, M., Anderson, V., Rapee, R. M., Nicholson, J. M., Efron, D., Lee, K., Markopoulos, Z., & Hiscock, H. (2018). Managing Anxiety in Children With ADHD Using Cognitive-Behavioral Therapy: A Pilot Randomized Controlled Trial. *Journal of Attention Disorders*, *22*(5), 515–520. https://doi.org/10.1177/1087054715584054

Silverman, W. K., Saavedra, L. M., & Pina, A. A. (2001). Test-Retest Reliability of Anxiety Symptoms and Diagnoses with the Anxiety

Disorders Interview Schedule for DSM-IV: Child and Parent Versions. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(8), 937–944. https://doi.org/10.1097/00004583-200108000-00016

Spence, S. H. (1997). *Spence Children's Anxiety Scale (SCAS)* [Database record]. APA PsycTests. https://doi.org/10.1037/t10518-000

Spence, S. H., Barrett, P. M., & Turner, C. M. (2003). Psychometric properties of the Spence Children's Anxiety Scale with young adolescents. *Journal of Anxiety Disorders*, *17*(6), 605–625. https://doi.org/10.1016/S0887-6185(02)00236-0

Swift, P., Cyhlarova, E., Goldie, I., & O'Sullivan, C. (2014). Living with anxiety: Understanding the role and impact of anxiety in our lives. In *Mental Health Fondation* (Vol. 43, Issue 1).

Walters, T., Simkiss, N. J., Snowden, R. J., & Gray, N. S. (2021). Secondary school students' perception of the online teaching experience during COVID-19: The impact on mental wellbeing and specific learning difficulties. *British Journal of Educational Psychology*. https://doi.org/10.1111/bjep.12475

Zafeiriou, M. E., & Gulliford, A. (2020). A grounded theory of educational psychologists' mental health casework in schools: connection, direction and reconstruction through consultation. *Educational Psychology in Practice*, *36*(4), 422–442. https://doi.org/10.1080/02667363.2020.1818553

Appendices

Appendix A

Study excluded at full text screening

Excluded Study	Rationale for Exclusion
Gould, K. L., Porter, M., Lyneham, H. J., & Hudson, J. L. (2018). Cognitive-Behavioral Therapy for Children With Anxiety and Comorbid Attention-Deficit/Hyperactivity Disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *57*(7), 481-490.e2. https://doi.org/10.1016/j.jaac.2018.03.021	Exclusion Criteria: 6

Appendix B

Mapping the field

	Author and geographical distribution	Participants (*n*)	Study/ Design	Intervention	Context of intervention & Setting	Interventionist	Outcome variables measured	Follow up	Results	Effect Size
1	Johnsen et al. (2019) Denmark	7- to 16-year-olds who met the criteria for an anxiety disorder*. *n*=54	Cohort (Open trial)	Cool Kids CBT programme with no adaptations	10x 2-hour weekly group sessions over 12 weeks. University & shopping mall	Student therapists	Anxiety levels using CALIS-Y$^{\nabla}$ & SCAS-Y$^{\nabla\nabla}$	3- month follow up with 1-hour booster	SCAS-Y $^{\nabla\nabla}$ p < .001 CALIS-Y$^{\nabla}$ p < .001	SCAS-Y $^{\nabla\nabla}$ g = 1.08* CALIS-Y $^{\nabla}$ g = .99*
2	Sciberras et al. (2018) Australia	Males aged 8 to 12 years old who met the criteria for an anxiety disorder* and ADHD. *n*=11	Randomised controlled trial	Cool Kids CBT programme used on individual basis with some adaptations (four in total)	8x 1-hour weekly sessions & 2x 1 hour bi-weekly sessions (10 weeks in total). Unknown setting	Trained facilitators	Anxiety levels using CALIS-C$^{\nabla}$ & SCAS-C$^{\nabla\nabla}$	5- month follow up	Not reported	SCAS-C$^{\nabla\nabla}$ d = 0.2** CALIS-C$^{\nabla}$ d = 0.4***
3	Arendt et al. (2016a) Denmark	Males and females aged 7 to 16-years old who met the criteria for an anxiety disorder*. *n*=98	Cohort (Open trial)	Cook Kids CBT programme translated to Danish	10x 2-hour group sessions. University	Clinical psychology & graduate students	Anxiety levels using SCAS-C$^{\nabla\nabla}$	3-month follow up with 1-hour booster	Not reported	SCAS-C$^{\nabla\nabla}$ d = 0.86****

4	Arendt et al. (2016b) Denmark	7- to 16-year-olds who met the criteria for an anxiety disorder*. *n*=109	Randomised wait-list controlled trial	Cool Kids CBT programme with Danish translated resources	10x 2-hour weekly sessions. University & shopping mall	Psychologists & graduate students	Anxiety levels using CALIS-Y & SCAS-Y	3-month follow up with 1 hour booster	SCAS-Y $p < .001$ CALIS-Y $p < .008$	SCAS-Y $d = 1.08$ CALIS-Y $d = 0.63$
5	McLellan et al. (2017) Australia	9 – 12-year-olds who met the criteria for an anxiety disorder*. *n*=16	Cohort study	Cool Kids CBT programme adapted for online use	10x 45-minute weekly sessions & 4x 30-minute optional phone calls to parent training. Video conferencing via school setting	Clinical psychologists and identified school adult	Anxiety levels using CALIS-C & SCAS-C	None	SCAS-C $p < .001$ CALIS-C $p < .001$	SCAS-C $d = 1.45$ CALIS-C $d = 0.65$****

Note: Each study used the Anxiety Disorder Interview Schedule for DSM-IV. **g* is for Hedges' g which is a measure of effect size comparable to Cohen's *d.* The parameters for both are the same. ** *d* is an effect size referred to as Cohen's d. *** Effect sizes for Sciberras et al. (2018) were reported for 5-month follow up only. **** This effect size was manually calculated by the author of this review. $^{\nabla}$ Lyneham et al., 2013 $^{\nabla\nabla}$ Spence, 1997

Appendix C

Weight of Evidence A – Methodological Quality

The following coding protocol was adapted for the current review from Law et al. (1998a). It is a specific coding protocol for quantitative studies. The 'design' indicator was removed as this is directly determined in Weight of Evidence B. References to occupational therapy were removed as this review was focussing on anxiety in school age pupils not related to occupational therapy needs. The guidance document (Law et al., 1998b) was used to create a numerical score-based system allowing for coding (see Table 1).

Table 1

Coding Protocol for Weight of Evidence A

Dimension 1: Study Purpose	Criteria
Score of 0:	No clear purpose statement is given.
Score of 1:	There is a partial statement of the study purpose in the abstract/introduction.
Score of 2:	The study purpose is clear, detailing the importance and relevance of the topic.
Dimension 2: Literature	**Criteria**
Score of 0:	The literature review gives minimal relevant information and no reference to the importance of the implications of this research.
Score of 1:	The literature review gives a broad overview of information related to the study and gives a weak justification for the importance of the topic.

Score of 2:	The literature review gives a clear and purposeful synthesis of relevant information including a summary of past research and strong justification for the importance of the topic.
Dimension 3: Design	Criteria
Score of 0:	The design is defined, and no control group is used.
Score of 1:	The design is not clearly defined, and a control group is used.
Score of 2:	The design is clearly defined, and a control group is used.
Dimension 4: Sample	Criteria
Score of 0:	Sample characteristics are not clear, and it is not clear if informed consent was given. It is unclear how the sample was obtained.
Score of 1:	Sample characteristics are defined with ambiguity, sample size is given, informed consent was given. It is clear how the sample was obtained.
Score of 2:	Clearly defined characteristics of the sample, sample size is given, it is clear informed consent was given. It is clear how the sample was obtained.
Dimension 5: Outcomes	Criteria
Score of 0:	Outcomes are described. The measures are not described/ reliability or validity is not addressed. The frequency of the measures is not clear (pre/post-test and follow up).
Score of 1:	Outcomes are not clearly described. The measures are described, reliable and valid. The frequency of the measures is clear (pre/post-test and follow up).

Score of 2:	Outcomes are clearly described. The measures are described, reliable and valid. The frequency of the measures is clear (pre/post-test and follow up).
Dimension 6: Interventions	Criteria
Score of 0:	Intervention is described with minimal detail making it difficult to replicate.
Score of 1:	Intervention is described, detailing who delivered the intervention, how often treatment was received and the setting for the intervention for each group (if appropriate).
Score of 2:	Intervention is described, detailing who delivered the intervention, how often treatment was received, the setting for the intervention for each group (if appropriate) and acknowledges how cointervention was minimised.
Dimension 7: Results	
Score of 0:	Statistical methods and results are not clear, statistical significance is reported without explanation and attrition is not recognised.
Score of 1:	Only some of the criteria for a score of 2 are evident.
Score of 2:	Appropriate statistical methods are used, results are reported in terms of statistical significance, effect sizes are identified accurately, and attrition is recognised.
Dimension 8: Conclusion and Clinical Implications	Criteria
Score of 0:	The conclusion is not clear, and findings are not appropriate.

Score of 1:	Appropriate conclusions were detailed, implications of conclusions were given but no recommendations for further study are given.
Score of 2:	Appropriate conclusions were detailed, implications of conclusions were given and recommendations for further study are given.

*RCT – Randomised controlled trial

Table 2

WoE A overall weighting scores for studies

Study Citation	Study Purpose	Literature	Sample	Outcomes	Interventions	Results	Conclusion and Clinical Implications	WoE A
Johnsen et al. (2019)	2	2	2	2	2	2	2	2
Sciberras et al. (2018)	1	1	1	0	2	1	2	1.14
Arendt et al. (2016a)	2	2	0	2	1	1	1	1.29
Arendt et al. (2016b)	2	2	2	2	2	2	1	1.86
McLellan et al. (2017)	2	2	2	0	1	2	2	1.58

Note: WoE A ratings are described as 'Low' for scores ≤1.0, 'Medium' for scores 1.1 - 1.5, and 'High' for scores ≥ 1.6. Overall weightings were an average of the seven criteria.

Appendix D

Weight of Evidence B

Weight of Evidence B (WoE B) is a review-specific judgement in which consideration of the relevance of the methodology is used for the research identified (Gough, 2007). The criteria selected, based on the Typology of Evidence criteria by Petticrew and Roberts (2003) can be found in Table 1.

Table 1

WoE B Criteria

Study Design	WoE B rating	Rationale
Randomised controlled trial (RCT)	2 (High)	RCTs are the most effective designs for questions related to effectiveness (Petticrew & Roberts, 2003).
Quasi-experimental, Cohort or Open Trial* studies	1 (Medium)	The next most effective study type for effectiveness questions are Quasi-experimental, Cohort or Open Trial* studies (Petticrew & Roberts, 2003).
Qualitative research, survey, case-control study and non-experimental evaluations.	0 (Low)	Qualitative research, survey, case-control study and non-experimental evaluations are considered the least effective design for effectiveness questions (Petticrew & Roberts, 2003).

**Note*: For the purpose of this review, Cohort and Open Trials will be given the same weighting due to their comparability.

Table 2

WoE B overall weighting scores for studies

Study	Overall WoE B
Johnsen et al. (2019)	1 (Low)
Sciberras et al. (2018)	2 (High)
Arendt et al. (2016a)	1 (Low)
Arendt et al. (2016b)	1 (Low)
McLellan et al. (2017)	1 (Low)

Note: WoE B ratings are described as 'Low' for scores ≤1.0, 'Medium' for scores 1.1 - 1.5, and 'High' for scores ≥ 1.6.

Appendix E

WoE C Criteria and Rationale

Weight of Evidence C (WoE C) is a review-specific judgement that considers how relevant the research is in relation to the review question (Gough, 2007). The criteria used for WoE C can be found in Table 1.

Table 1

Weight of Evidence C Criteria

Criteria	Scoring	Rationale
A: Intervention and Setting	2 = The intervention is adapted for minimal parental involvement and use in schools. 1= The intervention is adapted for minimal parental involvement 0 = There are no adaptation to the intervention.	This review is assessing the effectiveness of a manualised CBT intervention for reducing anxiety in school aged pupils with the intention of the intervention being carried out in school. Therefore, studies that adapt the Cool Kids intervention for minimal to no parental input will be more relevant.
B: Participants	2 = CYP with a recognised anxiety disorder according to a valid and reliable measure (e.g. the Anxiety Disorders Schedule for DSM -IV). 1= CYP with anxiety and comorbidity with another identified need. 0= No clear information regarding the needs of the participants.	This review considers the effectiveness of a CBT intervention in the reduction of a diagnosed anxiety disorder. Therefore, studies with comorbidity will be less relevant to the review question.

C: Instructor	2 = The intervention was delivered in collaboration with teaching staff 1 = The intervention was delivered by trainees and/or trained staff 0 = The intervention was delivered by a trained psychologist staff only.	This review aims to make recommendations for use of the CBT programme in UK school settings. Therefore, studies that include teaching staff within a school setting will be more relevant to the research question.
D: Location	2 = Conducted in OECD country with similar education systems/transition stages. 1 = Conducted in OECD country with dissimilar education systems/transition stages in comparison to the UK. 0 = Not conducted in an OECD member country	This review aims to make recommendations for use of the CBT programme in UK school settings. Therefore, studies conducted in countries (including those that are members of the Organisation for Economic Co-Operation and Development (OECD, 2021) with similar education systems to the UK will be more relevant.

Table 2

Weight of Evidence C

Study	A	B	C	D	WoE C
Johnsen et al. (2019)	0	2	1	1	1 (Low)
Sciberras et al. (2018)	0	1	0	2	0.75 (Low)
Arendt et al. (2016a)	0	2	1	1	1 (Low)
Arendt et al. (2016b)	0	2	1	1	1 (Low)
McLellan et al. (2017)	2	2	2	2	2 (High)

Note: WoE C ratings are described as 'Low' for scores ≤1.0, 'Medium' for scores 1.1 - 1.5, and 'High' for scores ≥ 1.6. Overall weightings were an average of the four scores.

Appendix F

Weight of Evidence A adapted Coding Protocol (Example)

Critical Review Form – Quantitative Studies

Law, M., Stewart, D., Pollock, N., Letts, L., Bosch, J., & Westmorland, M., 1998 McMaster University

Citation:	
Comments	
Study purpose: Was the purpose stated clearly? Yes/No	Outline the purpose of the study. How does the study apply to your research question?
Literature: Was relevant background literature reviewed? Yes/No	Describe the justification of the need for this study.
Design: Randomized (RCT) Cohort Single Case Design Before and after Case-control Cross-sectional Case Study	Describe the study design. Was the design appropriate for the study question? (e.g., for knowledge level about this issue, outcomes, ethical issues, etc.) Specify any biases that may have been operating and the direction of their influence on the results.

Sample: N= Was the sample described in detail? Yes/No Was the sample justified? Yes/No/N/A	Sampling (who; characteristics; how many; how was sampling done?) If more than one group, was there similarity between the groups? Describe ethics procedures. Was informed consent obtained?
Outcomes: Were the outcome measures reliable? Yes/No/Not addressed Were the outcome measures valid? Yes/No/Not addressed	Specify the frequency of outcome measurement (i.e. pre, post, follow-up) Outcome areas (e.g. self-care, productivity, leisure): List measures used:
Intervention: Intervention was described in detail? Yes/No/Not addressed Contamination was avoided: Yes/No/Not addressed/N/A	Provide a short description of the intervention (focus, who delivered it, how often, setting). Could the intervention be replicated?

Cointervention was avoided: Yes/No/Not addressed/ N/A	
Results: Results were reported in terms of statistical significance? Yes/No/Not addressed/ N/A Were the analysis method(s) appropriate? Yes/No/Not addressed Clinical Importance was reported? Yes/No/Not addressed	What were the results? Were they statistically significant (i.e., $p < 0.05$)? If not statistically significant, was study big enough to show an important difference if it should occur? If there were multiple outcomes, was that taken into account for the statistical analysis? What was the clinical importance of the results? Were differences between groups clinically meaningful? (if applicable)
Drop-outs were reported? Yes/No	Did any participants drop out from the study? Why? (Were reasons given and were drop-outs handled appropriately?)

Conclusions and clinical implications: Conclusions were appropriate given study methods and results Yes/No	What did the study conclude? What are the implications of these results? What were the main limitations or biases in the study?

Appendix G - Completed Weight of Evidence A Coding Protocols for 5 reviewed studies

Critical Review Form – Quantitative Studies

Law, M., Stewart, D., Pollock, N., Letts, L., Bosch, J., & Westmorland, M., 1998 McMaster University

Citation: Johnsen, D. B., Arendt, K., & Thastum, M. (2019). The efficacy of manualized Cognitive Behavior Therapy conducted by student-therapists treating Danish youths with anxiety using a benchmark comparison. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology*, *7*, 68–80. https://doi.org/10.21307/sjcapp-2019-010	
Comments	
Study purpose: Was the purpose stated clearly? Yes/No	Outline the purpose of the study. How does the study apply to your research question? To investigate the effects of student therapist (ST) delivered intervention on treating youths with anxiety using Cool Kids. These are compared to outcomes achieved by professional therapists (PT). I want to know if the deliverer of the training is important as my question relates to using the programme in school with the potential of it being delivered by school staff.
Literature: Was relevant background	Describe the justification of the need for this study. The paper detailed issues related to youth anxiety and how efficacious CBT was in treating these issues. The hope of the paper was to find out if the use of ST would produce

literature reviewed? Yes/No	similar results to PT therefore giving the opportunity to increase the number of therapists available to deliver the programme.
Design: Randomized (RCT) Cohort (Open Trial) Single Case Design Before and after Case-control Cross-sectional Case Study	Describe the study design. Was the design appropriate for the study question? (e.g., for knowledge level about this issue, outcomes, ethical issues, etc.) This question is considered under Weight of Evidence B so is being removed. Specify any biases that may have been operating and the direction of their influence on the results. No biases identified.
Sample: N= 54 in intervention/56 from control group used for benchmarking Was the sample described in detail? Yes/No Was the sample justified? Yes/No/N/A	Sampling (who; characteristics; how many; how was sampling done?) If more than one group, was there similarity between the groups? Procedures for both groups were identical for inclusion and assessment. They met criteria for primary diagnosis for an anxiety disorder as defined by the DSM-IV. Exclusion criteria were clear. The groups were not consistent in ages researched. Gender was not accounted for in the sample detailing. The participants sought help from the CEBU. The groups were not fully comparable as between group characteristics are not matched.

	Describe ethics procedures. Was informed consent obtained? Written consent forms were completed by all families.
Outcomes: Were the outcome measures reliable? Yes/No/Not addressed Were the outcome measures valid? Yes/No/Not addressed	Specify the frequency of outcome measurement (i.e. pre, post, follow-up) 12-week programme. Pre-test, post-test and a 3 month follow up. Follow up allows for a more complex understanding of the effect of the intervention (Gersten et al., 2005). Outcome areas (e.g. self-care, productivity, leisure): Reduction in reported anxiety. List measures used: For this question, we are focusing on the results of the CALIS Youth and SCAS Youth. This is because we want to see if the youth results show changes in levels of anxiety.
Intervention: Intervention was described in detail? Yes/No/Not addressed Contamination was avoided: Yes/No/Not addressed/N/A	Provide a short description of the intervention (focus, who delivered it, how often, setting). Could the intervention be replicated? Used cool kids intervention which involves youth and parents. 10x 2-hour weekly group sessions over 12 weeks. Parents received psychoeducation in parent management strategies. Sessions took place at CEBU with one session in a shopping mall. 3-months after a 1-hour booster session

Cointervention was avoided: Yes/No/Not addressed/ N/A	was offered. Treatment in PT group was delivered by a psychologist and 3 graduate psychology students. Treatment in the ST group was purely delivered by 3 graduate students. Although there is a lack of detail enabling this intervention to be replicable, this is because it is a manualised programmed designed for delivery by trained professionals, therefore by limiting the level of detail, it minimises the risk of untrained individuals attempting to deliver the programme. Participants were encouraged not to engage in other forms of treatment or change their psychopharmacological medication. Contamination was not a concern as the programme can only be delivered by trained individuals.
Results: Results were reported in terms of statistical significance? Yes/No/Not addressed/ N/A Were the analysis method(s) appropriate? Yes/No/Not addressed	What were the results? Were they statistically significant (i.e., p < 0.05)? If not statistically significant, was study big enough to show an important difference if it should occur? If there were multiple outcomes, was that taken into account for the statistical analysis? Pre/post SCAS Y ST - p<0.001, g=1.08 and PT- p<0.001, g=1.03 CALIS Y ST- p<0.001, g=0.99 and PT p<0.001, g=0.61 Large overall effect size supported by statistically significant results. Pre/follow up

Clinical Importance was reported? Yes/No/Not addressed	SCAS Y ST - p<0.01, g=1.30 and PT p<0.01, g=1.33 CALIS Y ST p<0.01, g=0.80 and PT p<0.01, g=0.88 What was the clinical importance of the results? Were differences between groups clinically meaningful? (if applicable) Yes (see conclusions below).
Drop-outs were reported? Yes/No	Did any participants drop out from the study? Why? (Were reasons given and were drop-outs handled appropriately?) Attrition was noted in detail but reasons were not given.
Conclusions and clinical implications: Conclusions were appropriate given study methods and results Yes/No	What did the study conclude? What are the implications of these results? What were the main limitations or biases in the study? Conclusion was that ST group with limited training could provide CBT to youth with anxiety disorders with comparable treatment outcomes to the PT group. A limitation was the difference in mean age for the groups as severity of anxiety and onset is heterogeneous.