



Acceptance and commitment therapy interventions in secondary schools and their impact on students' mental health and well-being: A systematic review

Laura Knight^{*}, Victoria Samuel

South Wales Doctoral Programme in Clinical Psychology, Cardiff University, 70 Park Place, Cardiff, CF103AT, UK

ARTICLE INFO

Keywords:

Acceptance and commitment therapy'
'Systematic review'
'School'
'Mental health'
'Wellbeing'

ABSTRACT

In order to meet the growing need for mental health provision for young people, more attention has turned to schools to provide evidence-based interventions. Acceptance and Commitment Therapy (ACT) has been demonstrated in recent reviews and meta-analyses to be effective with young people, however to date no systematic reviews have examined the use of ACT as a school-based intervention.

This systematic review aimed to evaluate the methodological quality and examine the effectiveness of all peer-reviewed literature on ACT interventions based in secondary schools.

The PsycInfo, Scopus and Web of Science databases were searched for studies published in any year reporting on the use of ACT interventions based in secondary schools aiming to prevent or reduce mental health difficulties or promote wellbeing. Both universal and targeted studies were eligible for inclusion. Nine studies met inclusion criteria, with a total of 1324 participants across studies (age range 13–21 years).

Outcomes measured across all studies were depression, anxiety, anger, psychological capital, stress, wellbeing, life satisfaction, psychological health, emotional problems and mental health symptoms. Six studies also used process measures to explore different constructs linked to psychological flexibility, the mechanism of change in ACT. There was significant variation in methodological quality across studies.

Despite methodological weaknesses across studies, there are some promising results to show support for the use of ACT as a school-based intervention. As existing studies were heterogeneous with regard to design and outcomes measured, this review was unable to draw firm conclusions regarding the efficacy of ACT or the moderating influence of program type, program format or program delivery. More highly powered studies comparing ACT to other active treatments are needed in order to explore these questions further.

By the age of 18 years, approximately 20% of young people worldwide will have experienced a mental health problem (Kessler et al., 2007; Kieling et al., 2011). Poor mental health can impact upon many areas of a young person's life including poor engagement with education, increased health risk behaviours as well as increased risk of self-harm and suicide (Collins & Dozois, 2008; Patel et al., 2007). In an analysis of National Health Surveys in the UK between 1995 and 2014, Pitchforth et al. (2019) found a consistent increase in long-standing mental health conditions in young people aged 4–24 years. Over this 19-year period, the prevalence of mental health conditions increased sixfold in England, more than doubled in Scotland and increased by more than half in Wales.

There is an increasing recognition of the importance of early life

experiences for lifetime mental health problems, which further highlights the necessity to address the mental health needs of young people. Research conducted in the United States has found that 50% of adults with mental health problems first experienced them prior to age 15, and 75% of life-time mental health problems appear by age 24 years (Kessler et al., 2005). Research shows that mental health prevention in young people is key, as this is a sensitive period during the lifespan where protective factors such as building resilience could have significant and long-lasting consequences (Black et al., 2017).

Despite the increasing mental health needs of young people, statistics show that in the UK, one in three Child and Adolescent Mental Health Services (CAMHS) referrals made by schools are not accepted, and one in six referrals not accepted overall (NSPCC, 2017). Restricted access to

^{*} Corresponding author.

E-mail address: laura.knight@wales.nhs.uk (L. Knight).

<https://doi.org/10.1016/j.jcbs.2022.06.006>

Received 15 November 2021; Received in revised form 22 June 2022; Accepted 24 June 2022

Available online 30 June 2022

2212-1447/© 2022 The Authors. Published by Elsevier Inc. on behalf of Association for Contextual Behavioral Science. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

specialist services has meant that increasing attention has turned to mental health promotion and prevention in schools, due to their broad scope and existing structures to support child development (Domitrovich et al., 2010; Masia-Warner et al., 2006).

Schools are a key environment to provide mental health programmes for young people outside of clinical settings as they are safe, cost-effective and flexible places in which a diverse range of interventions can be offered (Marks, 2012). Wolpert et al. (2011) highlight how schools can provide a more accessible route for young people to access support, particularly for those from socio-economically disadvantaged families. School-based support has been associated with reduced stigma and increased engagement, especially among ethnic minority adolescents (Stephan et al., 2007). Therefore, school-based intervention programs provide a promising opportunity for low-threshold care, with the potential to reach adolescents who may be unlikely to access support in clinical settings.

The need to engage schools in supporting the mental health of young people has been recognised in UK policies and guidance. In 2017, the UK Government published ‘Transforming children and young people’s mental health provision: a green paper’, which detailed proposals for expanding access to mental health provision for young people, with a specific focus on additional support through schools and colleges. In Wales, the ‘Curriculum for Wales Guidance’ outlines plans to build health and wellbeing into the core of the new curriculum by defining it as one of the six ‘Areas of Learning Experience’ for Welsh schools from 2022 onwards (Welsh Government, 2020).

School-based interventions for mental health and wellbeing can be broadly grouped into three types; universal, selective and indicative approaches (Neil & Christensen, 2009). Both selective and indicative interventions are often referred to in research as ‘targeted’ interventions. Universal interventions are offered to all students regardless of risk or symptom status and are often aimed at enhancing wellbeing, resilience and promoting positive mental health (Barrett & Turner, 2004). Research has demonstrated that school-based staff generally have a preference for universal interventions due to their broad application, as well as the reduced time and stigma associated with running interventions that do not require students to be screened for mental health symptoms. (Horowitz et al., 2007). Conversely, selective intervention programs target students deemed at risk of mental health problems, due to individual or environmental characteristics such as socio-economic background. Indicative approaches are aimed at students identified as having existing low-moderate symptoms of a mental health problem, commonly anxiety or depression.

There have been a number of systematic reviews conducted to date on the effectiveness of school-based interventions. Reviews exploring the effectiveness of universal interventions have generally found small effects of the interventions on outcomes including anxiety, depression and externalising problems, suggesting limited effectiveness (Caldwell et al., 2019; Dray et al., 2017; Mackenzie & Williams, 2018). The studies included in these reviews predominantly based their interventions on a Cognitive Behavioural Therapy (CBT) approach.

A number of recent systematic reviews have compared universal, selective and indicative interventions delivered in schools. Corrieri et al. (2014) compared universal and indicative interventions for both depression and anxiety and found that although both types of interventions showed similar levels of effectiveness across outcomes, only the indicative programs maintained their benefits at follow up. Werner-Seidler et al. (2017) similarly compared these two types of interventions on depression and anxiety outcomes and found that whilst the outcomes for anxiety were comparable, universal interventions produced smaller effects for depression than targeted programs. Across both reviews, small to moderate effects for the interventions were found. Feiss et al. (2019) examined the outcomes of both universal and targeted interventions for depression, anxiety and stress. It was found that universal and targeted interventions were both effective at significantly reducing anxiety, however universal interventions were more effective

in a higher dose of the intervention whereas targeted interventions were not affected by dose. Targeted interventions were more effective for both depression and stress than universal interventions, however only significant results for depression were found. In this study, none of the benefits were maintained at follow up for either intervention type.

Two reviews evaluated the moderating factor of intervention content on program effectiveness. Werner-Seidler et al. (2017) found that program content did not moderate program effectiveness, whereas Dray et al. (2017) found that CBT interventions were more effective than non-CBT based interventions including positive psychology, mindfulness and social and emotional learning. Of the studies reviewed in Dray et al. (2017), 54% used a CBT approach.

In summary, there are several existing reviews demonstrating small effects of school-based interventions on mental health symptomatology, with slightly higher levels of effectiveness reported for targeted compared to universal interventions. The approach predominantly used across studies to inform the interventions is CBT, and there appears to be a distinct lack of a comparable research base using other therapeutic approaches. Of note, the studies reviewed have primarily evaluated the impact of the interventions on decreasing symptoms of poor mental health as opposed to increasing wellbeing, a factor that may be more relevant for universal interventions that emphasise prevention.

Acceptance and Commitment Therapy (ACT), an approach demonstrated in recent meta-analyses to be effective in young people across a range of outcomes (Fang & Ding, 2020a; Swain et al., 2015), was not used in any of the studies included in the current reviews of school-based interventions. Gillard et al. (2018) have identified ACT as a coherent model that has the potential to support schools in promoting wellbeing in children due to its clear health benefits.

Acceptance and Commitment Therapy (ACT) is a third wave therapeutic approach that uses acceptance and mindfulness strategies, together with identification of values and commitment to values-based living (Forman & Herbert, 2009; Hayes et al., 2006). The primary goal of ACT is not to reduce mental health symptoms but to increase psychological flexibility (Hayes, 2004). Psychological flexibility is defined as “the ability to be in the present moment with full awareness and openness to our experience and to take action guided by our values” (Harris, 2019, p.12). ACT is a transdiagnostic approach, which can be used as a treatment for a range of both mental health and physical health conditions such as chronic pain (Swain et al., 2015). This suggests ACT may be particularly suitable for universal interventions as it does not depend upon a disorder-specific formulation model.

ACT is based on Relational Frame Theory (RFT), which emphasises the role of human language development and cognition, specifically our capacity for identifying and creating relational links between stimuli (Hayes, 2004; Hayes et al., 2006). Functional contextualism is the philosophical stance behind RFT, which highlights the importance of context, and the function of internal experiences such as thoughts, emotions and memories (Hofmann & Asmundson, 2008). ACT posits that it is not the content of internal experiences that causes distress but the context in which they take place (Hayes et al., 2006; Hayes et al., 2004). Emotional distress is perceived as resulting from the experiencing of painful or difficult thoughts and feelings as intolerable, and the use of avoidance or suppression of these experiences as a way to escape distress (Luoma et al., 2007). The process of avoidance and suppression has the paradoxical effect of increasing the salience of the distressing internal experiences, which subsequently reduces a person’s ability to live a valued and meaningful life. Therefore, the focus of intervention in ACT is a person’s relationship with their internal experiences, rather than altering the internal experiences themselves.

In ACT, psychological flexibility is targeted using six inter-relational core therapeutic processes that form a “hexaflex” model of psychological flexibility: acceptance of internal experiences; cognitive defusion (interpreting thoughts as thoughts, as opposed to facts); mindfulness (present moment awareness, without judgement); self-as-context (detaching from unhelpful narratives about the self); identification of

personal values; and committed action towards a valued life (Luoma et al., 2007).

There are several elements of ACT which suggest it may be particularly suitable for young people. ACT relies less on talking during active therapy and uses experiential exercises and metaphor to introduce and practise key ideas. The use of experiential exercises and metaphors to link abstract concepts to concrete examples is particularly encouraged when working with adolescents, as this helps to support the cognitive shift from concrete to abstract thinking that occurs during adolescence (Greco et al., 2005; Halliburton & Cooper, 2015). Additionally, the focus on identification of values and commitment to valued action helps the adolescent to apply new skills and learning to their wider context, as opposed to the primary focus of symptom-reduction often found in other approaches such as CBT (Hofmann & Asmundson, 2008). This focus on identifying values can be particularly important for adolescents who are exploring their sense of identity and striving for autonomy (Casey et al., 2008).

There are two existing systematic reviews that have looked specifically at the use of ACT with young people. A systematic review by Swain et al. (2015) examined the use of ACT with children and young people aged 8–18 years across both physical and mental health difficulties. The authors concluded that in young people, ACT is more effective than control conditions across several problem domains. A more recent meta-analysis by Fang and Ding (2020a), examined 14 randomised controlled trials (RCTs) on the efficacy of ACT in children and adolescents. From their findings, the authors concluded that ACT is more effective than treatment as usual and untreated comparison groups in treating anxiety and depression, but was not superior to CBT. It was also found that ACT led to increases in quality of life and wellbeing compared to the untreated group, however ACT did not outperform CBT or treatment as usual. It was concluded that more high-quality research with improved methodology is needed to understand the efficacy of ACT for young people.

1. Aims and review question

Existing systematic reviews focused on the use of mental health interventions in school settings have found positive yet small effects when using predominantly CBT-based approaches. Recent systematic reviews have presented good evidence for the effectiveness of ACT with young people, however to date no systematic reviews have examined the literature on the use of ACT in school settings. Therefore, this systematic review aims to examine ACT interventions in secondary schools and their impact on students' mental health and wellbeing.

This review focuses on secondary school age children, as research has suggested that there is little evidence for the effectiveness of ACT in children under 11 years (Swain et al., 2015).

Specific research questions are:

- 1) How has ACT been applied to secondary school mental health and wellbeing interventions within existing studies?
- 2) How effective are ACT interventions based in secondary schools on improving students' mental health and wellbeing?
- 3) How methodologically robust are these studies?

2. Method

2.1. Search and screening procedures

PsycInfo, Scopus and Web of Science databases were electronically searched for published literature. These databases were chosen to give access to articles published in journals related to psychology and health. A list of keywords and terms were developed to identify relevant literature. From the preliminary searches it was clear that broad search terms were needed in order to capture all relevant studies. The search terms included were as follows:

“acceptance and commitment therapy”

AND

“school*” OR “adolescen*” OR “student*” OR “child*” OR “young”

Titles and abstracts of studies from the initial database searches were screened according to pre-determined inclusion and exclusion criteria (see Table 1). The full texts were then retrieved and screened according to the same inclusion and exclusion criteria.

For each included study, manual searches of reference lists were conducted and citation searches undertaken to locate additional potential studies for inclusion. These additional studies were then subject to the same screening procedures as those identified through initial database searches.

2.2. Eligible studies

2022 records were identified through the initial database searches. This was reduced to 1379 following removal of duplicates. An additional record was identified through manual searching of reference lists and citation searches. Twenty of these records met eligibility criteria following an initial screen of abstracts and titles, which then reduced to eight records following screening of the full text articles. See Appendix A for a list of the excluded articles and the reasons for exclusion. One article that met eligibility criteria included two empirical studies (Livheim et al., (a and b) 2015), therefore a total of nine studies are included in a narrative synthesis. See Fig. 1 for an overview of the study selection process.

2.3. Data extraction, synthesis and quality assessment

A data extraction sheet was developed in order to retrieve the relevant information from each included study. The data extracted included setting, total number of participants, participant demographics, study design, comparison conditions, mental health and/or wellbeing outcome measures, ACT process measures, data points, intervention format and length, therapist training, statistical analysis and outcomes. Relevant outcomes were any statistically significant reductions in mental health related symptoms or improvements in wellbeing, and whether the effects of the intervention were maintained at follow up. Due to the heterogeneity of the studies that met the inclusion criteria, a narrative synthesis of results was the most appropriate method for this review.

The quality of each included study was assessed using the ‘Psychotherapy Outcome Study Methodology Rating Form’ (POMRF) (Öst, 2008) (Appendix B). This is a 22-item scale that allows for assessment of a range of methodological elements including sample characteristics, research design, randomisation, the psychometric properties of outcome measures, assessment of statistical power, statistical analysis methods and bias in reporting of results. This scale was chosen as it includes elements specific to studies which evaluate psychological interventions, such as therapist training, therapist competence and therapeutic modality adherence. The POMRF has been used in a number of published systematic reviews of the ACT literature for adults and children (Fang & Ding, 2020a; Graham et al., 2016; Kelson et al., 2019; Swain et al., 2013, 2015). Each item is rated on a 3-point scale from 0 (poor) to 2 (good). Overall POMRF scores range from 0 to 44, with higher overall scores indicative of greater methodological rigour. In terms of psychometric properties, the POMRF has been found to have good internal consistency (0.86) and interrater reliability within the range 0.50–1.00 with a mean of 0.75 (Öst, 2008).

For the purposes of this review, a number of amendments were made to ensure the POMRF was a relevant tool in assessing all included studies, and scores could be compared. Items two (severity and chronicity of the disorder) and four (reliability of the diagnosis in question) on the POMRF were excluded from the quality assessment as four of the

Table 1
Inclusion and exclusion criteria.

Criteria	Inclusion	Exclusion	Rationale
Participants	The age range will generally be 11–18 as this represents the majority of young people attending secondary school. However, young people aged up to age 21 will be included if still attending a secondary school setting.	Participants younger than 11 or older than 21.	Adolescence is a key time for intervention in order to prevent lifetime mental health difficulties (Kessler et al., 2005). Research has suggested that there is little evidence for the effectiveness of ACT in children under 11 years (Swain et al., 2015).
Setting	Studies based in secondary schools.	Study is in a home setting, university, college or primary school. Studies where students were recruited from secondary schools however the intervention was not school-based.	Secondary schools are considered a key environment to deliver mental health programmes for young people outside of healthcare settings, as they are safe, cost-effective and flexible places in which a diverse range of interventions can be offered (Marks, 2012).
Intervention	Studies where Acceptance and Commitment Therapy has been used as a school-based intervention. The intervention can be either targeted to specific groups of students or non-targeted (universal). The intervention can be delivered to groups or individuals. The intervention can be of any duration, including single-session interventions.	Any study which only uses specific parts of Acceptance and Commitment Therapy in the intervention (e.g., mindfulness) or uses it in combination with one or more other therapeutic approaches.	This review aims to examine the impact of Acceptance and Commitment Therapy interventions.
Type of study	Intervention studies of all design types, from randomised controlled trials (RCT) to case studies, will be included within this review.	Review papers or observational, correlational or qualitative studies.	To ensure access to all studies that examine the effectiveness of Acceptance and Commitment Therapy interventions.
Outcome variables	At least one of the primary outcome measures is related to mental health or wellbeing.	Primary outcome measures that are related to other areas e.g., academic achievement/performance.	There is a growing concern about the mental health and well-being of children and young people in the UK, with increasing demand for specialist services as well as increased hospital admissions (Pitchforth et al., 2019).
Date	No date of publication limits will be applied to these searches.	No date of publication limits will be applied to these searches.	To ensure access to all relevant articles.
Language	Searches will be limited to only those publications written in English.	Publications not written in English.	No access to a translator.
Type of publication	Empirical studies published in peer reviewed journals.	Conference papers, book chapters, discussion papers and <u>grey literature</u> .	To ensure quality of studies.

studies included in this review used a universal, non-targeted intervention, therefore participants did not have a mental health diagnosis. Additional items that referenced a mental health diagnosis were items three (representativeness of the sample) and five (specificity of outcome measures). For studies using a non-targeted intervention, representativeness of the sample was interpreted as whether the sample of study participants reflected the whole school population demographics, or whether this subset of participants shared a specific characteristic such as gender or level of academic ability. Additionally, specificity of outcome measures was interpreted as whether the outcome measures selected allowed for specific measurement of the outcome variables identified in the aims of the research. As a result of these amendments, the total POMRF scores ranged from 0 to 40. These amendments ensured that comparisons could be made across all studies included in the review.

The quality of each study was rated on the POMRF by the author and an independent second rater. Where there was discrepancy in the scores, both the author and second rater presented a rationale for their scoring in order to facilitate discussion and reach a consensus.

3. Results

Table 2 provides an overview of the nine included studies. The studies included a total of 1324 participants. Despite the searches having no limit on publication date, all studies were recent, with the earliest published in 2014 (Pahnke et al., 2014).

3.1. Assessment of methodological quality

The results of the assessment of methodological quality revealed a

high level of variability among included studies (see Table 3). Overall POMRF scores ranged from 13 to 21 out of a total of 40 points, with a mean score of 16.7 (SD = 2.69). As Öst (2008) did not include cut-off scores for the POMRF, the current review followed the protocol set out in Swain et al. (2013) and employed standard deviations (rounded to the nearest whole number) to enable the calculation of a POMRF rating to compare methodological quality across the studies. Studies more than one SD below the mean POMRF score were rated “Below average” (range 0–14), those within one SD of the mean “average” (15–19), and those more than one SD above the mean “Above average” (20+). As demonstrated in Table 3, there were two studies in the below average range, five in the average range and two studies rating above average.

The following sections highlight common methodological strengths and weaknesses occurring across the nine included studies.

3.2. Participant demographics and representativeness of sample

The included studies were located across six different countries, with three of the studies based in Australia (Burckhardt et al., 2017; Livheim et al., (a) 2015; Smith et al., 2020). The majority of the schools included in the studies were state schools, with the exception of one study based in a private school (Burckhardt et al., 2017) and one study based in a specialist school for students with Autistic Spectrum Disorder (Pahnke et al., 2014). Aside from age and gender, the demographics of the students were sparsely reported. Participant ethnicity was not reported in any of the studies. Gender representation was variable across the studies. Four of the studies reported a fairly even gender distribution of between 42 and 53% female participants (Burckhardt et al., 2017; Fang and Ding, 2020b; Puolakanaho et al., 2019; Van der Gucht et al., 2017), two studies had over 70% female participants (Livheim et al. (a and b),

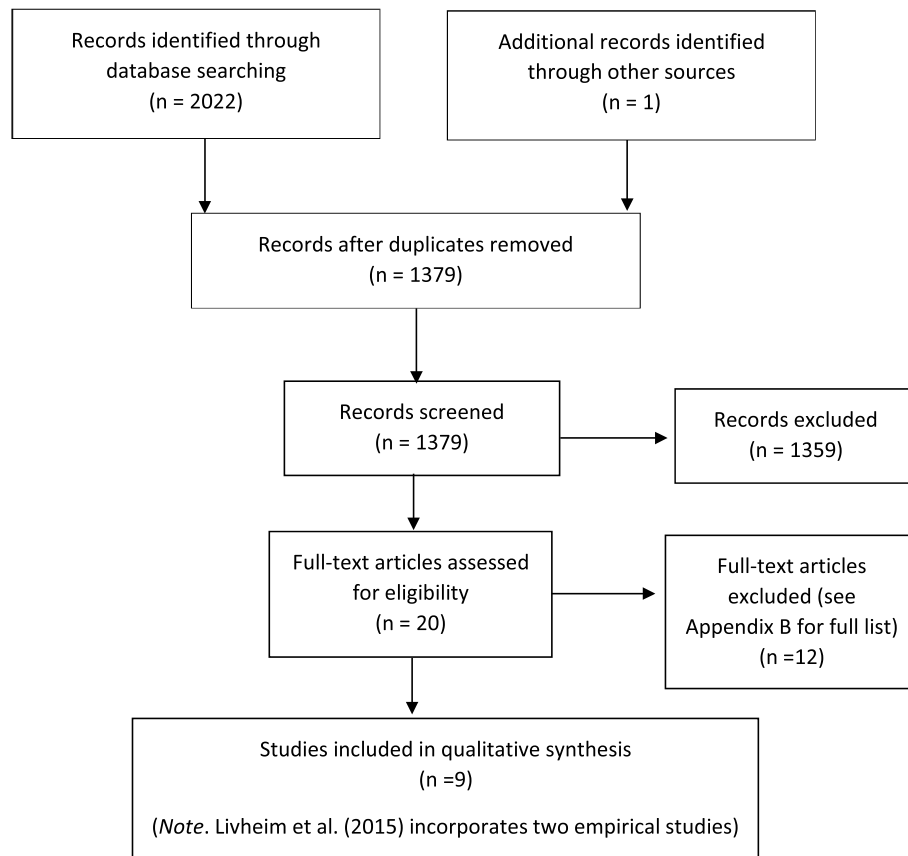


Fig. 1. Study selection process.

2015) and one study had only female participants (Smith et al., 2020). Pahnke et al., 2014 had majority male participants at 75%. One study did not report the gender distribution of their participants (Takahashi et al., 2020).

Only three studies made reference to the socioeconomic catchment area of schools included in the study (Burckhardt et al., 2017; Fang and Ding, 2020b; Livheim et al., (b) 2015). Burckhardt et al. (2017) stated that students in the school were “socio-economically advantaged compared with other students in the state of New South Wales and Australia, with 76% in the top quartile on a measure of socio-educational advantage” (p.3). Fang and Ding (2020b) stated that the school was in a ‘poverty-stricken area’ with a high percentage of ‘left-behind children’. In China, “left-behind children” refers to children under 18 years old who remain at home while one or both parents migrate to other places for work without living together for at least six months (Cheng & Sun, 2015). The authors reference prior research which states that left-behind children and adolescents are more likely to experience school maladaptation and mental health problems (Liu et al., 2007; Liu & Zhao, 2016). Livheim et al. (b, 2015) stated of their study setting that “the only notable differences from other regular public high school are that 100% of the students at this school qualified academically to study at upper secondary school at the age of 15 compared to 89%, which is the national average rate, and parent’s income level was twice as high as the Swedish family mean. (p.13)”

Studies which used a targeted intervention were generally found to have recruited a representative sample of students with that particular disorder and were not found to have used excessively strict exclusion criteria as indicated on the POMRF. Common exclusion criteria were high risk students expressing suicidality or psychotic symptoms, and students already receiving ongoing psychological treatment.

3.3. Study design and randomisation

Five studies described themselves as ‘pilot’ or ‘feasibility’ research (Burckhardt et al., 2017; Livheim et al. (a and b), 2015; Pahnke et al., 2014; Smith et al., 2020). Three of the studies were randomised controlled trials with two using cluster randomisation of school classes (Livheim et al. (b) 2015; Van der Gucht et al., 2017), and the other using individual randomisation (Puolakanaho et al., 2019). Two studies described using a ‘quasi-randomised design’, both of which involved cluster randomisation of school classes (Burckhardt et al., 2017; Pahnke et al., 2014). It was not clear in either of these studies which element of the cluster randomisation was considered ‘quasi’; the process of randomisation appeared to follow the same process as in other studies where cluster randomisation was also used and not referred to as ‘quasi’. Three of the studies were classified in their reports as a between-group design due to a lack of participant randomisation or insufficient randomisation stringency to be classified as an RCT (Fang and Ding, 2020b; Livheim et al., (a) 2015; Takahashi et al., 2020). Smith et al. (2020) used a within-group design, and therefore was the only study with no comparison group.

A control group was used in eight out of nine of the studies. Livheim et al., (a and b) (2015) used treatment as usual (TAU) comparison conditions with different treatment hours compared to the ACT intervention. The treatment as usual conditions consisted of ‘12 weeks of monitoring’ from the school counsellor (Livheim et al., (a) 2015) or ‘individual support’ from the school nurse (Livheim et al., (b) 2015). No further detail was provided regarding the type of monitoring and support that the students received, therefore it is not possible to ascertain the level of active treatment that was received or what approach was used. Six studies used a no treatment comparison group, where students attended their usual lessons. A treatment method that in previous research has been found effective for the disorder in question is the most

Table 2
Overview of included studies.

Study	N	Participant demographics (age and % female)	Country	Study design	Conditions	Mental health/ wellbeing outcome measure(s)	ACT process measure(s)	Treatment length	Targeted/non-targeted intervention	Individual/ Group intervention	Therapist
Burckhardt et al. (2017)	48	Age range: 14–16. 42% female	Australia	Feasibility study: Quasi-RCT (cluster randomisation)	ACT vs normal lessons	Flourishing Scale Depression Anxiety Stress Scale Short Form (DASS-21)	None	7 workshops and teacher-led exercises, once a week for 7 weeks. All workshops and teacher led classes were 25 min in duration.	Non-targeted	Group	Authors of paper who developed the ACT program, with support from teachers.
, Fang and Ding, (2020b)	35	Mean age: 13.23 46% female	China	Between-groups design	ACT vs normal lessons	Utrecht work Engagement Scale-student (UWES-s), Positive Psychological Capital Questionnaire (PPQ)	Acceptance and Action Questionnaire e II (AAQ- II)	Ten x 1-h workshops across five weeks.	Targeted to students at risk of mental health problems	Group	Graduate student in psychology, who was trained by the first author and had studied ACT.
Van der Gucht et al. (2017)	586	Age range: 14–21. 53% female	Belgium	Randomised controlled trial. Cluster randomisation.	ACT vs normal lessons	The Youth Self Report (YSR) World Health Organization Quality of Life questionnaire (WHOQoL-Bref)	Avoidance and Fusion Questionnaire for Youth (AFQ- Y)	Four x 120 min sessions delivered over four weeks.	Non-targeted	Group	Two teachers per school, who attended two days of ACT training by experienced ACT trainers.
Livheim et al. (2015) (a)	66	Age range: 12.5 and 17.75 years. 88% female	Australia	Pilot study: Between-groups	ACT vs. treatment as usual (12 weeks of monitoring from school counsellor)	The Reynolds Adolescent Depression Scale-2 (RADSD-2) Strengths and Difficulties Questionnaire (SDQ)	Avoidance and Fusion Questionnaire for Youth (AFQ- Y8)	Eight-week group, duration unspecified.	Targeted to students experiencing mild to moderate depressive symptoms	Group	Registered psychologists and co-facilitated by clinical psychology graduate students or the school's own counsellor.
Livheim et al. (2015) (b)	32	Age range: 14–15 years. 72% females	Sweden	Pilot study: Randomised Controlled Trial	ACT vs. treatment as usual (individual support from school nurse)	The General Health Questionnaire (GHQ- 12) Strengths and Difficulties Questionnaire (SDQ) Perceived Stress Scale Depression, Anxiety and Stress Scale (DASS- 21) Satisfaction with Life Scale	Avoidance and Fusion Questionnaire for Youth (AFQ- Y17) Mindful Attention Awareness Scale (MAAS)	Six x 90 min sessions, delivered over six weeks.	Targeted to students experiencing mild to moderate depressive symptoms	Group	Two clinical psychology major students, with clinical training in CBT and four days ACT-training.
Pahnke et al. (2014)	28	Age range: 13–21 years 75% male	Sweden, (specialist school for ASD)	Pilot study, Quasi- RCT (cluster randomisation).	ACT vs. normal lessons	Stress Survey Schedule (teacher- and self- ratings) Strengths and Difficulties Questionnaires (SDQ; teacher- and self- ratings), Beck Youth Inventories (BYI)-	None	Six-week group comprising two x 40-min group sessions per week and 6- to 12-min of daily mindfulness exercises.	Targeted to students with high- functioning ASD	Group	Group sessions delivered by a graduate psychology student; daily exercises delivered by classroom teacher.

(continued on next page)

Table 2 (continued)

Study	N	Participant demographics (age and % female)	Country	Study design	Conditions	Mental health/wellbeing outcome measure(s)	ACT process measure(s)	Treatment length	Targeted/non-targeted intervention	Individual/Group intervention	Therapist
Puolakanaho et al. (2019)	249	Mean age: 15.27 years 49% females	Finland	Randomised Controlled Trial	ACT vs. normal lessons	anxiety, depression and anger subscales (self-ratings). Overall stress scale developed for study School stress scale adapted from the Health Behaviour in School-Aged Children (HBSC) study	None	Five-week online program, duration unspecified.	Non targeted	Individual	Coaching and support from 31 undergraduate psychology students, supervision by a licensed psychologist
Smith et al. (2020)	10	Age range: 13–15 years. 100% female	Australia	Pilot study, within-group design	N/A- within group design	Beck Youth Inventory (BYI-II)- anxiety and depression sub-scales	Avoidance and Fusion Questionnaire for Youth (AFQ-Y 17)	Six x 1-h sessions, delivered over six weeks.	Targeted to students with mild-moderate depression and/or anxiety	Group	Two psychologists
Takahashi et al. (2020)	270	Age range: 14–15 years. Gender distribution not reported.	Japan	Between-groups design	ACT vs. normal lessons	Strengths and Difficulties Questionnaire (SDQ)	Value of Young Age scale – VOYAGE (Ishizu, Otsuki, Shimoda, & Nagata, 2016)	Six x bi-weekly group sessions, each lasting 50 min.	Non-targeted	Group	Clinical psychologist

Table 3

POMRF score and rating.

Study	POMRF score	POMRF rating
Takahashi et al. (2020)	13	Below average
Smith et al. (2020)	14	Below average
Livheim et al. (2015) (b)	15	Average
Pahnke et al. (2014)	15	Average
Fang and Ding, (2020b)	17	Average
Burckhardt et al. (2017)	17	Average
Puolakanaho et al. (2019)	18	Average
Livheim et al. (2015) (a)	20	Above average
Van der Gucht et al. (2017)	21	Above average

stringent comparison condition to use and therefore this criterion was not fulfilled by any of the studies included in this review (Öst, 2008)

3.4. Outcome measures: specificity, reliability and validity

The primary outcome measures selected varied significantly across all studies. A variety of disorder-specific and global measures of mental health and wellbeing were used, depending on the aims of the study. All measures used were self-report, with the exceptions of the Stress Survey Schedule (Grodén et al., 2001) and the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 2001) used in Pahnke et al. (2014) which contained a teacher rating as well as a student self-report rating.

The reliability and validity of outcome measures was variable across studies, with several measures selected for use that had not been validated within a youth sample. The Depression Anxiety and Stress Scale (DASS-21) (Antony et al., 1998) was used in two studies (Burckhardt et al., 2017; Livheim et al. (b) 2015), despite concerns in the literature regarding the appropriateness of this scale for an adolescent population. Two studies have found the three-factor structure of the DASS-21 to be invalid when used with an adolescent population (Szabó, 2010; Moore et al., 2017). These authors have also noted that emotional differentiation is still developing in younger respondents and they may not be able to fully appreciate the differentiation between depression, anxiety, and stress as reflected in the DASS-21 items. Furthermore, Szabó et al. (2010) contended that the DASS contained several expressions and words that might not be familiar to adolescents.

An ‘overall stress measure’ was used in Puolakanaho et al. (2019), which had only been validated in an adult population (Elo et al., 2003). There is a distinct lack of detail regarding the structure of this measure, with no reference made to number of items or whether any subscales were included.

Three other scales used, the General Health Questionnaire (GHQ-12) (Gao et al., 2004), the Satisfaction with Life Scale (Diener et al., 1985), and the Flourishing Scale (Diener et al., 2010) were all originally developed for use with an adult population, however have since been validated in adolescent samples (Duan & Xie, 2019; Neto, 1993; Tait et al., 2003).

ACT process measures were used in six of the studies, with three exceptions (Burckhardt et al., 2017; Pahnke et al., 2014; Puolakanaho et al., 2019). There was higher consistency amongst the ACT process measures used than with the primary outcome measures, with four studies opting for the Avoidance and Fusion Questionnaire for Youth (AFQ-Y) (Greco et al., 2008). The AFQ-Y has good internal consistency (Cronbach’s alpha = 0.90) and has good convergent validity against established measures of psychological distress, as well as ACT-specific measures (Greco et al., 2008).

The other ACT process measures used were a Chinese version of the Acceptance and Action Questionnaire II (AAQ-II) (Cao et al., 2013), the MAAS: Mindful Attention Awareness Scale (Carlson & Brown, 2005), and the Value of Young Age scale – VOYAGE (Ishizu et al., 2016). The AAQ-II has been validated within a Chinese adolescent sample (Cao et al., 2013), and the VOYAGE within a Japanese adolescent sample (Ishizu et al., 2016). It is unclear why the MAAS was used in Livheim

et al. (b) (2015), rather than the adolescent version (MAAS-A) published by Brown et al. (2011) which has been validated in youth samples. The studies which included an ACT process measure only measured either one or two ACT process variables. Therefore, no studies achieved a comprehensive assessment of psychological flexibility.

A factor which impacted the validity and reliability of both primary and secondary outcome measures used in studies from non-English speaking countries was the limited availability of outcome measures validated within a youth sample in the local language. Van der Gucht et al. (2017) used a validated Dutch version of the Youth Self Report (Verhulst et al., 1997), however Livheim et al. (b) (2015) translated and back-translated existing outcome measures without conducting a subsequent validity or reliability analysis. COSMIN guidance (Mokkink et al., 2019) recommends that cognitive interviewing should be used post-translation to check comprehensibility of items. Three studies in non-English speaking samples did not make reference to whether outcome measures were translated for their study (Pahnke et al., 2014; Puolakanaho et al., 2019; Takahashi et al., 2020), therefore the validity and reliability of these measures is difficult to ascertain.

3.5. Intervention delivery: format, therapist competence and adherence

Eight of the included studies delivered the ACT intervention in a group or lecture format (range of group participants was 9–60), and one study used an online program accessed by individuals who received supplementary weekly online coaching (Puolakanaho et al., 2019).

Burckhardt et al. (2017) delivered the intervention to the largest group of 60 students via lecture-style presentations for the psycho-educational elements, however experiential exercises were also used in smaller groups in between lectures.

All studies obtained a high score on the POMRF for ‘manualised, replicable, specific treatment programs’. A number of the interventions used across studies were based on existing ACT programs or manuals. In Fang and Ding (2020b), a translated version of the “ACT Made Simple” manual (Harris, 2019) was used to create a workshop based on the six modules of ACT. Van der Gucht et al. (2017) adapted a universal ACT prevention program (De Groot, 2005; Livheim, 2004) to the Flemish school context. All sessions included a psycho-educational part focused on theory and background, as well as experiential exercises and homework assignments. In both Livheim et al. (2015) studies, the intervention used was the ACT Experiential Adolescent Group, which is a manualized 8-week group program (Hayes & Rowse, 2008). This program uses experiential mediums, for example painting and role-play, to facilitate adolescents’ experience of the six ACT processes. This program was translated to Swedish and tested on a non-clinical group ahead of the main study in Livheim et al. (b) (2015). In Pahnke et al. (2014), an ACT protocol (Hayes et al., 2003) was adapted to meet the needs of the target population of young people with an Autistic Spectrum Condition. Skills training based on the six components of ACT was provided with the aim of developing “participants’ ability to cope with daily hassles and stressful situations, to break behavioural avoidance patterns, and to develop a broader behavioural repertoire” (p.4, Pahnke et al., 2014). Three studies (Puolakanaho et al., 2019; Smith et al., 2020; Takahashi et al., 2020) based their intervention on the book ‘Get Out of Your Mind and Into Your Life For Teens’ which introduces the six components of ACT via a new format created for an adolescent population named BOLD Warrior skills (Breathing deeply and slowing down, Observing, Listening to your values and Deciding on actions and doing them) (Bailey et al., 2012).

The majority of interventions covered all six core components of ACT, with two exceptions. Burckhardt et al. (2017) chose to exclude ‘self-as context’ from the intervention, as the developer found this component to be a difficult and confusing concept to transmit to adolescents. Takahashi et al. (2020) only made reference to four components of ACT in their paper (values, defusion, acceptance and committed action), however present moment exercises such as mindful breathing

are included in the description of each session.

The duration of the interventions delivered across studies was between four and ten weeks with session length varying from 25 min to 120 min. The study which reported the lowest number of total intervention hours was Smith et al. (2020) where the intervention had a total duration of 6 h. The highest number of total intervention hours was 10 h (Fang & Ding, 2020a, 2020b). Duration of the intervention was unspecified in two studies (Livheim et al., (a) 2015; Puolakanaho et al., 2019).

With regard to therapist training and competence, there was significant variability across the studies. Four studies used students as the primary therapist (Fang and Ding, 2020b; Livheim et al. (b), 2015; Puolakanaho et al., 2019; Pahnke et al., 2014), one study trained teachers to deliver the intervention (Van der Gucht et al., 2017), two studies used psychologists (Takahashi et al., 2020; Livheim et al., (a) 2015) and one study used specialists in ACT (Burckhardt et al. (2017)). In Burckhardt et al. (2017) the therapist delivering the intervention was the main author of the paper. In those studies where the primary therapists had less experience in ACT, unspecified or variable levels of supervision were offered, ranging from none to session-by-session supervision (Puolakanaho et al., 2019).

Checks for adherence to the treatment protocol and therapist competence were sparse across the literature, with only four studies making any attempt to monitor adherence to treatment protocol through use of supervision sessions (Fang and Ding, 2020b; Van der Gucht et al., 2017; Livheim et al., 2015 (a); Pahnke et al., 2014). None of the studies used a tool to monitor fidelity to the treatment protocol directly during intervention sessions.

3.6. Power, data points and statistical reporting

Power calculations were reported and followed in only one of the studies (Puolakanaho et al., 2019). All studies collected outcome data pre- and post-intervention, however only three studies included a follow up data point (Burckhardt et al., 2017; Pahnke et al., 2014; Van der Gucht et al., 2017). Of these three studies, only Van der Gucht et al. (2017) included a follow up at one year, which is the minimum criteria necessary to obtain the full score on the POMRF for this item.

Across the nine studies, statistical analysis methods were generally well-matched to the research design and the results comprehensively reported. Six of the studies (Burckhardt et al., 2017; Van der Gucht et al., 2017; Takahashi et al., 2020; Livheim et al. (a and b) 2015; Puolakanaho et al., 2019), used a linear mixed modelling approach to analysis which is recommended when there is longitudinal data or clustered data of students within classes/schools (Verbeke & Molenberghs, 2000). A score of zero was allocated to Pahnke et al. (2014) on the POMRF as repeated measures analysis of variance (ANOVA) was used to analyse the main effects and interaction effects despite clustered data. Repeated measures ANOVA does not take into account the lack of independence often seen in clustered data, as this approach assumes spherical errors. The use of statistical methods with underlying assumptions that do not reflect the data can have significant consequences for the accuracy and replicability of scientific results (Oleson et al., 2019).

3.7. Outcomes

There is an inconclusive picture of the effectiveness of ACT-based interventions from studies in the current review, due to the variability in outcomes and low methodological quality of many studies. Outcome data is reported in Tables 4 and 5, which includes reporting of statistical significance and effect sizes where presented in the studies at post-treatment and follow-up. All significance values reported are for a time by condition interaction, with the exception of Smith et al. (2020) where a within-subjects design was used.

Significance values and effect sizes are reported separately for mental health and wellbeing outcomes (Table 4) and process measure

outcomes (Table 5). Values are reported only for outcomes related to mental health, wellbeing and quality of life, which for some studies means only specific subscales of more general measures are reported, for example the ‘emotional symptoms’ subscale of the Strengths and Difficulties Questionnaire (SDQ). Where possible, effect sizes are reported even for non-significant findings to investigate whether Type 2 errors were made due to studies being underpowered. The majority of studies reported Cohen’s *d* effect sizes, which have been interpreted in this review according to Cohen’s criteria (1988) of 0.2 as small, 0.5 as medium and over 0.8 as large. In Pahnke et al. (2014), effect sizes were expressed as partial eta squared and were interpreted using the guidelines proposed by Cohen (1988) of 0.01 as small, 0.06 as medium and 0.14 as large.

Table 4
Mental health and wellbeing outcomes.

Study name and conditions	Outcome variable	Outcome measure	P	Effect sizes (ES)
Burckhardt et al., 2017. ACT vs. normal lessons.	Anxiety	Depression, Anxiety and Stress Scale (DASS-21)	.19	BG pre-post = small ES ($d = .28$), BG pre-follow up = medium ES ($d = .55$)
	Depression	Depression, Anxiety and Stress Scale (DASS-21)	.79	Both BG pre-post ($d = .31$) ES and BG pre-follow up ($d = .34$) ES = small to medium
	Stress	Depression, Anxiety and Stress Scale (DASS-21)	.14	Both BG pre-post ES ($d = .63$) and BG pre-follow-up ES ($d = .75$) = medium to large
	Wellbeing	Flourishing Scale	.57	BG pre-post = small ES, BG pre-follow up = small to medium ES
Fang and Ding, (2020b) ACT vs. normal lessons.	Psychological capital	Positive Psychological Capital Questionnaire (PPQ)	.006*	BG pre-post ES = large (1.4)
Van der Gucht et al. (2017). ACT vs. normal lessons.	Anxiety	Youth Self Report	>0.5	ACT pre-post = small to medium ES ($d = -0.33$), ACT pre-follow up = small ES ($d = -0.26$)
	Depression	Youth Self Report	>0.5	Control pre-post = small ES ($d = -0.28$), control pre-follow up = small ES ($d = -0.16$)
	Psychological health	World Health Organization Quality of Life questionnaire (WHOQoL-Bref)- psychological health subscale	>0.5	ACT pre-post = small to medium ES ($d = -0.34$), ACT pre-follow up = small ES ($d = -0.25$)
				Control pre-post = small ES ($d = -0.27$), control pre-follow up = small ES ($d = -0.21$)
Livheim et al. (a) (2015). ACT vs. TAU.	Depression	Reynolds Adolescent Depression Scale (RAD5-2)	0.008*	ACT pre-post = small ES ($d = 0.16$), ACT pre-follow up = small-medium ES ($d = 0.35$)
Livheim et al. (b) (2015). ACT vs. TAU.	Stress	Perceived Stress Scale	0.009*	Control pre-post = small ES ($d = 0.08$), control pre-follow up = small ES ($d = 0.09$)
	Life satisfaction	Satisfaction with Life Scale	0.126	BG pre-post = large ES ($d = 1.20$)
	Anxiety	Depression, Anxiety and Stress Scale (DASS-21)	0.057	ES not reported
	Depression	Depression, Anxiety and Stress Scale (DASS-21)	0.742	BG pre-post = large ES ($d = 0.8$)
	Stress	Depression, Anxiety and Stress Scale (DASS-21)	0.971	ES not reported
	Mental health symptoms	General health Questionnaire (GHQ-12)	0.130	ES not reported
	Stress	Stress Survey Schedule (student ratings)	.044*	ES not reported
Pahnke et al. (2014). ACT vs. normal lessons.	Stress	Stress Survey Schedule (teacher ratings)	.045*	BG overall ES = large ($\eta^2 = .11$)
	Emotional problems	Strengths and Difficulties Questionnaire (SDQ)- Emotional Symptoms subscale	.13	BG overall ES = large ($\eta^2 = .12$)
	Anxiety	Beck Youth Inventory (BYI)	.286	BG overall ES = moderate-large ($\eta^2 = .08$)
	Depression	Beck Youth Inventory (BYI)	.083	Not reported
	Anger	Beck Youth Inventory (BYI)	.023*	BG overall ES = large ($\eta^2 = .1$)
	BYI total	Beck Youth Inventory (BYI)	.032*	BG overall ES = large ($\eta^2 = .14$)
	Stress	Stress measure designed for study	0.037*	BG overall ES = large ($\eta^2 = .12$)
Puolakanaho et al. (2019). ACT vs. normal lessons			(Per protocol analysis)	BG pre-post ES = small ($d = 0.22$)
	Stress	Stress measure designed for study	>.05 (Intention-to-treat analysis)	ES not reported
Smith et al. (2020). No control group.	Anxiety	Beck Youth Inventory (BYI)	.007**	WG pre-post ES = medium to large ($d = 0.74$)
	Depression	Beck Youth Inventory (BYI)	.277	WG pre-post ES = small to medium ($d = 0.31$)
Takahashi et al. (2020). ACT vs. normal lessons.	Emotional problems	Strengths and Difficulties Questionnaire (SDQ)- Emotional Symptoms subscale	1.00	ES not reported

Note: *statistically significant group by time interaction ($p < .05$), **statistically significant effect of time (within-group design), BG = between-groups, WG = within-group, ES = effect size, TAU = treatment as usual.

3.7.1. Depression

Depression was measured in six of the studies (Burckhardt et al., 2017; Van der Gucht et al., 2017; Livheim et al. (a and b) 2015; Smith et al., 2020; Pahnke et al., 2014). The only statistically significant score was in Livheim et al. (a) (2015), which reported a large effect size. This study also scored ‘above average’ with regard to methodological rigour in comparison to the other studies.

3.7.2. Anxiety

Anxiety was measured in five studies (Burckhardt et al., 2017; Van der Gucht et al., 2017; Livheim et al. (b) 2015; Smith et al., 2020; Pahnke et al., 2014). The only statistically significant score was in Smith et al. (2020), with a medium to large effect size. In Livheim et al. (b)

(2015), the outcome was marginally significant with a large effect size. However, both of these studies scored as having relatively poor methodological rigour in comparison to the other studies. Smith et al. (2020) was the only study in this review with no comparison condition, therefore caution must be used in interpreting these results. A methodological review of studies including psychological treatments showed that the 'pre-post-test' design consistently overestimates effectiveness by an average of 61% compared to studies with a control group (Wilson & Lipsey, 2001). Several confounding variables could be the cause for this including regression to the mean, effects of testing and increases in the maturity and experience of participants over time (Marsden & Torgerson, 2012).

3.7.3. Stress

Stress was measured in four of the studies (Burckhardt et al., 2017; Livheim et al. (b) 2015; Pahnke et al., 2014; Puolakanaho et al., 2019). In Livheim et al. (b) (2015) a statistically significant outcome on the Perceived Stress Scale (PSS) with a large effect size was found. However, this study also measured stress through use of the Depression, Anxiety and Stress Scale (DASS-21) and did not find a significant result. It is important to note that the DASS-21 has not been validated for use in youth samples, and several studies have reported that the three-factor structure of the DASS-21 is invalid when used with an adolescent population (Szabó, 2010; Moore et al., 2017).

Livheim et al., (b) (2015) has relatively low overall methodological rigour in comparison to the other studies, therefore any significant results should be interpreted with caution regardless of the outcome measures used.

A significant outcome for stress was also found in Pahnke et al. (2014), with large effect sizes across both self-ratings and teacher-ratings. This study was rated as having 'average' methodological quality in relation to the other studies included in the review. Puolakanaho et al. (2019) found a significant effect for stress with a low effect size in the 'per-protocol analysis' which included only those participants who completed treatment, but not in the intention-to-treat analysis which included data from all participants.

3.7.4. Overall measures of mental health

Several overall measures of mental health were included across four studies with outcomes termed 'emotional problems', 'psychological health' and 'mental health symptoms' dependent on the measure used (Van der Gucht et al., 2017; Livheim et al., (b) 2015; Pahnke et al., 2014; Takahashi et al., 2020). These measures were all demonstrated to have good reliability and validity within youth samples. All findings were non-significant with low effect sizes, with the exception of Pahnke et al. (2014) where a non-significant effect yet large effect size for 'emotional problems' was found.

3.7.5. Other outcomes

There were several other outcomes measured less frequently across the nine studies. Burckhardt et al. (2017) measured wellbeing, and findings were non-significant and the effect size small. Psychological capital was measured in Fang and Ding (2020b) and they found a significant outcome with a large effect size. Psychological capital has been defined as an individual's positive psychological state of development, and is characterized by self-efficacy, optimism, perseverance towards goals and resilience (Luthans et al., 2006). Anger was measured in Pahnke et al. (2014), and a significant outcome was found, with a large effect size.

3.7.6. Follow up

Three studies included a follow up data point (Burckhardt et al., 2017; Pahnke et al., 2014; Van der Gucht et al., 2017). Only Pahnke et al. (2014) found significant results at follow up for both stress and anger. In Burckhardt et al. (2017), there were medium to large effect sizes at follow up for both stress and anxiety despite a non-significant

Table 5

Process measures.

Study	Outcome variable	Outcome measure	P	Effect sizes
Fang and Ding, (2020b)	Psychological Inflexibility	Acceptance and Action Questionnaire (AAQ-II)	.47	BG pre-post ES = small-medium (d = 0.39)
Van der Gucht (2017)	Psychological Inflexibility	Avoidance and Fusion Questionnaire (AFQ-Y)	NS	BG pre-post (d = -0.02) and pre- follow up (d = 0.15) ES = small
Livheim et al., (a) (2015)	Psychological Inflexibility	Avoidance and Fusion Questionnaire (AFQ-Y)	.021*	BG pre-post ES = medium to large (0.73)
Livheim et al., (b) (2015)	Psychological Inflexibility	Avoidance and Fusion Questionnaire (AFQ-Y)	0.56	ES not reported
	Mindful awareness	Mindful Attention Awareness Scale (MAAS)	.067	BG pre-post ES = medium to large (d = 0.75)
Smith et al. (2020)	Psychological Inflexibility	Avoidance and Fusion Questionnaire (AFQ-Y)	.022**	WG pre-post ES = small to medium (d = 0.38)
Takahashi et al. (2020)	Clarification of values and commitment	Value of Young Age Scale (VOYAGE)	1.00	ES not reported
	Continuation of avoidance	Value of Young Age Scale (VOYAGE)	.036*	ES not reported

Note: *statistically significant group by time interaction ($p < .05$), **statistically significant effect of time (within-group design), BG = between-groups, WG = within-group, ES = effect size.

finding.

3.7.7. ACT process measures

Across the five studies that used measures of psychological flexibility, three studies found significant results (Livheim et al., (a) 2015; Smith et al., 2020; Takahashi et al., 2020). Two of these studies (Livheim et al. (a) 2015; Smith et al., 2020), used the same outcome measure, the Avoidance and Fusion Questionnaire for Youth (AFQ-Y). In Livheim et al. (a) (2015), a large effect size was found, however in Smith et al. (2020) a small to medium effect size was found. No control group was used in Smith et al. (2020), and this study was deemed to have below average methodological quality on the POMRF. Takahashi et al. (2020) found a significant result for the 'continuation of avoidance' subscale on the VOYAGE, but not the 'clarification of value and commitment' subscale. No effect sizes were presented for this study.

3.7.8. Universal vs targeted interventions

One out of the six studies that demonstrated significant results used a universal intervention (Puolakanaho et al., 2019). In this study, a small effect size for stress was found. Across the five targeted interventions that demonstrated significant results, effect sizes were medium to large (Fang and Ding, 2020b; Livheim et al. (a and b), 2015; Smith et al., 2020; Pahnke et al., 2014). The average score for methodological quality for universal interventions was slightly higher at 17.25 (range 13–21), compared to the average score for targeted interventions which was 16.2 (range 14–20).

3.7.9. Impact of sample size

The low sample size reported across the majority of the studies in this review is likely to have had a significant impact upon statistical power. In two of the studies (Burckhardt et al., 2017; Pahnke et al., 2014), large

effect sizes were found despite no statistically significant results. These studies both had a low sample size of 28 and 48 participants, respectively, suggesting that more highly powered research may be needed to obtain significance.

4. Discussion

The aim of the current systematic review was to evaluate the methodological quality and examine the effectiveness of all peer-reviewed literature on ACT interventions based in secondary schools. Nine studies adapted ACT protocols for use in school settings across a range of targeted interventions for mild to moderate mental health difficulties, and universal, non-targeted interventions.

The existing evidence for the effectiveness of ACT-based interventions delivered in school settings on improving mental health and wellbeing is mixed. This review found statistically significant results across six studies for outcomes of depression (Livheim et al. (a), 2015), psychological capital (Fang and Ding, 2020b), stress (Livheim et al.(b), 2015; Pahnke et al., 2014; Puolakanaho et al., 2019) and anxiety (Smith et al., 2020). Three studies found no significant findings across any of the outcomes measured (Burckhardt et al., 2017; Takahashi et al., 2020; Van der Gucht et al., 2017).

With regard to program type, studies that used a targeted intervention performed significantly better than studies that used a universal intervention. This finding aligns with previous research that demonstrates higher levels of effectiveness for targeted compared to universal school-based interventions (Corrieri et al., 2014; Feiss et al., 2019; Werner-Seidler et al., 2017). In this review, there were no significant differences between the universal and targeted interventions with regard to methodological quality, treatment dose, or the experience of the therapists delivering the interventions. The difference in outcomes may reflect a potential difficulty in using outcome measures based on mental health symptoms to quantify the effectiveness of universal interventions. Many students accessing universal interventions may not be presenting with any current difficulties with their mental health and wellbeing and therefore there may be a floor effect to the degree of symptom reduction possible. It may be that future studies examining the effects of universal interventions may benefit from using outcome measures that capture aspects of general wellbeing such as resilience and life satisfaction.

Additionally, it is possible that students accessing a targeted intervention may demonstrate higher levels of therapeutic engagement, as a result of being motivated to address a current problem and therefore finding the intervention content more applicable. It may be useful for future studies to examine levels of engagement amongst students attending targeted compared to universal interventions, for example, by looking at treatment completion rates or program satisfaction.

It is clear from the relatively recent publication dates across all studies that the literature on ACT interventions in schools is in its infancy, with five of the included studies identified as 'pilot' or 'feasibility' research. Studies were heterogeneous with regard to design and outcomes measured; therefore, it is difficult to draw firm conclusions regarding the efficacy of ACT or the moderating influence of program type, program format and delivery. Additionally, as may be expected in a newly developing research area, many methodological limitations were identified. The most common methodological weaknesses across studies were a low sample size, lack of a follow up data point, lack of checks for treatment adherence and therapist competence, and lack of comparison with another active treatment. No studies reported participant ethnicity, which limits this review's assessment of the representativeness of the samples, and potentially the generalisability of these findings more broadly. The average POMRF rating in the current study is 16.7 which is significantly lower than in the most recent meta-analysis on the efficacy of ACT for children by Fang and Ding (2020b) which stated a mean POMRF score of 22.85. This suggests that the methodological quality of school-based studies with ACT currently lags behind the main ACT literature for young people.

The study which received the highest score for methodological quality in this review found no significant results in any measured variables (Van der Gucht et al., 2017). Several reasons for the lack of significant findings were presented by the authors. Van der Gucht et al. (2017) hypothesised that use of a brief treatment program of four sessions was as a potential reason for the lack of significant outcomes. Four sessions were the fewest number of intervention sessions used across all the nine studies. Additionally, use of teachers as the ACT facilitators was suggested as potentially affecting outcomes, with the authors concluding that teachers may need to be supported by a mental health professional. This conclusion is supported by similar research conducted by Wahl et al. (2014), which compared a depression prevention program delivered by either teachers or psychologists and found only the program facilitated by psychologists to have any significant impact on outcomes. Not all of the studies in this review used psychologists as facilitators, however of interest is that Van der Gucht et al. (2017) was the only study that did not use facilitators with a background in psychology or prior mental health training.

It is clear that more stringent checks on therapist competence and adherence are needed across all studies included in this review. No fidelity checks were included across any of the nine studies, impacting upon the internal validity of the studies. Conclusive statements about treatment effectiveness cannot be drawn without consideration of treatment fidelity (Borrelli, 2011; Murphy & Gutman, 2012). Tools such as the recently developed ACT fidelity measure (ACT-FM) (O'Neill et al., 2019) may be a valuable inclusion in further studies.

4.1. Limitations of the current review

In the current systematic review, only peer review articles were included for quality purposes, however in doing so the review failed to account for unpublished 'grey' literature. As this is an emerging and expanding area, reviews of the grey literature may be helpful. Additionally, concerns have been raised around publication bias leading to subsequent inflated effect sizes, therefore examinations of 'grey literature' may help appease these concerns (Strauss et al., 2014).

The use of the term "Acceptance and Commitment Training" is increasing, especially in application with organizations or individuals who do not have a clinical diagnosis, both of which are relevant to a school context. It is possible that the search terminology used in this review may have excluded any studies using this term rather than "therapy". Future reviews may find it beneficial to consider inclusion of this term.

The POMRF rating scale (Öst, 2008) was selected to assess methodological quality across studies, due to its inclusion of elements specific to the evaluation of psychological interventions and its use in a number of published systematic reviews of the ACT literature (Fang and Ding, 2020b; Graham et al., 2016; Kelson et al., 2019; Swain et al., 2013, 2015). However, this measure focuses primarily on clinically diagnosable difficulties, which was not applicable to studies with a preventative focus, and is not in keeping with the transdiagnostic approach used in ACT.

Additionally, due to the self-report nature of the majority of the outcome measures used by the included studies, statement eight on the POMRF ('assessor training') was not relevant to most of the studies, resulting in a low mean score on this item. This is a weakness of using the POMRF as a quality assessment tool, as it is biased towards studies where the outcome measures are administered by trained professionals. There are also no standardised criteria for interpreting POMRF scores, therefore comparisons of the quality of studies to the wider literature is difficult.

5. Conclusions

Despite methodological weaknesses across studies, there is some evidence to show support for the use of ACT as a school-based

intervention. However, more highly powered studies are needed in order to draw any firm conclusions regarding the effectiveness of interventions. More rigorous methodological processes in future research will aid understanding of effects; for example, the extent to which intervention format, therapist competence or adherence to the protocol may be impacting results. As is often the case with emerging intervention approaches, methodological quality can suffer due to lack of funding and resources, as was noted in the earlier days of CBT research (Gaudiano, 2009). However, in spite of these issues, there is a sense of growing momentum in the adolescent ACT literature and the reviewed studies highlight the recent efforts of the ACT community to address the need for evidence-based school interventions, in keeping with UK government guidance ‘Transforming Children and Young People’s Mental Health Provision’ that stipulates a need for more evidence-based

approaches to support mental health in schools (Department for Education, 2017).

Funding

This project was funded by Cardiff University.

Declaration of competing interest

I have no conflicts of interest to declare.

Acknowledgements

None.

Appendix A. Excluded articles with reasons

Excluded Articles

Reference	Reason for exclusion
Burckhardt, R., Manicavasagar, V., Batterham, P. J., & Hadzi-Pavlovic, D. (2016). A randomized controlled trial of strong minds: A school-based mental health program combining acceptance and commitment therapy and positive psychology. <i>Journal of school psychology</i> , 57, 41–52.	ACT intervention combined with positive psychology.
Gómez, M. J., Luciano, C., Pérez-Blarrina, M., Ruiz, F. J., Valdivia-Salas, S., & Gil-Luciano, B. (2014). Brief ACT protocol in at-risk adolescents with conduct disorder and impulsivity. <i>International Journal of Psychology and Psychological Therapy</i> , 14(3), 307–332.	ACT intervention not based in a secondary school. Primary outcome measures related to behaviour rather than mental health and wellbeing.
Habibollahi, A., & Soltanizadeh, M. (2016). Efficacy of Acceptance and Commitment Therapy on Body Dissatisfaction and Fear of Negative Evaluation in Girl adolescents with Body Dysmorphic Disorder. <i>Journal of Mazandaran University of Medical Sciences</i> , 25(134), 278–290.	Full text in Arabic.
Kelly, M. (2015). <i>Effectiveness of Acceptance and Commitment Group Therapy for the Treatment of Perfectionistic and Anxious Adolescent Girls</i> . Palo Alto University.	ACT intervention not based in a secondary school. Dissertation paper not published in a peer reviewed journal.
Brookshier, A. R. (2016). <i>Ameliorating Anxiety in the School Setting Using Acceptance and Commitment Therapy and Mindfulness</i> (Doctoral dissertation, Northern Arizona University).	Dissertation paper not published in a peer reviewed journal.
Mendoza, J. J. (2018). <i>The effectiveness of a modified acceptance and commitment therapy treatment module within a Northern California high school</i> (Dissertation Abstracts International: Section B: The Sciences and Engineering). Requested from Cardiff uni library.	Dissertation paper not published in a peer reviewed journal.
Noormohamadi, S. M., Arefi, M., Afshaini, K., & Kakabarade, K. (2019). The effect of acceptance and commitment therapy on the mental health of students with an emotional breakdown. <i>International Journal of Adolescent Medicine and Health</i> , 1(ahead-of-print).	Full text not available.
Pentcheva-Burns, M. (2015). <i>Effects of Acceptance and Commitment Therapy for diverse symptomatology in immigrant youth</i> . Hofstra University.	Dissertation paper not published in a peer reviewed journal.
Theodore-Oklot, C., Orsillo, S. M., Lee, J. K., & Vernig, P. M. (2014). A pilot of an acceptance- based risk reduction program for relational aggression for adolescents. <i>Journal of Contextual Behavioural Science</i> , 3(2), 109–116.	ACT intervention combined with psycho- education on relational aggression. Participants recruited from secondary school but intervention not school-based.
Azadeh, S. M., Kazemi-Zahrani, H., & Besharat, M. A. (2016). Effectiveness of acceptance and commitment therapy on interpersonal problems and psychological flexibility in female high school students with social anxiety disorder. <i>Global journal of health science</i> , 8(3), 131.	
Bernal-Manrique, K. N., García-Martín, M. B., & Ruiz, F. J. (2020). Effect of acceptance and commitment therapy in improving interpersonal skills in adolescents: A randomized waitlist control trial. <i>Journal of Contextual Behavioral Science</i> , 17, 86–94.	Primary outcome measure not related to mental health or wellbeing.
Murrell, A. R., Steinberg, D. S., Connolly, M. L., Hulsey, T., & Hogan, E. (2015). Acting out to ACTing on: A preliminary investigation in youth with ADHD and co-morbid disorders. <i>Journal of Child and Family Studies</i> , 24(7), 2174–2181.	ACT intervention not based in a secondary school.

Appendix B. Psychotherapy Outcome Study Methodology Rating Form (POMRF)

- Clarity of sample description
 - Poor. Vague description of sample (e.g. only mentioned whether patients were diagnosed with the disorder).
 - Fair. Fair description of sample (e.g. mentioned inclusion/exclusion criteria, demographics, etc.).
 - Good. Good description of sample (e.g. mentioned inclusion/exclusion criteria, demographics, and the prevalence of comorbid disorders).
- Severity/chronicity of the disorder
 - Poor. Severity/chronicity was not reported and/or subsyndromal patients were included in the sample.
 - Fair. All patients met the criteria for the disorder. Sample includes acute (01 yr) and/or low severity.
 - Good. Sample consisted entirely of chronic (41 yr) patients of at least moderate severity.
- Representativeness of the sample
 - Poor. Sample is very different from patients seeking treatment for the disorder (e.g. there are strict exclusion criteria).
 - Fair. Sample is somewhat representative of patients seeking treatment for the disorder (e.g. patients were only excluded if they met criteria for other major disorders).

- 3 Good. Sample is very representative of patients seeking treatment for the disorder (e.g. authors made efforts to ensure representativeness of sample).
4. Reliability of the diagnosis in question
 - 1 Poor. The diagnostic process was not reported, or not assessed with structured interviews by a trained interviewer.
 - 2 Fair. The diagnosis was assessed with structured interview by a trained interviewer.
 - 3 Good. The diagnosis was assessed with structured interview by a trained interviewer and adequate inter-rater reliability was demonstrated (e.g. kappa coefficient).
5. Specificity of outcome measures
 - 0 Poor. Very broad outcome measures, not specific to the disorder (e.g. SCL-90R total score).
 - 1 Fair. Moderately specific outcome measures.
 - 2 Good. Specific outcome measures, such as a measure for each symptom cluster.
6. Reliability and validity of outcome measures
 - 1 Poor. Measures have unknown psychometric properties, or properties that fail to meet current standards of acceptability.
 - 2 Fair. Some, but not all measures have known or adequate psychometric properties.
 - 3 Good. All measures have good psychometric properties. The outcome measures are the best available for the authors' purpose.
7. Use of blind evaluators
 - 1 Poor. Blind assessor was not used (e.g. assessor was the therapist, assessor was not blind to treatment condition, or the authors do not specify).
 - 2 Fair. Blind assessor was used, but no checks were used to assess the blind.
 - 3 Good. Blind assessor was used in correct fashion. Checks were used to assess whether the assessor was aware of treatment condition.
8. Assessor training
 - 1 Poor. Assessor training and accuracy are not specified, or are unacceptable.
 - 2 Fair. Minimum criterion for assessor training is specified (e.g. assessor has had specific training in the use of the outcome measure), but accuracy is not monitored or reported.
 - 3 Good. Minimum criterion of assessor training is specified. Inter-rater reliability was checked, and/or assessment procedures were calibrated during the study to prevent evaluator drift.
9. Assignment to treatment
 - 1 Poor. Biased assignment, e.g. patients selected their own therapy or were assigned in another non-random fashion, or there is only one group.
 - 2 Fair. Random or stratified assignment. There may be some systematic bias but not enough to pose a serious threat to internal validity. There may be therapist by treatment confounds. N may be too small to protect against bias.
 - 3 Good. Random or stratified assignment, and patients are randomly assigned to therapists within condition. When theoretically different treatments are used, each treatment is provided by a large enough number of different therapists. N is large enough to protect against bias.
10. Design
 - 1 Poor. Active treatment vs. WLC, or briefly described TAU.
 - 2 Fair. Active treatment vs. TAU with good description, or placebo condition.
 - 3 Good. Active treatment vs. another previously empirically documented active treatment.
11. Power analysis
 - 0 Poor. No power analysis was made prior to the initiation of the study.
 - 1 Fair. A power analysis based on an estimated effect size was used.
 - 2 Good. A data-informed power analysis was made and the sample size was decided accordingly.
12. Assessment points
 - 0 Poor. Only pre- and post-treatment, or pre- and follow-up.
 - 1 Fair. Pre-, post-, and follow-up <1 year.
 - 2 Good. Pre-, post-, and follow-up >1 year.
13. Manualized, replicable, specific treatment programs
 - 1 Poor. Description of treatment procedure is unclear, and treatment is not based on a publicly available, detailed treatment manual. Patients may be receiving multiple forms of treatment at once in an uncontrolled manner.
 - 2 Fair. Treatment is not designed for the disorder, or description of the treatment is generally clear and based on a publicly available, detailed treatment manual, but there are some ambiguities about the procedure. Patients may have received additional forms of treatment, but this is balanced between groups or otherwise controlled.
 - 3 Good. Treatment is designed for the disorder. A detailed treatment manual is available, and/or treatment is explained in sufficient detail for replication. No ambiguities about the treatment procedure. Patients receive only the treatment in question.
14. Number of therapists
 - 0 Poor. Only one therapist, i.e., complete confounding between therapy and therapist.
 - 1 Fair. At least two therapists, but the effect of therapist on outcome is not analyzed.
 - 2 Good. Three, or more therapists, and the effect of therapist on outcome is analyzed.
15. Therapist training/experience
 - 0 Poor. Very limited clinical experience of the treatment and/or disorder (e.g. students).
 - 1 Fair. Some clinical experience of the treatment and/or disorder.
 - 2 Good. Long clinical experience of the treatment and the disorder (e.g. practicing therapists).
16. Checks for treatment adherence
 - 1 Poor. No checks were made to assure that the intervention was consistent with protocol.
 - 2 Fair. Some checks were made (e.g. assessed a proportion of therapy tapes).
 - 3 Good. Frequent checks were made (e.g. weekly supervision of each session using a detailed rating form).
17. Checks for therapist competence
 - 0 Poor. No checks were made to assure that the intervention was delivered competently.
 - 1 Fair. Some checks were made (e.g. assessed a proportion of therapy tapes).

- 2 Good. Frequent checks were made (e.g. weekly supervision of each session using a detailed rating form).
18. Control of concomitant treatments (e.g. medications)
 - 1 Poor. No attempt to control for concomitant treatments, or no information about concomitant treatments provided. Patients may have been receiving other forms of treatment in addition to the study treatment.
 - 2 Fair. Asked patients to keep medications stable and/or to discontinue other psychological therapies during the treatment.
 - 3 Good. Ensured that patients did not receive any other treatments (medical or psychological) during the study.
19. Handling of attrition
 - 1 Poor. Proportions of attrition are not described, or described but no dropout analysis is performed.
 - 2 Fair. Proportions of attrition are described, and dropout analysis or intent-to-treat analysis is performed.
 - 3 Good. No attrition, or proportions of attrition are described, dropout analysis is performed, and results are presented as intent-to-treat analysis.
20. Statistical analyses and presentation of results
 - 0 Poor. Inadequate statistical methods are used and/or data are not fully presented.
 - 1 Fair. Adequate statistical methods are used but data are not fully presented.
 - 2 Good. Adequate statistical methods are used and data are presented with M and SD.
21. Clinical significance
 - 1 Poor. No presentation of clinical significance was done.
 - 2 Fair. An arbitrary criterion for clinical significance was used and the conditions were compared regarding percent clinically improved.
 - 3 Good. Jacobson's criteria for clinical significance were used and presented for a selection (or all) of the outcome measures, and conditions were compared regarding percent clinically improved.
22. Equality of therapy hours (for non-WLC designs only)
 - 1 Poor. Conditions differ markedly (>20% difference in therapy hours).
 - 2 Fair. Conditions differ somewhat (10–19% difference in therapy hours).
 - 2 Good. Conditions do not differ (<10% difference in therapy hours).

References

- Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment*, 10(2), 176. <https://doi.org/10.1037/1040-3590.10.2.176>
- Bailey, A., Ciarrochi, J., & Hayes, L. (2012). *Get out of your mind and into your life for teens: A guide to living an extraordinary life*. New Harbinger Publications.
- Barrett, P. M., & Turner, C. M. (2004). Prevention of childhood anxiety and depression. In P. M. Barrett, & T. H. Ollendick (Eds.), *Handbook of interventions that work with children and adolescents: Prevention and treatment* (pp. 429–474). John Wiley & Sons Ltd. <https://doi.org/10.1002/9780470753385.ch18>
- Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., Lu, C., ... Lancet Early Childhood Development Series Steering Committee. (2017). Early childhood development coming of age: Science through the life course. *The Lancet*, 389(10064), 77–90. [https://doi.org/10.1016/S0140-6736\(16\)3189-7](https://doi.org/10.1016/S0140-6736(16)3189-7)
- Borrelli, B. (2011). The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *Journal of Public Health Dentistry*, 71, S52–S63. <https://doi.org/10.1111/j.1752-7325.2011.00233.x>
- Brown, K. W., West, A. M., Loverich, T. M., & Biegel, G. M. (2011). Assessing adolescent mindfulness: Validation of an adapted mindful attention awareness scale in adolescent normative and psychiatric populations. *Psychological Assessment*, 23(4), 1023. <https://doi.org/10.1037/a0021338>
- Burckhardt, R., Manicavasagar, V., Batterham, P. J., Hadzi-Pavlovic, D., & Shand, F. (2017). Acceptance and commitment therapy universal prevention program for adolescents: A feasibility study. *Child and Adolescent Psychiatry and Mental Health*, 11(1), 1–10. <https://doi.org/10.1186/s13034-017-0164-5>
- Caldwell, D. M., Davies, S. R., Hetrick, S. E., Palmer, J. C., Caro, P., López-López, J. A., ... Welton, N. J. (2019). School-based interventions to prevent anxiety and depression in children and young people: A systematic review and network meta-analysis. *The Lancet Psychiatry*, 6(12), 1011–1020. [https://doi.org/10.1016/S2215-0366\(19\)30403-1](https://doi.org/10.1016/S2215-0366(19)30403-1)
- Cao, J., Ji, Y., & Zhu, Z. H. (2013). Reliability and validity of the Chinese version of the Acceptance and Action Questionnaire (AAQ-II) in college students. *Chinese Mental Health Journal*, 27(11), 873–877.
- Carlson, L. E., & Brown, K. W. (2005). Validation of the mindful attention awareness scale in a cancer population. *Journal of Psychosomatic Research*, 58(1), 29–33. <https://doi.org/10.1016/j.jpsychores.2004.04.366>
- Casey, B. J., Getz, S., & Galvan, A. (2008). The adolescent brain. *Developmental Review*, 28(1), 62–77. <https://doi.org/10.1016/j.dr.2007.08.003>
- Cheng, J., & Sun, Y. H. (2015). Depression and anxiety among left-behind children in China: A systematic review. *Child: Care, Health and Development*, 41(4), 515–523. <https://doi.org/10.1111/cch.12221>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Psychology Press.
- Collins, K. A., & Dozois, D. J. (2008). What are the active ingredients in preventative interventions for depression? *Clinical Psychology: Science and Practice*, 15(4), 313–330. <https://doi.org/10.1111/j.1468-2850.2008.00143.x>
- Corrieri, S., Heider, D., Conrad, I., Blume, A., König, H. H., & Riedel-Heller, S. G. (2014). School-based prevention programs for depression and anxiety in adolescence: A systematic review. *Health Promotion International*, 29(3), 427–441. <https://doi.org/10.1093/heapro/dat001>
- De Groot, F. (2005). Primaire preventie van psychische problemen op basis van Acceptance & Commitment Therapy. *Psychopraxis*, 7(2), 52–56.
- Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13
- Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D. W., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research*, 97(2), 143–156. <https://doi.org/10.1007/s11205-009-9493-y>
- Domitrovich, C. E., Bradshaw, C. P., Greenberg, M. T., Embry, D., Poduska, J. M., & Ialongo, N. S. (2010). Integrated models of school-based prevention: Logic and theory. *Psychology in the Schools*, 47(1), 71–88. <https://doi.org/10.1002/pits.20452>
- Dray, J., Bowman, J., Campbell, E., Freund, M., Wolfenden, L., Hodder, R. K., ... Wiggers, J. (2017). Systematic review of universal resilience-focused interventions targeting child and adolescent mental health in the school setting. *Journal of the American Academy of Child & Adolescent Psychiatry*, 56(10), 813–824. <https://doi.org/10.1016/j.jaac.2017.07.780>
- Duan, W., & Xie, D. (2019). Measuring adolescent flourishing: Psychometric properties of flourishing scale in a sample of Chinese adolescents. *Journal of Psychoeducational Assessment*, 37(1), 131–135. <https://doi.org/10.1177/0734282916655504>
- Elo, A. L., Leppänen, A., & Jahkola, A. (2003). Validity of a single-item measure of stress symptoms. *Scandinavian Journal of Work, Environment & Health*, 444–451. <https://doi.org/10.5271/sjweh.752>
- Fang, S., & Ding, D. (2020a). A meta-analysis of the efficacy of acceptance and commitment therapy for children. *Journal of Contextual Behavioral Science*, 15, 225–234. <https://doi.org/10.1016/j.jcbs.2020.01.007>
- Fang, S., & Ding, D. (2020b). The efficacy of group-based acceptance and commitment therapy on psychological capital and school engagement: A pilot study among Chinese adolescents. *Journal of Contextual Behavioral Science*, 16, 134–143. <https://doi.org/10.1016/j.jcbs.2020.04.005>
- Feiss, R., Dolinger, S. B., Merritt, M., Reiche, E., Martin, K., Yanes, J. A., ... Pangelinan, M. (2019). A systematic review and meta-analysis of school-based stress, anxiety, and depression prevention programs for adolescents. *Journal of Youth and Adolescence*, 48(9), 1668–1685. <https://doi.org/10.1007/s10964-019-01085-0>
- Forman, E. M., & Herbert, J. D. (2009). New directions in cognitive behavior therapy: Acceptance-based therapies. In W. T. O'Donohue, & J. E. Fisher (Eds.), *General principles and empirically supported techniques of cognitive behavior therapy* (pp. 77–101). John Wiley & Sons, Inc.
- Gao, F., Luo, N., Thumboo, J., Fones, C., Li, S. C., & Cheung, Y. B. (2004). Does the 12-item General Health Questionnaire contain multiple factors and do we need them? *Health and Quality of Life Outcomes*, 2(1), 1–7. <https://doi.org/10.1186/1477-7525-2-63>
- Gillard, D., Flaxman, P., & Hooper, N. (2018). Acceptance and commitment therapy: Applications for educational psychologists within schools. *Educational Psychology in Practice*, 34(3), 272–281. <https://doi.org/10.1080/02667363.2018.1446911>
- Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(11), 1337–1345. <https://doi.org/10.1097/00004583-200111000-00015>
- Graham, C. D., Gouick, J., Krahe, C., & Gillanders, D. (2016). A systematic review of the use of Acceptance and Commitment Therapy (ACT) in chronic disease and long-term

- conditions. *Clinical Psychology Review*, 46, 46–58. <https://doi.org/10.1016/j.cpr.2016.04.009>
- Greco, L. A., Blackledge, J. T., Coyne, L. W., & Ehrenreich, J. (2005). Integrating Acceptance and mindfulness into treatments for child and adolescent anxiety disorders. In *Acceptance and mindfulness-based approaches to anxiety* (pp. 301–322). Boston, MA: Springer. https://doi.org/10.1007/0-387-25989-9_12
- Greco, L. A., Lambert, W., & Baer, R. A. (2008). Psychological inflexibility in childhood and adolescence: Development and evaluation of the avoidance and fusion questionnaire for youth. *Psychological Assessment*, 20(2), 93. <https://doi.org/10.1037/1040-3590.20.2.93>
- Groden, J., Diller, A., Bausman, M., Velicer, W., Norman, G., & Cautela, J. (2001). The development of a stress survey schedule for persons with autism and other developmental disabilities. *Journal Of Autism And Developmental Disorders*, 31(2), 207–217. <https://doi.org/10.1023/A:1010755300436>
- Halliburton, A. E., & Cooper, L. D. (2015). Applications and adaptations of acceptance and commitment therapy (ACT) for adolescents. *Journal of Contextual Behavioral Science*, 4(1), 1–11. <https://doi.org/10.1016/j.jcbs.2015.01.002>
- Harris, R. (2019). *ACT made simple: An easy-to-read primer on acceptance and commitment therapy*. New Harbinger Publications.
- Hayes, S. C. (2004). Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behavior Therapy*, 35(4), 639–665. [https://doi.org/10.1016/S0005-7894\(04\)80013-3](https://doi.org/10.1016/S0005-7894(04)80013-3)
- Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and commitment therapy: Model, processes and outcomes. *Behaviour Research and Therapy*, 44(1), 1–25. <https://doi.org/10.1016/j.brat.2005.06.006>
- Hayes, L. L., & Rowse, J. (2008). Acceptance and commitment therapy: Experiential adolescent group [online]. Available at: Association for Contextual Behavioral Science website: <http://contextualpsychology.org/treatment/protocols>
- Hofmann, S. G., & Asmundson, G. J. (2008). Acceptance and mindfulness-based therapy: New wave or old hat? *Clinical Psychology Review*, 28(1), 1–16. <https://doi.org/10.1016/j.cpr.2007.09.003>
- Horowitz, J. L., Garber, J., Ciesla, J. A., Young, J. F., & Mufson, L. (2007). Prevention of depressive symptoms in adolescents: A randomized trial of cognitive-behavioral and interpersonal prevention programs. *Journal of Consulting and Clinical Psychology*, 75 (5), 693. <https://doi.org/10.1037/0022-006X.75.5.693>
- Ishizu, K., Otsuki, T., Shimoda, Y., & Nagata, M. (2016). Development of the Value of young age scale (1): Examination of its reliability and validity [poster presentation]. In *The 58th annual conference of the Japanese association of educational psychology*. October 8–10 Japan: Kagawa.
- Kelson, J., Rollin, A., Ridout, B., & Campbell, A. (2019). Internet-delivered acceptance and commitment therapy for anxiety treatment: Systematic review. *Journal of Medical Internet Research*, 21(1), Article e12530. <https://doi.org/10.2196/12530>
- Kessler, R. C., Angermeyer, M., Anthony, J. C., et al. (2007). Lifetime prevalence and age-of-onset distributions of mental disorders in the world health organization's world mental health survey initiative. *World Psychiatry*, 6, 168–176.
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry*, 62(6), 593–602. <https://doi.org/10.1001/archpsyc.62.6.593>
- Kieling, C., Baker-Henningham, H., Belfer, M., Conti, G., Ertem, I., Omigbodun, O., ... Rahman, A. (2011). Child and adolescent mental health worldwide: Evidence for action. *The Lancet*, 378(9801), 1515–1525. [https://doi.org/10.1016/S0140-6736\(11\)60827-1](https://doi.org/10.1016/S0140-6736(11)60827-1)
- Liu, X., Fan, X. H., & Shen, J. L. (2007). Relationship between social support and problem behaviors of the left-home-kids in junior middle school. *Psychological Development and Education*, 23(3), 98–102.
- Liu, X., & Zhao, J. (2016). Chinese migrant adolescents' perceived discrimination and psychological well-being: The moderating roles of group identity and the type of school. *PLoS One*, 11(1), Article e0146559. <https://doi.org/10.1371/journal.pone.0146559>
- Livheim, F. (2004). *Acceptance and commitment therapy in school settings: To cope with stress, a randomized controlled trial*. Sweden: Uppsala University. Unpublished master's thesis.
- Livheim, F., Hayes, L., Ghaderi, A., Magnusdottir, T., Högföldt, A., Rowse, J., ... Tengström, A. (2015). The effectiveness of acceptance and commitment therapy for adolescent mental health: Swedish and Australian pilot outcomes. *Journal of Child and Family Studies*, 24(4), 1016–1030. <https://doi.org/10.1007/s10826-014-9912-9>
- Luoma, J. B., Hayes, S. C., & Walser, R. D. (2007). *Learning ACT: An acceptance & commitment therapy skills-training manual for therapists*. New Harbinger Publications.
- Luthans, F., Vogelgesang, G. R., & Lester, P. B. (2006). Developing the psychological capital of resiliency. *Human Resource Development Review*, 5(1), 25–44. <https://doi.org/10.1177/1534484305285335>
- Mackenzie, K., & Williams, C. (2018). Universal, school-based interventions to promote mental and emotional well-being: What is being done in the UK and does it work? A systematic review. *BMJ Open*, 8(9), Article e022560. <https://doi.org/10.1136/bmjopen-2018-022560>
- Marks, R. (Ed.). (2012). *Health literacy and school-based health education*. Emerald Group Publishing.
- Marsden, E., & Torgerson, C. J. (2012). Single group, pre-and post-test research designs: Some methodological concerns. *Oxford Review of Education*, 38(5), 583–616. <https://doi.org/10.1080/03054985.2012.731208>
- Masia-Warner, C., Nangle, D. W., & Hansen, D. J. (2006). Bringing evidence-based child mental health services to the schools: General issues and specific populations. *Education & Treatment of Children*, 29(2), 165–172.
- Mokkink, L. B., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C., & Terwee, C. B. (2019). COSMIN Study Design checklist for Patient-reported outcome measurement instruments. https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf. (Accessed December 2020). <https://doi.org/10.1136/gutjnl-2020-320729>
- Moore, S. A., Dowdy, E., & Furlong, M. J. (2017). Using the depression, anxiety, stress scales–21 with US adolescents: An alternate models analysis. *Journal of Psychoeducational Assessment*, 35(6), 581–598. <https://doi.org/10.1177/07342829166651537>
- Murphy, S. L., & Gutman, S. A. (2012). Intervention fidelity: A necessary aspect of intervention effectiveness studies. *American Journal of Occupational Therapy*, 66(4), 387–388. <https://doi.org/10.5014/ajot.2010.005405>
- National Society for the Protection of Cruelty to Children. (2017). School referrals for mental health treatment rise by over a third [online]. Available at: www.nspcc.org.uk/what-wedo/news-opinion/one-third-increase-in-school-referrals-for-mental-health-treatment/. (Accessed September 2020) Accessed.
- Neil, A. L., & Christensen, H. (2009). Efficacy and effectiveness of school-based prevention and early intervention programs for anxiety. *Clinical Psychology Review*, 29(3), 208–215. <https://doi.org/10.1016/j.cpr.2009.01.002>
- Neto, F. (1993). The satisfaction with life scale: Psychometrics properties in an adolescent sample. *Journal of Youth and Adolescence*, 22(2), 125–134. <https://doi.org/10.1007/BF01536648>
- O'Neill, L., Latchford, G., McCracken, L. M., & Graham, C. D. (2019). The development of the acceptance and commitment therapy fidelity measure (ACT-FM): A delphi study and field test. *Journal of Contextual Behavioral Science*, 14, 111–118. <https://doi.org/10.1016/j.jcbs.2019.08.008>
- Oleson, J. J., Brown, G. D., & McCreery, R. (2019). Essential statistical concepts for research in speech, language, and hearing sciences. *Journal of Speech, Language, and Hearing Research*, 62(3), 489–497. https://doi.org/10.1044/2018_JSLHR-S-ASTM-18-0239
- Öst, L. G. (2008). Efficacy of the third wave of behavioral therapies: A systematic review and meta-analysis. *Behaviour Research and Therapy*, 46(3), 296–321. <https://doi.org/10.1016/j.brat.2007.12.005>
- Pahnke, J., Lundgren, T., Hursti, T., & Hirvikoski, T. (2014). Outcomes of an acceptance and commitment therapy-based skills training group for students with high-functioning autism spectrum disorder: A quasi-experimental pilot study. *Autism*, 18 (8), 953–964. <https://doi.org/10.1177/1362361313501091>
- Patel, V., Flisher, A. J., Hetrick, S., & McGorry, P. (2007). Mental health of young people: A global public-health challenge. *The Lancet*, 369(9569), 1302–1313. [https://doi.org/10.1016/S0140-6736\(07\)60368-7](https://doi.org/10.1016/S0140-6736(07)60368-7)
- Pitchforth, J., Fahy, K., Ford, T., Wolpert, M., Viner, R. M., & Hargreaves, D. S. (2019). Mental health and well-being trends among children and young people in the UK, 1995–2014: Analysis of repeated cross-sectional national health surveys. *Psychological Medicine*, 49(8), 1275–1285. <https://doi.org/10.1017/S0033291718001757>
- Puolakanaho, A., Lappalainen, R., Lappalainen, P., Muotka, J. S., Hirvonen, R., Eklund, K. M., ... Kiuru, N. (2019). Reducing stress and enhancing academic buoyancy among adolescents using a brief web-based program based on acceptance and commitment therapy: A randomized controlled trial. *Journal of Youth and Adolescence*, 48(2), 287–305. <https://doi.org/10.1007/s10964-018-0973-8>
- Smith, K. J., Oxman, L. N., & Hayes, L. (2020). ACT for adolescents: Impact of a pilot group intervention on psychological wellbeing. *Journal of Psychologists and Counsellors in Schools*, 30(2), 248–254. <https://doi.org/10.1017/jgc.2019.33>
- Stephan, S. H., Weist, M., Kataoka, S., Adelsheim, S., & Mills, C. (2007). Transformation of children's mental health services: The role of school mental health. *Psychiatric Services*, 58(10), 1330–1338. <https://doi.org/10.1176/ps.2007.58.10.1330>
- Strauss, C., Cavanagh, K., Oliver, A., & Pettman, D. (2014). Mindfulness-based interventions for people diagnosed with a current episode of an anxiety or depressive disorder: A meta-analysis of randomised controlled trials. *PLoS One*, 9(4), Article e96110. <https://doi.org/10.1371/journal.pone.0096110>
- Swain, J., Hancock, K., Dixon, A., & Bowman, J. (2015). Acceptance and commitment therapy for children: A systematic review of intervention studies. *Journal of Contextual Behavioral Science*, 4(2), 73–85. <https://doi.org/10.1016/j.jcbs.2015.02.001>
- Swain, J., Hancock, K., Hainsworth, C., & Bowman, J. (2013). Acceptance and commitment therapy in the treatment of anxiety: A systematic review. *Clinical Psychology Review*, 33(8), 965–978. <https://doi.org/10.1016/j.cpr.2013.07.002>
- Szabó, M. (2010). The short version of the Depression Anxiety Stress Scales (DASS-21): Factor structure in a young adolescent sample. *Journal of Adolescence*, 33(1), 1–8. <https://doi.org/10.1016/j.adolescence.2009.05.014>
- Tait, R. J., French, D. J., & Hulse, G. K. (2003). Validity and psychometric properties of the General Health Questionnaire-12 in young Australian adolescents. *Australian and New Zealand Journal of Psychiatry*, 37(3), 374–381. <https://doi.org/10.1046/j.1440-1614.2003.01133.x>
- Takahashi, F., Ishizu, K., Matsubara, K., Ohtsuki, T., & Shimoda, Y. (2020). Acceptance and commitment therapy as a school-based group intervention for adolescents: An open-label trial. *Journal of Contextual Behavioral Science*, 16, 71–79. <https://doi.org/10.1016/j.jcbs.2020.03.001>
- UK Government. (2017). Transforming children and young people's mental health provision: A green paper [online]. <https://www.gov.uk/government/consultations/transforming-children-and-young-peoples-mental-health-provision-a-green-paper>. (Accessed September 2020).
- Van der Gucht, K., Griffith, J. W., Hellemans, R., Bockstaele, M., Pascal-Claes, F., & Raes, F. (2017). Acceptance and commitment therapy (ACT) for adolescents: Outcomes of a large-sample, school-based, cluster-randomized controlled trial. *Mindfulness*, 8(2), 408–416. <https://doi.org/10.1007/s12671-016-0612-y>
- Verbeke, G., & Molenberghs, G. (2000). *A model for Longitudinal Data. Linear mixed models for longitudinal data* (pp. 19–29). https://doi.org/10.1007/978-0-387-22775-7_3

- Verhulst, F. C., Van der Ende, J., & Koot, H. M. (1997). *Manual for the youth self-report (YSR) [in Dutch]*. Rotterdam: Erasmus University/Department of Child and Adolescent Psychiatry, Sophia Children's Hospital.
- Wahl, M. S., Adelson, J. L., Patak, M. A., Pössel, P., & Hautzinger, M. (2014). Teachers or psychologists: Who should facilitate depression prevention programs in schools? *International Journal of Environmental Research and Public Health*, 11(5), 5294–5316. <https://doi.org/10.3390/ijerph110505294>
- Welsh Government. (2020). Curriculum for Wales guidance [online]. Available at: <https://hwb.gov.wales/storage/b44ad45b-ff78-430a-9423-36feb86aaf7e/curriculum-for-wales-guidance.pdf>. (Accessed January 2021) Accessed.
- 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>
- Wilson, D. B., & Lipsey, M. W. (2001). The role of method in treatment effectiveness research: Evidence from meta-analysis. *Psychological Methods*, 6(4), 413. <https://doi.org/10.1037/1082-989X.6.4.413>