



Treatment Fidelity in Brief Versus Standard-Length School-Based Interventions for Youth with Anxiety

Elisabeth Husabo^{1,2} · Bente S. M. Haugland^{1,3} · Bryce D. McLeod⁴ · Valborg Baste⁵ · Åshild Tellefsen Haaland⁶ · Jon Fauskanger Bjaastad^{2,7} · Asle Hoffart^{8,9} · Solfrid Raknes² · Krister W. Fjermestad⁹ · Ronald M. Rapee¹⁰ · Terje Ogden¹¹ · Gro Janne Wergeland^{1,12}

Accepted: 4 June 2021 / Published online: 13 June 2021
© The Author(s) 2021

Abstract

To examine treatment fidelity in a randomized controlled trial of indicated school-based cognitive-behavioral therapy (CBT) delivered in groups to youth with anxiety. We investigated whether adherence and competence (a) differed across brief and standard-length CBT, and (b) if adherence and competence predicted change in anxiety symptoms and impairment. Method: Sessions were observationally coded with the *Competence and Adherence Scale for CBT for Anxiety Disorders in Youth*. Coders ($N=7$) rated 104 sessions from 52 groups delivered by 32 facilitators (M age = 43.2 years, $SD=8.1$) to 295 youth (M age = 14.0 years, $SD=0.8$). Outcomes were youth- and parent-reported anxiety symptoms and impairment at post-intervention and 1-year follow-up. Linear mixed effect models were used to analyze whether fidelity predicted clinical outcomes. Results: Levels of adherence and competence were adequate in both programs, but higher in brief compared to standard-length CBT ($p < .001$ and $p = .010$, respectively). Neither adherence nor competence predicted clinical outcomes at any timepoints. Conclusion: Higher levels of adherence and competence in brief CBT suggest that it may be easier for novice CBT providers to achieve fidelity in simplified and less flexible interventions. Contrary to expectation, adherence and competence did not predict clinical outcomes.

Keywords Treatment fidelity · Adherence · Competence · School-based · Brief CBT

Introduction

School settings represent an ideal opportunity to provide support to youth at risk for anxiety disorders by offering effective interventions at school, during school hours (Lyon

& Bruns, 2019). Subclinical levels of anxiety are common among youth (Balazs et al., 2013), associated with significant impairment (Angold et al., 1999), and increase the likelihood that youth will develop anxiety disorders (Pine, 2007). The majority of studies of indicated school-based

✉ Elisabeth Husabo
elisabeth@husabo.no

¹ Division of Psychiatry, Department of Child and Adolescent Psychiatry, Haukeland University Hospital, N-1400, 5021 Bergen, Norway

² RKBU West-Regional Centre for Child and Youth Mental Health and Child Welfare, NORCE Norwegian Research Centre, N-7810, 5020 Bergen, Norway

³ Department of Clinical Psychology, University of Bergen, N-5015 Bergen, Norway

⁴ Department of Psychology, Virginia Commonwealth University, PO Box 842018, Richmond, VA 23284-2018, USA

⁵ NORCE Norwegian Research Centre, N-7810, 5020 Bergen, Norway

⁶ Department of Child and Adolescent Mental Health, Sorlandet Hospital HF, N-416 Lundsiden, 4604 Kristiansand S, Norway

⁷ Division of Psychiatry, Stavanger University Hospital, N-4060 Stavanger, Norway

⁸ Research Institute, Modum Bad Psychiatric Centre, Badeveien 287, 3370 Vikersund, Norway

⁹ Department of Psychology, University of Oslo, Blindern, Po box 1094, 0317 Oslo, Norway

¹⁰ Centre for Emotional Health, Department of Psychology, Macquarie University, Sydney, NSW 2109, Australia

¹¹ The Norwegian Centre for Child Behavioral Development, N-7053, Majorstuen, 0306 Oslo, Norway

¹² Department of Clinical Medicine, University of Bergen, PB 7800, 5020 Bergen, Norway

prevention provided to youth with anxiety has been on cognitive behavioral therapy (CBT), and these studies have shown small to moderate effects (e.g., Haugland et al., 2020; Stoll et al., 2020; Werner-Seidler et al., 2017). However, barriers to implementing CBT programs in school settings exist. For example, schools may not have (enough) mental health professionals, lack the time and resources required to implement a program, or be reluctant to invest time for training and consultation (Rasmussen et al., 2019; Weist et al., 2017). Addressing these barriers may help improve the implementation and the effectiveness of indicated CBT programs in school settings.

Design of CBT Programs

Many CBT programs are designed to be delivered by mental health professionals (Muggeo et al., 2017; Werner-Seidler et al., 2017). Yet the extent to which mental health professionals are available to deliver CBT programs in schools varies across school districts (Langley et al., 2010). If treatments can be delivered by personnel with less mental health training in schools (e.g., school health nurses), then this may increase the number of individuals available to deliver mental health treatments (Kakuma et al., 2011). However, if a mismatch between the design of CBT programs and the training background of school personnel exists, then it may serve to undermine the effectiveness by impacting the delivery of these programs (Lyon & Koerner, 2016).

CBT programs can be designed to make them easy to deliver for school personnel with minimal prior CBT training (i.e., novice CBT providers; Bennett-Levy et al., 2010). First, program materials can be made simple (e.g., minimize the level of possible tailoring of the interventions to each client), because program protocols with higher complexity may require more clinical experience to deliver with treatment fidelity (Lyon & Koerner, 2016). Second, more structure can be introduced (e.g., more detailed instructions, assigning number of minutes that should be spent on each component in the sessions) to increase support for providers. A high degree of flexibility may create uncertainty in how to deliver a program (Chorpita & Daleiden, 2014). It is possible that increased structure and simple presentation materials for program content may make it easier for novice CBT providers to deliver CBT programs in school settings, but few studies have tested this question.

Treatment Fidelity

Measurement of treatment fidelity can be used to determine if efforts to create simplified and more structured CBT programs are successful (Proctor et al., 2011). Treatment fidelity concerns the extent to which a program is delivered as designed (Schoenwald et al., 2011). In mental

health research, the term often includes three components: (a) Adherence—the degree to which the core elements of a program are delivered, (b) Competence—how skillfully the core elements of a program are delivered, and (c) Differentiation—the extent to which non-prescribed program content is avoided (Perepletchikova et al., 2007; Southam-Gerow et al., 2021). In this paper, we focus on two of these components, adherence, and competence, since our focus is on characterizing the delivery of two interventions.

One way to gauge if efforts to make CBT programs easy to deliver by novice CBT providers are successful is to determine if treatment fidelity is higher in a brief, simple and structured program than in a standard-length program, which may be more complex and with more room for flexibility. We tested this in a study where both programs were delivered by the same providers, training, and supervision were held constant across groups, and where both programs were effective in reducing anxiety symptoms (Haugland et al., 2020).

Beyond understanding the impact of program design on treatment fidelity, examining if treatment fidelity predicts clinical outcomes in school settings can inform implementation efforts. In the treatment evaluation literature, findings have been mixed regarding the relation between treatment fidelity and clinical outcomes (Collyer et al., 2019; Rapley & Loades, 2018; Webb et al., 2010). Some studies have found that higher adherence (Hogue et al., 2008; Podell et al., 2013) or higher competence (Podell et al., 2013) predict improvements in clinical outcomes. For example, a study of CBT for anxiety in community mental health clinics (hereafter called *community clinics*) found that higher adherence predicted improved clinical outcomes, and that competence among clinicians with two years of CBT training predicted improved clinical outcomes (Bjaastad et al., 2018). Yet most studies have found no significant prediction by adherence (Hartnett et al., 2016; Heywood & Fergusson, 2016; Liber et al., 2010; Overbeek et al., 2013; Southam-Gerow et al., 2021) or competence (Garner et al., 2012; Hogue et al., 2008; Southam-Gerow et al., 2021) on clinical outcomes. Most studies that have investigated whether fidelity predicts outcome focus on mental health professionals delivering care in university or outpatient mental health settings (see Collyer et al., 2019; Webb et al., 2010). Thus, it is an open question whether these findings will generalize to novice CBT providers delivering CBT programs in school settings (Mellin et al., 2011).

To date, only two studies have looked at whether fidelity predicts clinical outcomes in CBT for anxiety in school-based programs. The first study, with two published articles, indicated that greater adherence in the use of CBT session structure components (e.g., agenda-setting and homework assignment) and higher competence in administering these components, predicted better clinical outcomes (i.e.,

symptom reduction and diagnostic recovery; Ginsburg et al., 2012). In contrast, greater adherence in delivering specific CBT modules (e.g., psychoeducation, exposure) did not predict better clinical outcomes (Becker et al., 2012). The authors hypothesized that the lack of prediction was due to providers not having enough training and supervision to deliver the program with optimal skills (Becker et al., 2012; Ginsburg et al., 2012). The second study found that provider-rated adherence predicted anxiety symptom improvement at 3-month follow-up (Caron et al., 2020). However, both studies were limited by small samples size ($N = 16$ and $N = 54$), making it difficult to draw conclusions about the relation between treatment fidelity and clinical outcomes.

Aims of the Study

The current study investigated competence and adherence in two school-based CBT programs and how treatment fidelity predicts outcomes. Three research questions were investigated: (a) do adherence and competence differ by CBT program (brief and standard-length), (b) do adherence and competence predict change in anxiety symptoms and impairment at post-intervention and at 1-year follow-up, and (c) do the brief and standard-length CBT programs differ in the prediction of outcome by adherence and competence? Investigating these questions can help determine what programs are most suitable to implement in schools by novice providers and ascertain whether the association between fidelity and clinical outcomes differs across CBT programs with different designs. To investigate these questions, the current study presents secondary analyses from a multisite, school-based randomized controlled trial (RCT) evaluating the effectiveness of two group-based indicated CBT programs for youth with elevated anxiety symptoms compared to a waitlist (WL) control: a 5-session CBT program called Vaag (Raknes et al., 2015) and a 10-session CBT program called Cool Kids (CK; Rapee et al., 2006).

Methods

In the RCT, randomization was determined prior to inclusion, and groups of five to eight youth from each school were randomized to one of three conditions: Vaag (15 groups; $n = 91$), CK (19 groups; $n = 118$), or WL (18 groups; WL; $n = 104$). Youth and parents completed assessments at pre-intervention, post-intervention, and at 1-year follow-up. The study was approved by the Regional Committee for Medical and Health Research Ethics, Western Norway (approval no 2013/2331). Participants including youth, parents and facilitators provided written informed assent/consent. See Haugland et al. (2017) and Haugland et al. (2020) for details about procedures and outcomes (NCT02279251, clinicaltrials.gov).

The RCT showed that school-based CBT was effective in reducing anxiety symptoms and impairment relative to WL (Haugland, 2020). The RCT involved 52 CBT groups (24 Vaag, 28 CK) delivered at school, during school hours. Within-group (pre-post) effect sizes (Cohen's d) for youth- and parent-reported anxiety ranged from 0.41–0.53 in Vaag and 0.62–0.67 in CK, respectively. Thus, within-group effect sizes were generally larger for CK compared to Vaag. Furthermore, Vaag was deemed *not* non-inferior to CK. Outcomes were maintained or improved at 1-year follow-up for both interventions (Haugland et al., 2020). Facilitators achieved adequate adherence and competence in both programs. Fidelity was measured by the Competence and Adherence Subscale for CBT for Anxiety Disorders in Youth (Bjaastad et al., 2016). Adherence and competence scores for each group (mean of 2 rated sessions) ranged from 3.17 to 5.75 ($M = 4.41$, $SD = 0.56$) for adherence and 2.75 to 5.88 ($M = 4.18$, $SD = 0.66$) for competence.

Participants and Procedure

Youth Participants

The RCT included 313 youth (M age = 14.0 years, $SD = 0.8$; range 12–15 years, 84% female, 16% male). They were from high (29.7%), medium (62.3%), and low (8.0%) social class families, defined by the Registrar General Social Class coding scheme (Currie et al., 2008). The youth were recruited from 18 schools (17 public, 1 private) in urban and rural areas, from October 2014 to November 2016. Youth were invited to participate in the study if (a) either self- or parent-reported youth anxiety symptoms were ≥ 25 on the Spence Children's Anxiety Scale (SCAS; Spence, 1998), and (b) interference in daily life was indicated by a score of ≥ 1 on the first question of the Child Anxiety Life Interference Scale (CALIS; self- and/or parent-report; Lyneham, 2013). See Table 1 for baseline youth and family characteristics.

Youth participants in the RCT who attended at least one CBT session ($N = 295$; M age = 14.0, $SD = 0.8$; range 12–15 years; 85% female, 15% male) were included in the present study. Of these, 252 were defined as treatment completers (attending ≥ 4 sessions of Vaag or ≥ 7 sessions of CK), with the remaining defined as non-completers ($n = 43$). Retention rates were not different between interventions ($p = 0.68$; Haugland et al., 2020).

Facilitators

Thirty-two facilitators (M age = 43.2 years, $SD = 8.1$, range 32–62 years; 93.8% female, 6.2% male) delivered the CBT programs. The facilitators were recruited from school and community services (school nurses $n = 21$, community psychologists $n = 5$, family therapist $n = 1$) or from community

Table 1 Youth- and parent-reported baseline youth and family characteristics

Variables	Brief CBT (<i>n</i> = 138)		Standard-length CBT (<i>n</i> = 157)		<i>t</i> or χ^2	<i>p</i>	Total (<i>N</i> = 295)	
	<i>M</i> (SD)	<i>n</i> (%)	<i>M</i> (SD)	<i>N</i> (%)			<i>M</i> (SD)	<i>N</i> (%)
Age	13.91 (0.86)		14.04 (0.81)		−1.36	.174	13.98 (0.84)	
Sex								
Male		21 (15.2)		24 (15.3)	.000	.987		45 (15.3)
Ethnicity					.73	.394		
Norwegian		132 (95.7)		153 (97.5)				285 (96.6)
Non-Norwegian		6 (4.3)		4 (2.5)				10 (3.4)
Family structure					0.15	.703		
Two-parent family		109 (79.0)		126 (80.3)				235 (79.7)
Single parent family		29 (21.0)		30 (19.1)				59 (20.0)
Social class					2.07	.356		
High		41 (29.7)		38 (24.2)				79 (26.9)
Medium		86 (62.3)		99 (63.1)				185 (62.9)
Low		11 (8.0)		19 (12.1)				30 (10.2)
<i>Anxiety measures</i>								
SCAS-C	41.97 (17.27)		42.69 (17.09)		−0.36	.174	42.35 (17.20)	
SCAS-P	31.45 (13.52)		31.59 (13.33)		−0.9	.928	31.52 (13.40)	
CALIS-C	11.50 (7.13)		12.69 (7.85)		−1.35	.178	12.13 (7.53)	
CALIS-P	14.19 (6.60)		14.68 (6.30)		−0.64	.522	14.45 (6.50)	

SCAS Spence Children's Anxiety Scale; CALIS Child Anxiety Life Interference Scale. Norwegian defined as one or both parents born in Norway. Parent reports consisted mainly of mothers' ratings (90.4%). Fathers' ratings were used when mothers' ratings were not available

clinics (social workers *n* = 3, psychiatric nurse *n* = 1, special education teachers *n* = 2). All delivered the interventions as part of their regular job. Among the facilitators, 83.9% had no prior CBT training, 75.0% had no prior experience working with anxiety, and they had an average of 6.70 years of experience working with youth (*SD* = 6.74, range 0–27; Haugland et al., 2020). Each facilitator delivered 1–8 groups (*M* = 3.25; *SD* = 1.80).

Interventions

Vaag and CK are manualized and include basic CBT components for anxiety (e.g., psychoeducation, cognitive restructuring, exposure). The components are introduced in the same order in both interventions. Skills are taught through goal setting, exposure plans, problem solving, and homework. Participants receive booklets with illustrative descriptions and fill-in-tasks to use in and between sessions. These materials differ between interventions when it comes to graphic design and terms used. The interventions also differ in the number of sessions, degree of structure, and the level of possible tailoring to each group member. Compared to the standard-length program, the treatment protocol for the brief program has more detailed instructions describing how much time should be used for each exercise and less room for adjustments to fit the individual youth. To reduce systematic bias, all facilitators received the same training,

both interventions were delivered at each school involved in the RCT, and the majority of facilitators (75.0%) administered both interventions. To control for treatment differentiation, we used a 3-item measure indicating whether materials from one program was applied in the other (e.g., self-help material in Vaag or realistic thinking schema in CK). There was no overlap in use of materials and concepts between the programs.

Vaag Vaag¹ (Raknes et al., 2015) is a 5-session CBT group intervention with sessions varying between 45 and 90 min (total 5.5 h). The first four sessions are delivered weekly, with the fifth session delivered five weeks later. Session two is a joint youth-parent session. Vaag was developed for the RCT (Haugland et al., 2017).

Cool Kids CK (Rapee et al., 2006) is a 10-session CBT intervention, with weekly sessions of 90 min (total 15 h, plus two 90-min parent-only sessions). The adolescent group-based school-version was applied (Rapee et al., 2006). CK has previously shown to demonstrate efficacy in the treatment of youth with anxiety disorders (Mychailyszyn, 2017), and as an indicated school-based prevention program (Mifsud & Rapee, 2005).

¹ Norwegian word meaning “dare” or “venture”.

Training, Supervision, and Treatment Fidelity Each group was delivered by two facilitators. Each facilitator received a four-day training workshop comprised of basic CBT principles for anxiety, assessment procedures, and both intervention manuals. Two additional two-day workshops were provided during the study to ensure cross-site consistency and prevent intervention drift. The facilitators received regular supervision by CBT experts ($N=10$). Eight supervisors were certified by the national CBT association by completing a CBT training program. Supervision was 3–4.5 h for each Vaag group and 6–10.5 h for each CK group. The difference in number of hours is due to the different number of sessions in the two interventions. Thus, the supervision was held constant relative to program duration. Supervision was delivered according to a plan entailing the duration, structure, and content of supervision, including feedback on video recordings of sessions (Haugland et al., 2020).

Anxiety Measures

Spence Children's Anxiety Scale

(SCAS-C/P; Spence, 1998) is a 38-item questionnaire assessing youth anxiety symptoms. Items are rated on a 4-point scale (0–3; *never* to *always*) yielding a maximum score of 114. The SCAS-C/P has demonstrated sound psychometric properties (Nauta et al., 2004; Spence, 1998), which have been replicated for youth in Scandinavian samples (Arendt et al., 2014; Olofsdotter et al., 2016). Studies using the Norwegian translation have been published (e.g., Fjermestad et al., 2020; Wergeland et al., 2014). Internal consistency in the total sample was good to excellent (SCAS-C, $\alpha=0.91$; SCAS-P Mothers, $\alpha=0.89$; SCAS-P Fathers, $\alpha=0.87$).

Child Anxiety Life Interference Scale

(CALIS-C/P; Lyneham et al., 2013) is a 9-item measure assessing impairment associated with anxiety in school, home, social life, and activities. Items are rated on a 5-point scale (0–4; *not at all* to *a great deal*) yielding a maximum score of 36. CALIS has demonstrated satisfactory psychometric properties (Lyneham et al., 2013), which have been replicated in Scandinavian samples (Johnsen et al., 2019; Kilburn et al., 2019). In addition to the current RCT, one other study using the Norwegian translation has been published (Raknes et al., 2017). Internal consistency for the CALIS in the total sample was adequate to good (Youth $\alpha=0.86$; Mothers $\alpha=0.79$, Fathers $\alpha=0.84$).

Adherence and Competence Measure

Competence and Adherence Scale for CBT

(CAS-CBT) for Anxiety Disorders in Youth (Bjaastad et al., 2016) is an 11-item observational instrument designed to assess adherence and competence in the delivery of manual-based CBT for youth anxiety. The CAS-CBT is based on the structure of the Cognitive Therapy Adherence and Competence Scale (Barber et al., 2003; Liese, Barber, & Beck, 1995), but adapted for CBT delivered to youth. The adherence items include (a) review of homework and presentation of new homework, (b) structure and progress, (c) parental involvement, (d) positive reinforcement, (e) facilitation of collaboration with the youths, (f) facilitation and completion of session goal 1 (e.g., facilitator helping the youths do exposure exercises), (g) facilitation and completion of session goal 2. Parent involvement was not scored as parents did not attend any session coded for the present study. This resulted in six adherence items. The competence items include (a) skill level in CBT structure, (b) flexibility in adjusting the intervention to the youths, (c) skill level on process and relational skills, and (d) skill level in administering the session goals. The two adherence items and the one competence item focusing on the two main goals of each session were operationalized by consulting the treatment manual. The Adherence subscale was calculated by adding the six item scores and averaging them. Likewise, the Competence subscale was calculated by adding the four item scores and averaging them.

Coders watch entire sessions and rate each item on a 7-point Likert-type extensiveness scale (Hogue et al., 1996). Adherence items are scored on a 0–6 scale reflecting the degree to which the facilitator delivers the intervention: 0 (*None*) to 6 (*Thorough*). Competence items are scored on a 0 to 6 scale reflecting the level of skills: 0 (*Poor skills*) to 6 (*Excellent skills*). Consistent with prior use of the CAS-CBT with group programs, both facilitators' behaviors were considered when producing scores on each item (see Bjaastad et al., 2018). For each session, an Adherence subscale score was produced by averaging the six adherence items, and a Competence subscale score was produced by averaging the four competence items. Bjaastad et al. (2016) and Harstad et al. (2021) found that the CAS-CBT total scale demonstrated good internal consistency ($\alpha=0.87$ and $\alpha=0.87$, respectively), and good to excellent interrater reliability (Adherence subscale ICC=0.83; Competence subscale ICC=0.64 and Adherence subscale ICC=0.87; Competence subscale ICC=0.63, respectively). Interrater reliability was calculated based on 23 sessions for the Adherence and Competence subscales using ICC (Shrout & Fleiss, 1979). The reliability coefficients represent the model ICC (2,1) based on a two-way random effects model. The

reliability coefficients showed good agreement for adherence (ICC = 0.63) and competence (ICC = 0.69; Cicchetti, 1994). In the present study, internal consistency for the CAS-CBT was excellent (Adherence $\alpha = 0.81$; Competence $\alpha = 0.87$; Total scale $\alpha = 0.91$).

CAS-CBT Coding and Session Sampling Procedures

Seven coders ($n = 6$ clinical psychologists, $n = 1$ child psychiatrist; 28.6% male; M age = 48.7 years, range = 31–66 years; all Norwegian) comprised the coding team. Each coder had clinical experience delivering CBT for anxiety and were trained in both interventions. Six of the coders were supervisors in the trial but did not rate their supervisees. Training of coders involved didactic instruction, discussions of the scoring manual, and reviews of sessions with developers of CAS-CBT. Each item was reviewed with discussion and examples from videotapes. Coders engaged in coding of three videos, and results were discussed to reach consensus ratings. Each coder then independently scored four sessions for certification. The study principal investigator was one of the CAS-CBT developers and scored master codes for the certification sessions. A coder was certified if s/he had no more than two out of 40 items (4 sessions X 10 items) deviate plus or minus 2 points from the master coders. All coders reached this criterion.

After coders were certified, recordings were randomly assigned to each coder while ensuring that no one rated their supervisee. All coders evaluated both Vaag and CK. During coding, the coding team remained in regular contact to prevent coder drift. All group sessions were videotaped (except out-of-office exposure tasks). Two sessions from each group were coded ($N = 104$ sessions), representing 20% of the sessions in CK (session 6 and 7), and 40% of the sessions in Vaag (session 3 and 4). These specific sessions were selected due to similar duration and content across interventions. Tapes of seven sessions of CK (12.5%) were missing. These were replaced by subsequent or preceding sessions. For interrater reliability, the expert coders double coded ≥ 2 randomly selected tapes evaluated by each of the other coders ($n = 23$; 23.9%).

Data Analyses

Preliminary Analyses

Analyses were performed in SPSS 25 and STATA (15.1). Sample bias analyses were conducted to determine whether the included 295 participants differed from the parent sample on demographic variables and outcome measures. Pre-intervention differences for youth participants on demographic and clinical variables were analyzed by t-tests (continuous

variables) and χ^2 tests (categorical variables). In addition, missing data patterns were examined.

Group Differences in Adherence and Competence Scores

To determine how to generate CAS-CBT Adherence and Competence Subscale scores, descriptive statistics for each subscale within the coded intervention sessions were evaluated along with the magnitude of the correlations for the subscales within and across the two coded sessions for each group. The correlations were interpreted following Rosenthal and Rosnow (1984) where r is “small” if 0.10–0.23, “medium” if 0.24–0.36, and “large” if > 0.36 . To analyze whether adherence (CAS-CBT Adherence subscale) or competence (CAS-CBT Competence subscale) differed between Vaag and CK, generalized linear models (GLM) with robust standard-length error estimates to account for clustering of facilitators were used. Mean CAS-CBT Adherence and Competence subscale scores with 95% confidence interval (CI) were calculated.

Treatment Fidelity Predicting Outcome

To account for multiple testing (four analyses, one per outcome), predictors were considered significant at a Bonferroni-corrected significance level of $\alpha = 0.0125$. For evaluation of whether treatment fidelity predicted youth clinical outcomes, linear mixed effect models (LMM) were conducted for adherence and competence. LMMs were used to analyze if adherence (CAS-CBT Adherence subscale) predicted change over time (pre-, post-, and follow-up) in each of the clinical outcomes (SCAS-C/P; CALIS-C/P), and to examine if the relation varied by CBT program. The LMMs included random intercepts to account for the data's hierarchical structure, with three levels of nesting: individual, intervention group, and school. Separate analyses were conducted for each anxiety measure. The model included time (pre-intervention, post-intervention, and follow-up), adherence, and an interaction term of adherence and time as fixed effects to examine if adherence was related to change in average level of anxiety over time. A model with the three-way interaction of intervention, adherence, and time was included to examine if the effect of adherence on outcome varied by intervention type. Likelihood-ratio test was applied to compare the different models and test if the interaction was significant. The same analyses were conducted for competence (CAS-CBT Competence subscale) as a predictor of outcome. All analyses were repeated with the treatment completers only, to investigate if outcomes in youth completing the interventions were impacted by treatment fidelity.

Effect sizes were generated by standardizing the adherence, competence and clinical outcome measures (i.e.,

$M=0$; $SD=1$; Ferron et al., 2008; Lorah, 2018). Clinical outcome measures were standardized within each time point. All analyses were repeated with the standardized variables. The resulting values are interpreted as the expected change in the number of standard deviations in the dependent variable, given a one standard deviation change in the independent variable (Lorah, 2018). For the current study, that means a negative effect size is interpreted as higher fidelity predicting better outcomes. The value of these effect sizes can be interpreted as smaller, more conservative correlation coefficients (r ; Ferguson, 2009), with an effect size of 0.2 being the recommended minimum representing a “practically” significant effect for social sciences, i.e., a small effect, ≥ 0.5 is a moderate effect and ≥ 0.8 is a strong effect (Ferguson, 2009).

Residual Analyses

Because the adherence and competence scores were highly correlated, the two treatment fidelity components were not included in the same model. However, the unique contribution of each variable was investigated by providing residual analyses. We generated a residual competence variable from a linear regression with adherence as independent variable and competence as dependent variable. The residual competence variable was used as an independent variable in addition to adherence in additional analyses with LMMs using the same strategy as when the two treatment fidelity components were analyzed separately.

Results

Preliminary Analyses

There were no differences between the current study sample ($n=295$) and the remaining participants in the Haugland et al. (2020) study ($n=18$). Further, there were no differences between Vaag ($n=142$) and Cool Kids ($n=160$) on youth demographic and baseline clinical characteristics (see Table 1). All facilitators received training in both interventions and 75% delivered both interventions. There were no missing values in the adherence and competence data. We evaluated patterns of missing data in youth- and parent-reported outcomes. Missing data ranged from 0.7 to 1.0% at pre-intervention, 12.9–14.6% post-intervention, and 31.5–35.9% at follow-up. Missing data occurred completely at random for all time points (Little’s missing-completely-at-random test: Pre-intervention $\chi^2=1.674$, $df=4$, $p=0.795$; Post-intervention $\chi^2=6.095$, $df=10$, $p=0.807$; Follow-up $\chi^2=14.707$, $df=10$, $p=0.143$). In the LMMs, we used full information maximum likelihood (FIML) missing data methodology (Wothke, 2000) in Stata, which is the

state-of-the-art method for handling missing data (Schafer & Graham, 2002).

Group Differences in Adherence and Competence Scores

The CAS-CBT Adherence subscale score for each group ranged from 3.17 to 5.75 ($M=4.41$, $SD=0.56$) and 2.75 to 5.88 ($M=4.18$, $SD=0.66$) for the Competence subscale. A medium correlation was found for adherence between the two coded sessions ($r=0.36$, $p=0.008$), whereas a large correlation was found for competence across the two coded sessions ($r=0.41$, $p=0.002$; Rosenthal & Rosnow, 1984). Within each session, correlations between adherence and competence scores were large (session 1 $r=0.83$, $p<0.001$; session 2 $r=0.86$, $p<0.001$; Rosenthal & Rosnow, 1984). Based on the pattern and magnitude of the correlations, we decided to create a single CAS-CBT Adherence and a single Competence subscale score for each group by averaging scores across the two coded sessions. The correlation between the CAS-CBT Adherence subscale score and CAS-CBT Competence subscale score was large ($r=0.86$, $p<0.001$).

Group comparisons indicated that the CAS-CBT Adherence scores were significantly higher in Vaag ($M=4.71$, 95% CI [4.51, 4.91]) than in CK ($M=4.15$, 95% CI [3.97, 4.34]; $p<0.001$). Also, CAS-CBT Competence scores were significantly higher in Vaag ($M=4.42$, 95% CI [4.17, 4.67]) than in CK ($M=3.97$, 95% CI [3.74, 4.20]; $p=0.010$).

Treatment Fidelity Predicting Outcome

We examined if treatment fidelity (CAS-CBT Adherence and Competence subscales) predicted change in anxiety symptoms or impairment (SCAS-C/P, CALIS-C/P; see Table 2). The CAS-CBT Adherence subscale and CAS-CBT Competence subscale did not significantly predict change in anxiety symptoms, all $p\geq 0.043$. All the effect sizes were lower than the recommended minimum for practical significance (Ferguson, 2009).

When the interaction between the CAS-CBT Adherence subscale and CBT program was examined, no significant interaction was found for any of the outcomes, all $p\geq 0.424$. Similarly, no interaction between the CAS-CBT Competence subscale and CBT program was found for any of the outcomes, all $p\geq 0.528$. Overall, these results indicate that treatment fidelity does not predict youth anxiety outcomes, nor does treatment fidelity predict youth anxiety outcomes differently across Vaag and CK.

The main model was also tested for whether adherence predicted outcome and whether competence predicted outcome for those defined as treatment completers. No significant predictions were found.

Table 2 Estimates of adherence and competence as predictors of change in anxiety symptoms and change in impairment from anxiety

	SCAS-C			SCAS-P			CALIS-C			CALIS-P		
	<i>b</i>	95% CI	<i>p</i>	Stand. β	95% CI	<i>p</i>	Stand. β	95% CI	<i>p</i>	Stand. β	95% CI	<i>p</i>
<i>Adherence</i>												
Intercept	49.11	[32.16–66.07]	<.001									
Post	-13.77	[-28.18–0.63]	.061	0.01	[-25.50–3.99]	0.007	0.03	[3.04–18.17]	.006	0.06	[5.82–18.72]	<.001
F/u	-12.45	[-28.58–3.68]	.130	-0.02	[-18.03–5.29]	.284	0.01	[-2.84–12.27]	.221	-0.02	[-6.42–6.70]	.967
Adh. x Post	-0.42	[-4.47–3.63]	.838	-0.01	[-1.27–4.43]	.287	0.07	[-2.30–1.34]	.606	-0.03	[-1.18–1.93]	.636
Adh. x F/u	-1.33	[-5.71–3.06]	.553	-0.42	[-4.18–1.92]	.467	-0.05	[-3.39–0.53]	.152	-0.10	[-2.42–0.90]	.372
<i>Competence</i>												
Intercept	45.97	[32.08–59.86]	<.001									
Post	-8.66	[-20.37–3.05]	.147	0.01	[-20.97–3.64]	.005	0.03	[3.85–16.13]	.001	0.01	[8.30–18.78]	<.001
F/u	-4.10	[-17.32–9.13]	.544	-0.02	[-14.85–4.15]	.269	0.01	[0.08–12.41]	.047	-0.03	[-6.28–4.43]	.735
Comp. x Post	-0.92	[-4.41–2.56]	.603	-0.03	[-1.38–3.45]	.402	0.06	[-1.95–1.14]	.605	-0.03	[-1.12–1.51]	.771
Comp. x F/u	-2.68	[-6.48–1.12]	.167	0.10	[-4.10–1.11]	.260	-0.08	[-3.42–0.05]	.043	-0.13	[-2.27–0.58]	.243

Linear mixed models were used. *Stand.* Standardized; *SCAS* Spence Children’s Anxiety Scale; *CALIS* Child Anxiety Life Interference Scale; F/u = follow-up. The standardized Betas are derived from analyses with standardized variables

Residual Analyses

As the CAS-CBT Adherence and Competence subscales were highly correlated ($r = 0.86$), LMMs with both adherence and a residual competence variable were conducted (see Table 3). No significant predictions by treatment fidelity were found for any of the youth anxiety outcomes, all $p \geq 0.055$.

Discussion

In this study, we evaluated treatment fidelity across brief and standard-length school-based CBT delivered by novice CBT providers, and the prediction of treatment fidelity on outcome. We found that both adherence and competence were significantly higher in brief CBT than in standard-length CBT. However, neither adherence nor competence predicted youth anxiety outcomes across measures, informants, or time points. There was no difference in prediction of outcome by treatment fidelity across brief and standard-length CBT.

Our findings suggest that it may have been easier for novice CBT providers in school health services to establish treatment fidelity in the brief program, than the standard-length program manual that was more complex and had more room for flexibility. The simple and structured design of the brief program might explain the difference in adherence and competence across the brief and standard-length CBT programs (Lyon & Koerner, 2016). Our findings are in line with a recent study, where school nurses delivered a brief anxiety intervention with adequate adherence, and the intervention was deemed feasible for the school setting (Ginsburg et al., 2019).

Having available school personnel, who may not have previous mental health experience, or not have previous knowledge of CBT delivering mental health interventions may help increase the availability of evidence-based interventions. This may be especially beneficial for youth with anxiety symptoms and anxiety disorders, who often do not receive help due to lack of available providers, and other barriers to treatment such as stigma (Reardon et al., 2018). Training novice CBT providers in delivering anxiety interventions will increase the number of available providers. Further, one could speculate that stigma may be a less prominent barrier to treatment when providers are not highly specialized mental health professionals (e.g., psychologist/psychiatrist) and the interventions are easily accessible in schools rather than in clinics. In light of a recent study indicating school health services may reach youth with elevated anxiety symptoms who otherwise do not receive help (Husabo et al., 2020), there are potential benefits of delivering indicated anxiety prevention in schools. As group

Table 3 Estimates of adherence and residualized competence as predictors of change in anxiety symptoms and change in impairment from anxiety

Variable	SCAS-C			SCAS-P			CALIS-C			CALIS-P		
	β	<i>p</i>	95% CI	β	<i>p</i>	95% CI	β	<i>p</i>	95% CI	β	<i>p</i>	95% CI
<i>Post</i>												
Adherence	−0.41	.841	[−4.43, 3.61]	1.61	.269	[−1.24, 4.46]	−0.49	.594	[−2.32, 1.33]	0.39	.629	[1.18, 1.95]
Com res	−2.39	.470	[−8.88, 4.09]	−0.33	.890	[−5.00, 4.34]	−0.16	.918	[−3.17, 2.85]	−0.40	.763	[−2.98, 2.19]
<i>Follow-up</i>												
Adherence	−1.52	.493	[−5.89, 2.84]	−1.18	.449	[−4.23, 1.87]	−1.57	.117	[−3.53, 0.39]	−0.78	.361	[−2.45, 0.89]
Com res	−5.71	.105	[−12.62, 1.19]	−2.24	.376	[−7.19, 2.72]	−2.12	.192	[5.31, 1.07]	−1.15	.412	[−3.88, 1.59]

Linear mixed models were used. *Com res* Residualized competence; *SCAS* Spence Children's Anxiety Scale; *CALIS* Child Anxiety Life Interference Scale

leaders achieved adherence and competence levels in both interventions observed in previous studies (Bjaastad et al., 2018; Harstad et al., 2021), this suggests that novice CBT providers can achieve treatment fidelity when training and supervision is provided. Our findings further indicate that such efforts may be even more successful when the design of the program matches available providers' training and experience (Lyon & Koerner, 2016).

Treatment fidelity did not predict any clinical outcomes at any time points. These findings are in line with a recent study on adherence and competence in a clinical study that also found no significant relationship between adherence or competence and youth outcomes (Southam-Gerow et al., 2021). Given the accumulation of findings suggesting that fidelity does not predict outcomes, it may be appropriate to consider other approaches to investigating this relation. As one example, researchers could investigate the role that other components of treatment fidelity play in promoting positive client outcomes (e.g., participant responsiveness). As another example, researchers could use statistical approaches that investigate how the fidelity-outcome association unfolds over treatment (e.g., using random intercept cross-lagged panel models; Selig & Little, 2012). These approaches may help further elucidate the relation between fidelity and outcome in mental health treatment.

Further, effect sizes in our study varied in direction and magnitude but were all below what can be interpreted as a small effect (Ferguson, 2009). Our effect size estimates are similar to results in a previous meta-analysis on treatment fidelity and clinical outcomes in adults, where non-significant effect-sizes were reported for adherence ($r=0.02$) and competence ($r=0.07$) (Webb et al., 2010). However, our effect sizes are lower than reported for adherence in a meta-analysis of mental health treatments for youth ($r=0.096$; Collyer et al., 2019). The same meta-analysis reported a non-significant effect size for competence ($r=0.026$; Collyer et al., 2019). This suggests that our findings are in line with the results for competence in the meta-analysis (Collyer

et al., 2019). The studies included in this meta-analysis differed regarding settings (e.g., school, community clinics and university clinics), with only one of the included studies conducted within schools. There is clearly a need for more studies with rigorous methodology on treatment fidelity and outcome from this setting.

The lack of associations between fidelity and outcomes in our study could have several explanations. One explanation could be that a threshold level may exist, where adherence and competence are needed to achieve positive outcomes but have no additional effect above a certain point (Collyer et al., 2019; Durlak & DuPre, 2008). As both programs were delivered with adequate treatment fidelity, adherence and competence may have been above such a threshold level. We did not evaluate the adherence and competence of specific intervention components (e.g., exposure or cognitive restructuring) separately. It might be that some, but not all, components must be delivered with adherence and competence in order to attain positive outcomes. Thus, as long as adequate adherence and competence are present for these core components, higher levels of adherence and competence may not lead to better outcomes (Durlak & DuPre, 2008). That said, in a previous school-based anxiety intervention study, the adherence to specific components did not predict outcome (Becker et al., 2012).

The seemingly contradictory finding of no association between treatment fidelity and outcomes in a study of interventions showing effect, underscore the limited knowledge in the field of what exact mechanisms account for change in clinical outcomes. It is possible that non-specific factors, such as the quality of the alliance could be investigated, but these factors have not consistently predicted effects in previous studies (Chiu et al., 2009; Liber et al., 2010; Southam-Gerow et al., 2021). Group cohesion, defined as a client's sense of belonging and bonding toward other members in group therapy, may also play an important role in predicting outcomes in group CBT (Lerner et al., 2013; see Luong et al., 2021), but this has not been investigated in group CBT

for youth anxiety. More research is needed to identify the factors that account for improvement within CBT for youth anxiety.

We found significant overlap between adherence and competence, which is consistent with previous studies with this instrument (Bjaastad et al., 2018; Harstad et al., 2021). However, other studies focused on CBT for youth anxiety have found less overlap (r s 0.55–0.65; McLeod et al., 2018). An open question in the field is the degree to which adherence and competence represent distinct treatment fidelity components. Previous research has found correlations between adherence and competence ranging from 0.30 (Carroll et al., 2000) to 0.96 (Barber et al., 2003). The correlations tend to be higher when the same coders score adherence and competence (e.g., Barber et al., 2003; Bjaastad et al., 2018) as opposed to when independent coders score the two components (e.g., McLeod et al., 2018). Our findings suggest that it may be difficult, even for expert coders, to distinguish between adherence and competence.

Strengths and Limitations

Strengths of our study include a large sample, documentation of effectiveness of the CBT programs (Haugland et al., 2020), robust methodology for assessing adherence and competence with the use of independent coders, and a treatment fidelity measure with evidence of score reliability (Bjaastad et al., 2016). To date, few studies use treatment fidelity measures with established psychometric properties as there are currently no consistent guidelines for implementation measures (Collyer et al., 2019). Further, the brief and standard-length CBT programs cover broadly the same CBT modules (e.g., cognitive restructuring, exposure), but the brief CBT program is shorter, more structured, and allows less flexibility. Novice CBT providers trained in both interventions delivered the programs with no overlap between them. The same training and supervision procedures were used for both interventions and both interventions were delivered in all schools.

The study also has limitations. Though the evidence supports the CAS-CBT score reliability of the Adherence and Competence subscales, no studies have provided support for the convergent validity of these subscales. The coding of adherence and competence by the same coders may have influenced the correlation between the two variables so that they have not been able to judge them separately (McLeod et al., 2016). For better understanding of treatment fidelity, separate measures for adherence and competence may be beneficial. Limited variability in the Adherence and Competence subscales may have influenced our ability to find an association between treatment fidelity and outcomes. We used average scores of two sessions to generate the CAS-CBT Adherence and Competence subscale scores. This

approach assumes that adherence and competence scores are stable over the course of treatment, which may not be the case (McLeod et al., 2018; Smith et al., 2017). It is possible that a measure of treatment fidelity more specifically targeted to the intervention manuals and/or the core components of the program would give a different result. The instrument we used may not tap sequencing and adaptation of program specific components such as cognitive restructuring and exposure in sufficient detail (see, e.g., Marques et al., 2019; Park et al., 2015). Further, including other aspects of treatment fidelity, e.g., participant responsiveness, could have added additional information. By obtaining information on the participants response to the way the providers administered the program content, we might have gained more understanding of the lack of association between adherence or fidelity and outcome. In addition, we could have tested whether participant responsiveness (a) was influenced by adherence and/or competence, and (b) influenced clinical outcomes.

Treatment fidelity of novice CBT providers may change over time with increasing experience. Distinguishing between treatment fidelity when delivering the first groups compared to later groups could be a way of testing this assumption. The youth and facilitators were primarily female and Norwegian, which could limit the generalizability of the findings.

Only two sessions from each group were coded, which may not be enough to provide an accurate estimate of treatment fidelity (Southam-Gerow et al., 2020). Moreover, this sampling plan resulted in a higher proportion of program content coded for the brief program due to the difference in the number of total sessions between the two programs (i.e., 40% of the brief program vs. 20% of the standard-length program). It is possible that this may have resulted in lower treatment fidelity ratings in the standard-length intervention.

Conclusion

This study contributes to the limited understanding of treatment fidelity and its relation to outcomes, specifically in school-based interventions. Both the brief and the standard-length intervention were delivered with adequate fidelity, by novice CBT providers. Higher levels of adherence and competence were found in brief compared to standard-length CBT, suggesting that it may be easier for novice CBT providers to establish fidelity in a program designed to be short and easy to deliver by non-mental health practitioners. Neither adherence nor competence predicted outcome in either of the two school-based interventions. As both fidelity and outcomes must be considered, further studies are needed to decide what interventions to implement in school settings.

Acknowledgements We are grateful to the school health services that participated in the project and the group leaders who administered the interventions and data collection. We want to express our deep gratitude to the adolescents and parents who took part in this study.

Funding Open access funding provided by NORCE Norwegian Research Centre AS. This work was supported by The Western Norway Health Authority (grant number 912026); the Norwegian Research Council (grant number 229020), the U.S.-Norway Fulbright Foundation and the Regional Centre for Child and Youth Mental Health and Child Welfare, NORCE Norwegian Research Centre. Funders were not involved in the research.

Declarations

Conflict of interest Dr. Raknes has received royalties and compensation from sales for the self-help material in the Vaag program (the Psychological First Aid kit). Dr. Ron Rapee is the author of Cool Kids program but receives no royalties. The remaining authors report no conflicts of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the Regional Committee for Medical and Health Research Ethics (2013/2331) and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Consent to Participate and for Publication Informed consent was obtained from all participants in the study. Specifically, informant consent and participant assent were obtained from all parents/caretakers and individual participants included in the study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Angold, A., Costello, E. J., Farmer, E. M. Z., Burns, B. J., & Erkanli, A. (1999). Impaired but undiagnosed. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38(2), 129–137. <https://doi.org/10.1097/00004583-199902000-00011>
- Arendt, K., Hougaard, E., & Thastum, M. (2014). Psychometric properties of the child and parent versions of Spence children's anxiety scale in a Danish community and clinical sample. *Journal of Anxiety Disorders*, 28(8), 947–956. <https://doi.org/10.1016/j.janxdis.2014.09.021>
- Balazs, J., Miklosi, M., Keresztesy, A., Hoven, C. W., Carli, V., Wasserman, C., Apter, A., Bobes, J., Brunner, R., Cosman, D., Cotter, P., Haring, C., Iosue, M., Kaess, M., Kahn, J. P., Keeley, H., Marusic, D., Postuvan, V., Resch, F., ... Wasserman, D. (2013). Adolescent subthreshold-depression and anxiety: psychopathology, functional impairment and increased suicide risk. *Journal of Child Psychology and Psychiatry*, 54(6), 670–677. <https://doi.org/10.1111/jcpp.12016>
- Barber, J. P., Liese, B. S., & Abrams, M. J. (2003). Development of the cognitive therapy adherence and competence scale. *Psychotherapy Research*, 13(2), 205–221. <https://doi.org/10.1093/ptr/kpg019>
- Becker, E. M., Becker, K. D., & Ginsburg, G. S. (2012). Modular cognitive behavioral therapy for youth with anxiety disorders: A closer look at the use of specific modules and their relation to treatment process and response. *School Mental Health*, 4(4), 243–253. <https://doi.org/10.1007/s12310-012-9080-2>
- Bennett-Levy, J., Richards, D., Farrand, P., Christensen, H., Griffiths, K., Kavanagh, D., Klein, B., Lau, M. A., Proudfoot, J., Ritterband, L., White, J., & Williams, C. (2010). *Oxford guide to low intensity CBT interventions*. Oxford University Press.
- Bjaastad, J. F., Haugland, B. S. M., Fjermestad, K. W., Torsheim, T., Havik, O. E., Heiervang, E. R., & Ost, L. G. (2016). Competence and adherence scale for cognitive behavioral therapy (CAS-CBT) for anxiety disorders in youth: Psychometric properties. *Psychological Assessment*, 28(8), 908–916. <https://doi.org/10.1037/pas0000230>
- Bjaastad, J. F., Wergeland, G. J. H., Haugland, B. S. M., Gjestad, R., Havik, O. E., Heiervang, E. R., & Ost, L. G. (2018). Do clinical experience, formal cognitive behavioural therapy training, adherence, and competence predict outcome in cognitive behavioural therapy for anxiety disorders in youth? *Clinical Psychology & Psychotherapy*. <https://doi.org/10.1002/cpp.2321>
- Caron, E. B., Drake, K. L., Stewart, C. E., Muggeo, M. A., & Ginsburg, G. S. (2020). Intervention adherence and self-efficacy as predictors of child outcomes in school nurse-delivered interventions for anxiety. *The Journal of School Nursing*. <https://doi.org/10.1177/1059840520925522>
- Carroll, K. M., Nich, C., Sifry, R. L., Nuro, K. F., Frankforter, T. L., Ball, S. A., Fenton, L., & Rounsaville, B. J. (2000). A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug and Alcohol Dependence*, 57(3), 225–238. [https://doi.org/10.1016/S0376-8716\(99\)00049-6](https://doi.org/10.1016/S0376-8716(99)00049-6)
- Chiu, A. W., McLeod, B. D., Har, K., & Wood, J. J. (2009). Child-therapist alliance and clinical outcomes in cognitive behavioral therapy for child anxiety disorders. *Journal of Child Psychology and Psychiatry*, 50(6), 751–758. <https://doi.org/10.1111/j.1469-7610.2008.01996.x>
- Chorpita, B. F., & Daleiden, E. L. (2014). Structuring the collaboration of science and service in pursuit of a shared vision. *Journal of Clinical Child and Adolescent Psychology*, 43(2), 323–338. <https://doi.org/10.1080/15374416.2013.828297>
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*. <https://doi.org/10.1037/1040-3590.6.4.284>
- Collyer, H., Eisler, I., & Woolgar, M. (2019). Systematic literature review and meta-analysis of the relationship between adherence, competence and outcome in psychotherapy for children and adolescents. *European Child and Adolescent Psychiatry*. <https://doi.org/10.1007/s00787-018-1265-2>
- Currie, C., Molcho, M., Boyce, W., Holstein, B., Torsheim, T., & Richter, M. (2008). Researching health inequalities in adolescents: the development of the Health Behaviour in School-Aged Children (HBSC) family affluence scale. *Social Science and Medicine*, 66(6), 1429–1436. <https://doi.org/10.1016/j.socscimed.2007.11.024>
- Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41(3–4), 327–350. <https://doi.org/10.1007/s10464-008-9165-0>

- Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40(5), 532–538. <https://doi.org/10.1037/a0015808>
- Ferron, J. M., Hiogarty, K. Y., Dedrick, R. F., Hess, M. R., Niles, J. D., & Kromrey, J. D. (2008). Reporting results from multilevel analyses. In A. A. O'Connell & D. B. McCoach (Eds.), *Multilevel modeling of educational data* (pp. 391–426). Information Age Publishing.
- Fjermestad, K. W., Wergeland, G. J., Rogde, A., Bjaastad, J. F., Heiervang, E., & Haugland, B. S. M. (2020). School-based targeted prevention compared to specialist mental health treatment for youth anxiety. *Child and Adolescent Mental Health*. <https://doi.org/10.1111/camh.12366>
- Garner, B. R., Godley, S. H., Dennis, M. L., Hunter, B. D., Bair, C. M., & Godley, M. D. (2012). Using pay for performance to improve treatment implementation for adolescent substance use disorders: Results from a cluster randomized trial. *Archives of Pediatrics and Adolescent Medicine*, 166(10), 938–944. <https://doi.org/10.1001/archpediatrics.2012.802>
- Ginsburg, G. S., Becker, K. D., Drazdowski, T. K., & Tein, J.-Y. (2012). Treating anxiety disorders in inner city schools: Results from a pilot randomized controlled trial comparing CBT and usual care. *Child and Youth Care Forum*, 41(1), 1–19. <https://doi.org/10.1007/s10566-011-9156-4>
- Ginsburg, G. S., Drake, K. L., Muggeo, M. A., Stewart, C. E., Pikulski, P. J., Zheng, D., & Harel, O. (2019). A pilot RCT of a school nurse delivered intervention to reduce student anxiety. *J Clin Child Adolesc Psychol*. <https://doi.org/10.1080/15374416.2019.1630833>
- Harstad, S., Bjaastad, J. F., Hjemdal, O., Compton, S., Waaktaar, T., & Aalberg, M. (2021). Competence and Adherence Scale for Cognitive Behavioural Therapy (CAS-CBT) for anxiety disorders in youth: reliability and factor structure. *Behav Cogn Psychother*. <https://doi.org/10.1017/S1352465821000217>
- Hartnett, D., Carr, A., & Sexton, T. (2016). The effectiveness of functional family therapy in reducing adolescent mental health risk and family adjustment difficulties in an Irish context. *Fam Process*, 55(2), 287–304. <https://doi.org/10.1111/famp.12195>
- Haugland, B. S. M., Haaland, A. T., Baste, V., Bjaastad, J. F., Hoffart, A., Rapee, R. M., Raknes, S., Himle, J. A., Husabo, E., & Wergeland, G. J. (2020). Effectiveness of brief and standard school-based cognitive-behavioral interventions for adolescents with anxiety: A randomized non-inferiority study. *J Am Acad Child Adolesc Psychiatry*. <https://doi.org/10.1016/j.jaac.2019.12.003>
- Haugland, B. S. M., Raknes, S., Haaland, A. T., Wergeland, G. J., Bjaastad, J. F., Baste, V., Himle, J., Rapee, R., & Hoffart, A. (2017). School-based cognitive behavioral interventions for anxious youth: Study protocol for a randomized controlled trial. *Trials*, 18(1), 100. <https://doi.org/10.1186/s13063-017-1831-9>
- Heywood, C., & Fergusson, D. (2016). A pilot study of functional family therapy in New Zealand. *New Zealand Journal of Psychology*, 45(3), 12–22.
- Hogue, A., Henderson, C. E., Dauber, S., Barajas, P. C., Fried, A., & Liddle, H. A. (2008). Treatment adherence, competence, and outcome in individual and family therapy for adolescent behavior problems. *Journal of Consulting and Clinical Psychology*, 76(4), 544–555. <https://doi.org/10.1037/0022-006x.76.4.544>
- Hogue, A., Liddle, H. A., & Rowe, C. (1996). Treatment adherence process research in family therapy: A rationale and some practical guidelines. *Psychotherapy Theory, Research, Practice, Training*, 33(2), 332–345. <https://doi.org/10.1037/0033-3204.33.2.332>
- Husabo, E., Haugland, B. S. M., McLeod, B. D., Ogden, T., Rapee, R. M., & Wergeland, G. J. (2020). Does school-based recruitment for anxiety interventions reach youth not otherwise identified? A comparison between a school-based sample and a clinical sample. *School Mental Health*. <https://doi.org/10.1007/s12310-019-09357-7>
- 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.
- Kakuma, R., Minas, H., van Ginneken, N., Dal Poz, M. R., Desiraju, K., Morris, J. E., Saxena, S., & Scheffler, R. M. (2011). Human resources for mental health care: Current situation and strategies for action. *Lancet*, 378(9803), 1654–1663. [https://doi.org/10.1016/S0140-6736\(11\)61093-3](https://doi.org/10.1016/S0140-6736(11)61093-3)
- Kilburn, T. R., Sorensen, M. J., Thastum, M., Rapee, R. M., Rask, C. U., Arendt, K. B., & Thomsen, P. H. (2019). Group-based cognitive behavioural therapy for anxiety disorder in children with autism spectrum disorder: A feasibility study. *Nordic Journal of Psychiatry*, 73(4–5), 273–280. <https://doi.org/10.1080/08039488.2019.1622153>
- Langley, A. K., Nadeem, E., Kataoka, S. H., Stein, B. D., & Jaycox, L. H. (2010). Evidence-based mental health programs in schools: Barriers and facilitators of successful implementation. *School Ment Health*, 2(3), 105–113. <https://doi.org/10.1007/s12310-010-9038-1>
- Lerner, M. D., McLeod, B. D., & Mikami, A. Y. (2013). Preliminary evaluation of an observational measure of group cohesion for group psychotherapy. *Journal of Clinical Psychology*, 69(3), 191–208. <https://doi.org/10.1002/jclp.21933>
- Liber, J. M., McLeod, B. D., Van Widenfelt, B. M., Goedhart, A. W., van der Leeden, A. J. M., Utens, E. M. W. J., & Treffers, P. D. A. (2010). Examining the relation between the therapeutic alliance, treatment adherence, and outcome of cognitive behavioral therapy for children with anxiety disorders. *Behavior Therapy*, 41(2), 172–186. <https://doi.org/10.1016/j.beth.2009.02.003>
- Lorah, J. (2018). Effect size measures for multilevel models: Definition, interpretation, and TIMSS example. *Large-scale Assessments in Education*. <https://doi.org/10.1186/s40536-018-0061-2>
- Luong, H. K., Drummond, S. P. A., & Norton, P. J. (2021). Can you see what I see? A comparison of client and observer perspectives of the alliance and group cohesion in CBT. *Cognitive Behaviour Therapy*. <https://doi.org/10.1080/16506073.2021.1898463>
- Lyneham, H. J., Sbulrati, 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>
- Lyon, A. R., & Bruns, E. J. (2019). From evidence to impact: Joining our best school mental health practices with our best implementation strategies. *School Mental Health: A Multidisciplinary Research and Practice Journal*. <https://doi.org/10.1007/s12310-018-09306-w>
- Lyon, A. R., & Koerner, K. (2016). User-centered design for psychosocial intervention development and implementation. *Clinical Psychology: Science and Practice*, 23(2), 180–200.
- Marques, L., Valentine, S. E., Kaysen, D., Mackintosh, M. A., Dixon De Silva, L. E., Ahles, E. M., Youn, S. J., Shtasel, D. L., Simon, N. M., & Wiltsey-Stirman, S. (2019). Provider fidelity and modifications to cognitive processing therapy in a diverse community health clinic: Associations with clinical change. *J Consult Clin Psychol*, 87(4), 357–369. <https://doi.org/10.1037/ccp0000384>
- McLeod, B. D., Southam-Gerow, M. A., Rodriguez, A., Quinoy, A. M., Arnold, C. C., Kendall, P. C., & Weisz, J. R. (2016). Development and initial psychometrics for a therapist competence instrument for CBT for youth anxiety. *J Clin Child Adolesc Psychol*. <https://doi.org/10.1080/15374416.2016.1253018>
- McLeod, B. D., Southam-Gerow, M. A., Rodriguez, A., Quinoy, A. M., Arnold, C. C., Kendall, P. C., & Weisz, J. R. (2018). Development and initial psychometrics for a therapist competence instrument for CBT for youth anxiety. *Journal of Clinical Child and*

- Adolescent Psychology*, 47(1), 47–60. <https://doi.org/10.1080/15374416.2016.1253018>
- Mellin, E. A., Anderson-Butcher, D., & Bronstein, L. (2011). Strengthening interprofessional team collaboration: Potential roles for school mental health professionals. *Advances in School Mental Health Promotion*, 4(2), 51–61.
- Mifsud, C., & Rapee, R. M. (2005). Early intervention for childhood anxiety in a school setting: Outcomes for an economically disadvantaged population. *Journal of the American Academy of Child and Adolescent Psychiatry*, 44(10), 996–1004. <https://doi.org/10.1097/01.chi.0000173294.13441.87>
- Muggeo, M. A., Stewart, C. E., Drake, K. L., & Ginsburg, G. S. (2017). A school nurse-delivered intervention for anxious children: An open trial. *School Mental Health*, 9(2), 157–171. <https://doi.org/10.1007/s12310-017-9211-x>
- 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-Psychologie Canadienne*, 58(2), 105–115. <https://doi.org/10.1037/cap000101>
- Nauta, M. H., Scholing, A., Rapee, R. M., Abbott, M., Spence, S. H., & Waters, A. (2004). A parent-report measure of children’s anxiety: Psychometric properties and comparison with child-report in a clinic and normal sample. *Behaviour Research and Therapy*, 42(7), 813–839. [https://doi.org/10.1016/s0005-7967\(03\)00200-6](https://doi.org/10.1016/s0005-7967(03)00200-6)
- Olofsdotter, S., Sonnby, K., Vadlin, S., Furmark, T., & Nilsson, K. W. (2016). Assessing adolescent anxiety in general psychiatric care: diagnostic accuracy of the swedish self-report and parent versions of the spence children’s anxiety scale. *Assessment*, 23(6), 744–757. <https://doi.org/10.1177/1073191115583858>
- Overbeek, M. M., de Schipper, J. C., Lamers-Winkelmann, F., & Schuengel, C. (2013). Effectiveness of specific factors in community-based intervention for child-witnesses of interparental violence: A randomized trial. *Child Abuse Negl*, 37(12), 1202–1214. <https://doi.org/10.1016/j.chiabu.2013.07.007>
- Park, A. L., Chorpita, B. F., Regan, J., Weisz, J. R., & Research Network on Youth Mental. (2015). Integrity of evidence-based practice: Are providers modifying practice content or practice sequencing? *Adm Policy Ment Health*, 42(2), 186–196. <https://doi.org/10.1007/s10488-014-0559-z>
- Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology*, 75(6), 829–841. <https://doi.org/10.1037/0022-006X.75.6.829>
- Pine, D. S. (2007). Research review: A neuroscience framework for pediatric anxiety disorders. *Journal of Child Psychology and Psychiatry*, 48(7), 631–648. <https://doi.org/10.1111/j.1469-7610.2007.01751.x>
- Podell, J. L., Kendall, P. C., Gosch, E. A., Compton, S. N., March, J. S., Albano, A. M., Rynn, M. A., Walkup, J. T., Sherrill, J. T., Ginsburg, G. S., Keeton, C. P., Birmaher, B., & Piacentini, J. C. (2013). Therapist factors and outcomes in CBT for anxiety in youth. *Professional Psychology-Research and Practice*, 44(2), 89–98. <https://doi.org/10.1037/a0031700>
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(2), 65–76.
- Raknes, S., Haaland, A. T., Haugland, B. S. M., Wergeland, G. J., Bjastad, J. F., Rogde, A. H., & Hoffart, A. (2015). *Et gruppetilbud for ungdom som vil tørre mer. [Groups for adolescents who would like to be braver]*. RKBV Vest.
- Raknes, S., Pallesen, S., Bjaastad, J. F., Wergeland, G. J., Hoffart, A., Dyregrov, K., Haland, A. T., & Haugland, B. S. M. (2017). Negative life events, social support, and self-efficacy in anxious adolescents. *Psychological Reports*, 120(4), 609–626. <https://doi.org/10.1177/0033294117699820>
- Rapee, R. M., Lyneham, H. J., Schniering, C. A., Wuthrich, V., Abbott, M. A., Hudson, J. L., & Wignall, A. (2006). *The cool kids child and adolescent anxiety program therapist manual (School Version)*. Macquarie University.
- Rapley, H. A., & Loades, M. E. (2018). A systematic review exploring therapist competence, adherence, and therapy outcomes in individual CBT for children and young people. *Psychother Res*. <https://doi.org/10.1080/10503307.2018.1464681>
- Rasmussen, L.-M.P., Patras, J., Neumer, S.-P., Adolfsen, F., Martinsen, K. D., Holen, S., Sund, A. M., & Martinussen, M. (2019). Facilitators and barriers to the implementation of EMOTION: An indicated intervention for young schoolchildren. *Scandinavian Journal of Educational Research*. <https://doi.org/10.1080/00313831.2019.1596976>
- Reardon, T., Harvey, K., Young, B., O’Brien, D., & Creswell, C. (2018). Barriers and facilitators to parents seeking and accessing professional support for anxiety disorders in children: Qualitative interview study. *European Child and Adolescent Psychiatry*. <https://doi.org/10.1007/s00787-018-1107-2>
- Rosenthal, R., & Rosnow, R. L. (1984). *Essentials of behavioral research: Methods and data analysis*. McGraw-Hill.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147–177. <https://doi.org/10.1037/1082-989X.7.2.147>
- Schoenwald, S. K., Garland, A. F., Chapman, J. E., Frazier, S. L., Sheidow, A. J., & Southam-Gerow, M. A. (2011). Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(1), 32–43.
- Selig, J. P., & Little, T. D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. In *Handbook of developmental research methods* (pp. 265–278). The Guilford Press.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations—uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428.
- Smith, M. M., McLeod, B. D., Southam-Gerow, M. A., Jensen-Doss, A., Kendall, P. C., & Weisz, J. R. (2017). Does the delivery of CBT for youth anxiety differ across research and practice settings? *Behavior Therapy*, 48(4), 501–516. <https://doi.org/10.1016/j.beth.2016.07.004>
- Southam-Gerow, M. A., Chapman, J. E., Martinez, R. G., McLeod, B. D., Hogue, A., Weisz, J. R., & Kendall, P. C. (2021). Are therapist adherence and competence related to clinical outcomes in cognitive-behavioral treatment for youth anxiety? *J Consult Clin Psychol*, 89(3), 188–199. <https://doi.org/10.1037/ccp0000538>
- Spence, S. H. (1998). A measure of anxiety symptoms among children. *Behaviour Research and Therapy*, 36(5), 545–566.
- Stoll, R. D., Pina, A. A., & Schleider, J. (2020). Brief, non-pharmacological, interventions for pediatric anxiety: Meta-analysis and evidence base status. *J Clin Child Adolesc Psychol*. <https://doi.org/10.1080/15374416.2020.1738237>
- Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *J Consult Clin Psychol*, 78(2), 200–211. <https://doi.org/10.1037/a0018912>
- Weist, M. D., Bruns, E. J., Whitaker, K., Wei, Y., Kutcher, S., Larsen, T., Holsen, I., Cooper, J. L., Geroski, A., & Short, K. H. (2017). School mental health promotion and intervention: Experiences from four nations. *School Psychology International*, 38(4), 343–362. <https://doi.org/10.1177/0143034317695379>
- Wergeland, G. J., Fjermestad, K. W., Marin, C. E., Haugland, B. S. M., Bjaastad, J. F., Oeding, K., Bjelland, I., Silverman, W. K.,

- Ost, L. G., Havik, O. E., & Heiervang, E. R. (2014). An effectiveness study of individual vs. group cognitive behavioral therapy for anxiety disorders in youth. *Behaviour Research and Therapy*, *57*, 1–12. <https://doi.org/10.1016/j.brat.2014.03.007>
- Werner-Seidler, A., Perry, Y., Calear, A. L., Newby, J. M., & Christensen, H. (2017). School-based depression and anxiety prevention programs for young people: A systematic review and meta-analysis. *Clinical Psychology Review*, *51*, 30–47. <https://doi.org/10.1016/j.cpr.2016.10.005>
- Wothke, W. (2000). Longitudinal and multigroup modeling with missing data. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data*. Lawrence Erlbaum Associates.
- Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.